# Adaptive semiparametric wavelet estimator and goodness-of-fit test for long-memory linear processes

Jean-Marc Bardet[a] and Hatem Bibi[a]

bardet@univ-paris1.fr, hatem.bibi@malix.univ-paris1.fr

[a] SAMM, Université Paris 1, 90 rue de Tolbiac, 75013 Paris, FRANCE.

**Keywords:** Long range dependence; linear processes; wavelet estimator; semiparametric estimator; adaptive estimator; adaptive goodness-of-fit test.

**MSC Classification:** Primary: 62M07, 62M09 Secondary 62M10; 62M15; 60F05.

**Abstract**

This paper is first devoted to the study of an adaptive wavelet-based estimator of the long-memory parameter for linear processes in a general semiparametric frame and as such is an extension of the previous contribution of Bardet *et al.* (2008) which only concerned Gaussian processes. Moreover, the definition of the long-memory parameter estimator is modified and asymptotic results are improved even in the Gaussian case. Finally an adaptive goodness-of-fit test is also built and easy to be employed: it is a chi-square type test. Simulations confirm the interesting properties of consistency and robustness of the adaptive estimator and test.

# 1 Introduction

Presently, long memory processes have become a widely-studied subject area and find frequent applications (see for instance Dhoukhan et al, 2003)

The best known long-memory stationary time series are the fractional Gaussian noises (fGn) with Hurst parameter $H$ and FARIMA$(p, d, q)$ processes. For both these time series, the spectral density $f$ in 0 follows power law: $f(\lambda) \sim C\,\lambda^{-2d}$ where $H = d + 1/2$ in the case of the fGn. This behavior of the spectral density generally defines a stationary long-memory (or long-range-dependent) process even if it needs the presence of a second order moment.

In this paper, we consider the general case of a linear process with a memory parameter $d$ and propose an adaptive wavelet-based estimator of this parameter, i.e. for $d < 1/2$ and $d' > 0$, we use the the following semiparametric framework for the present study:

**Assumption A$(d, d')$:** *$X = (X_t)_{t\in\mathbb{Z}}$ is a zero mean stationary linear process, i.e.*

$$X_t = \sum_{s\in\mathbb{Z}} \alpha(t-s)\xi_s, \quad t \in \mathbb{Z}, \quad \textit{where}$$

- *$(\xi_s)_{s\in\mathbb{Z}}$ is a sequence of independent identically distributed random variables such that the distribution of $\xi_0$ is symmetric, i.e. $\forall M \in \mathbb{R}$, $\Pr(\xi_0 > M) = \Pr(\xi_0 < -M)$, $\mathbb{E}\xi_0 = 0$, $Var\,\xi_0 = 1$ and $\mu_4 := \mathbb{E}\xi_0^4 < \infty$;*

- *$(\alpha(t))_{t\in\mathbb{Z}}$ is a sequence of real numbers such that there exist $c_d > 0$ and $c_{d'} \in \mathbb{R}$ satisfying*

$$|\widehat{\alpha}(\lambda)|^2 = \frac{1}{\lambda^{2d}}\left(c_d + c_{d'}|\lambda|^{d'}(1+\varepsilon(\lambda))\right) \quad \textit{for any} \quad \lambda \in [-\pi, \pi], \tag{1}$$

*where $\widehat{\alpha}(\lambda) := \frac{1}{2\pi}\sum_{k\in\mathbb{Z}} \alpha(k)e^{-ik\lambda}$ for $\lambda \in [-\pi, \pi]$ and with $\varepsilon(\lambda) \to 0$ $(\lambda \to 0)$.*

Thus, if $X$ satisfies Assumption A$(d, d')$, the spectral density $f$ of $X$ is such that

$$f(\lambda) = 2\pi\,|\widehat{\alpha}(\lambda)|^2 = \frac{2\pi}{\lambda^{2d}}\left(c_d + c_{d'}|\lambda|^{d'}(1+\varepsilon(\lambda))\right) \quad \text{for any} \quad \lambda \in [-\pi, \pi], \tag{2}$$

with $\varepsilon(\lambda) \to 0$ $(\lambda \to 0)$. Thus, if $d \in (0, 1/2)$, the process $X$ is a long-memory process, and if $d \leq 0$, it is a short-memory process (see Doukhan *et al.*, 2003).

After preliminary studies devoted to self-similar processes Abry *et al.* (1998), were the first to propose the use of a wavelet-based estimator for estimating $d$ by computing the log-log regression slope for different scales of wavelet coefficient sample variances. Bardet *et al.* (2000) provided proofs of the consistency of such an estimator in a Gaussian semiparametric frame. Moulines *et al.* (2007) not only improved these results, they also established a central limit theorem for the estimator of $d$ which they proved rate optimal for the minimax criterion. As to Roueff and Taqqu (2009a). They yielded similar results in a semiparametric frame for linear processes.

All of these studies used a wavelet analysis based on a discrete multi-resolution wavelet transform, which in particular allows to compute the wavelet coefficients with the fast Mallat's algorithm. Their results, however, are inferred from a semiparametric frame such as to (2) and consider the "optimal" scale used for the wavelet analysis which depends on the second order expansion $d'$ to be known although, in fact it is unknown. Two studies present automatic selection method for this "optimal" scale in the Gaussian semiparametric frame. The chi-square test according to Veitch *et al.* (2003) despite convincing numerical results, lacks sufficient evidence of consistency . Whereas, Bardet *et al.* (2008) proved the consistency of a procedure for choosing optimal scales based on the detection of the "most linear part" of the log-variogram graph. They consider that the "mother" wavelet is not necessarily associated with a multi-resolution analysis: although the computation cost is more important, it offers a larger wavelet function choice and scales are not limited to the power of 2.

The present paper is an extension of a previous study of Bardet *et al.* (2008). Improvements concern three following central issues:

1. The semiparametric Gaussian framework of Bardet *et al.* (2008) is extended to the semiparametric framework Assumption $\mathrm{A}(d, d')$ for linear processes. The same automatic procedure of the optimal scale selection allowed us to obtain adaptive estimators.

2. As in Bardet *et al.* (2008), the "mother" wavelet is not necessarily associated with a discrete multi-resolution transform. We also slightly modified the definition of the wavelet coefficient sample variance ("variogram"). The result of both these changes is a multidimensional central limit theorem satisfied by the logarithms of variograms with an extremely simple asymptotic covariance matrix (see (10)) depending only on $d$ and the Fourier transform of the wavelet function. Hence it is easy to compute an adaptive pseudo-generalized least square estimator (PGLSE) of $d$, satisfying a CLT with an asymptotic variance which is smaller than both the the adaptive (Bardet *et al.* (2008))and the non-adaptive (Roueff and Taqqu (2009)) ordinary least square estimator of $d$ . Simulations confirm the good performance of this PGLSE.

3. Finally, we used this PGLSE to perform an adaptive goodness-of-fit test. It represents a normalized sum

of the squared PGLS-distance between the PGLS-regression line and the points. We proved that this test statistic converges in distribution to a chi-square distribution.the asymptotic covariance matrix being easily approximated, the test is very simple test to compute. When $d > 0$ this test is a long-memory test. Moreover, simulations show that this test provides good properties of consistency under $H_0$ and reasonable properties of robustness under $H_1$.

In the light of these results, the present paper represents a conclusion to the study of Bardet *et al.* (2008). and the adaptive PGLS estimator and test an interesting extension of Roueff and Taqqu (2009).

The present paper is organized into 4 sections as follows.

Assumptions, definitions and a first multidimensional central limit theorem are the subject matter of Section 2.

The construction and consistency of the adaptive PGLS estimator and goodness-of-fit test in dealt with section 3.

In Section 4 features a Monte Carlo simulations-based demonstration of the convergence of the adaptive estimator, followed by a comparaison with efficient semiparametric estimators others than oures and investigations into the consistency and robustness properties of the adaptive goodness-of-fit test. Proofs figure in section 5.

## 2 Central limit theorem for the sample variance of wavelet coefficients

We let $\psi : \mathbb{R} \to \mathbb{R}$ the wavelet function , $k \in \mathbb{N}^*$. We shall consider the following assumption on $\psi$:

**Assumption** $\Psi(k)$**:** $\psi : \mathbb{R} \to \mathbb{R}$ *is such that*

1. *the support of* $\psi$ *is included in* $(0, 1)$*;*
2. $\displaystyle\int_0^1 \psi(t)\, dt = 0$*;*
3. $\psi \in \mathcal{C}^k(\mathbb{R})$.

Straightforward implications of these assumptions are $\psi^{(j)}(0) = \psi^{(j)}(1) = 0$ for any $0 \le j \le k$.

If we define $\widehat{\psi}(u)$ the Fourier transform of $\psi$ when $\psi$ satisfies Assumption $\Psi(k)$, *i.e.*

$$\widehat{\psi}(u) := \int_0^1 \psi(t)\, e^{-iut} dt,$$

Then $\widehat{\psi}(u) \sim C\, u^k$ $(u \to 0)$ with $C$ a real number not independant of $u$ and

$$\sup_{u\in\mathbb{R}} \big|u^k\, \widehat{\psi}(u)\big| \leq \sup_{x\in[0,1]} |\psi^{(k)}(x)|. \tag{3}$$

If $Y = (Y_t)_{t\in\mathbb{R}}$ is a continuous-time process, for $(a,b) \in \mathbb{R}_+^* \times \mathbb{R}$, the "classical" wavelet coefficient $d(a,b)$ of the process $Y$ for the scale $a$ and the shift $b$ is $d(a,b) := \frac{1}{\sqrt{a}} \int_{\mathbb{R}} \psi(\frac{t-b}{a}) Y_t\, dt$. However, since the process $X$ satisfying Assumption $\mathrm{A}(d,d')$ is a discrete-time process, we define the wavelet coefficients of $X$ by

$$e(a,b) := \frac{1}{\sqrt{a}} \sum_{t=1}^{N} X_t \psi(\frac{t-b}{a}) = \sum_{j=1}^{a} \Big(\frac{1}{\sqrt{a}}\, \psi(\frac{j}{a})\Big)\, X_{b+j} \tag{4}$$

for $(a,b) \in \mathbb{N}^* \times \mathbb{Z}$ (this definition of $e(a,b)$ also holds for $a \in \mathbb{R}_+^*$ to avoid the use of $[a]$, the integer part of $a$, we restrict it to $a \in \mathbb{N}^*$).

Let $(X_1, \ldots, X_N)$ be an observed path of $X$, $a \in \mathbb{N}^*$ and $b = 1, \ldots, N-a$. We use the usual convention $y = o(g(x))$ $(x \to \infty)$ when $\lim_{x\to\infty} y/g(x) = 0$,

**Property 1.** *Under Assumption $A(d,d')$ with $d < 1/2$ and $d' > 0$, and if $\psi$ satisfies Assumption $\Psi(k)$ with $k > d' - d + 1/2$, for $a \in \mathbb{N}^*$, then $(e(a,b))_{b\in\mathbb{Z}}$ is a zero mean stationary linear process and*

$$\mathbb{E}(e^2(a,0)) \quad = \quad 2\pi\, c_d \Big(K_{(\psi,2d)}\, a^{2d} + \frac{c_{d'}}{c_d} K_{(\psi,2d-d')}\, a^{2d-d'}\Big) + o\big(a^{2d-d'}\big) \quad \text{when} \quad a \to \infty, \tag{5}$$

$$\text{with} \qquad K_{(\psi,\alpha)} := \int_{-\infty}^{\infty} |\widehat{\psi}(u)|^2\, |u|^{-\alpha} du > 0 \quad \text{for all } \alpha < 1. \tag{6}$$

Refer to section 5 for the details results of all demonstrations.

Let $(X_1, \ldots, X_N)$ be an observed path of $X$ satisfying Assumption $\mathrm{A}(d,d')$. As soon as a consistent estimator of $\mathbb{E}(e^2(a,0))$ is provided, property 1 allows to make a log-log regression-based estimation of $2d$. Which allows us together with $a \in \{1, \ldots, N-1\}$ to consider the sample variance of the wavelet coefficients,

$$T_N(a) := \frac{1}{N-a} \sum_{b=1}^{N-a} e^2(a,b). \tag{7}$$

**Remark 1.** *In Bardet et al. (2000), (2008) or in Moulines et al. (2007) or Roueff and Taqqu (2009), this sample variance of wavelet coefficients is*

$$V_N(a) := \frac{1}{[N/a]} \sum_{b=1}^{[N/a]} e^2(a,ab) \tag{8}$$

*(with $a = 2^j$ in case of multiresolution analysis). Definition (7) has both a drawback and two advantages with respect to the usual definition (8): not being adapted to the fast Mallat's algorithm it is more time consuming.*

*Its advantage twofold : we have a simpler expression of the asymptotic variance* $(\gamma_{ij})_{1\le i,j\le \ell}$ *(see (10) below,* $\gamma_{ij} = 4\pi \, \frac{(r_i r'_j)^{1-2d}}{K^2_{(\psi,2d)}} \int_{-\infty}^{\infty} \frac{\big|\widehat{\psi}(r_i\lambda)\big|^2 \big|\widehat{\psi}(r_j\lambda)\big|^2}{|\lambda|^{4d}} d\lambda$*), furthermore, as inferred from the numerical approximations, this asymptotic variance is smaller that the one obtained with (8), i.e.*

$$\gamma'_{ij} = \frac{2(r_i r_j)^{2-2d}}{K^2_{(\psi,2d)} d_{ij}} \sum_{m=-\infty}^{\infty} \Big( \int_{-\infty}^{\infty} \frac{\widehat{\psi}(ur_i)\overline{\widehat{\psi}}(ur_j)}{|u|^{2d}} \cos(u\, d_{ij} m)\, du \Big)^2 \quad \text{with} \quad d_{ij} = GCD(r_i, r_j)$$

*(diagonal terms are nearly twice as small as with* $(r_1, \dots, r_\ell) = (1, \dots, \ell)$*).*

The following proposition specifying a multidimensional central limit theorem for a vector $(\log \widetilde{T}_N(a_i))_i$, which provides the first step towards obtaining by log-log regression-based definition of the asymptotic properties of the ordinary least square estimator :

**Proposition 1.** *Define* $\ell \in \mathbb{N} \setminus \{0,1\}$ *and* $(r_1, \cdots, r_\ell) \in (\mathbb{N}^*)^\ell$*. Under Assumption* $A(d,d')$ *with* $d < 1/2$ *and* $d' > 0$*, if* $\psi$ *satisfies Assumption* $\Psi(k)$ *with* $k \ge d' - d + 1/2$ *and if* $(a_n)_{n\in\mathbb{N}}$ *is such as* $N/a_N \underset{N\to\infty}{\longrightarrow} \infty$ *and* $a_N\, N^{-1/(1+2d')} \underset{N\to\infty}{\longrightarrow} \infty$*, then*

$$\sqrt{\frac{N}{a_N}} \Big( \log T_N(r_i a_N) - 2d \log(r_i a_N) - \log \big( \frac{c_d}{2\pi} K_{(\psi,2d)} \big) \Big)_{1\le i\le \ell} \overset{d}{\underset{N\to\infty}{\longrightarrow}} \mathcal{N}_\ell\big(0\,;\, \Gamma(r_1, \dots, r_\ell, \psi, d)\big), \tag{9}$$

*with* $\Gamma(r_1, \dots, r_\ell, \psi, d) = (\gamma_{ij})_{1\le i,j\le \ell}$ *the asymptotic covariance matrix such as*

$$\gamma_{ij} = 4\pi \, \frac{(r_i r'_j)^{1-2d}}{K^2_{(\psi,2d)}} \int_{-\infty}^{\infty} \frac{\big|\widehat{\psi}(r_i\lambda)\big|^2 \big|\widehat{\psi}(r_j\lambda)\big|^2}{\lambda^{4d}} d\lambda. \tag{10}$$

# 3 Adaptive estimator of the memory parameter and adaptive goodness-of-fit test

The CLT of Proposition 1 opens a certain number of perspectives. As we shall see, the simple expression of the asymptotic covariance matrix reveals to be very advantageous as compared to the complicated expression of the asymptotic covariance obtained in the case of a multiresolution analysis (see Roueff and Taqqu, 2009a). Proposition 1 confirms the consistency of estimator $\widehat{d}_N$ of $d$. Hence, we define

$$\widehat{d}_N(a_N) := \big(0 \ \frac{1}{2}\big) \, (Z'_{a_N} Z_{a_N})^{-1} Z'_{a_N} \big( \log T_N(r_i a_N) \big)_{1\le i\le \ell} \quad \text{with} \quad Z_{a_N} = \begin{pmatrix} 1 & \log(a_N) \\ 1 & \log(2a_N) \\ \vdots & \vdots \\ 1 & \log(\ell a_N) \end{pmatrix}. \tag{11}$$

**Remark 2.** *To minimize the asymptotic covariance matrix* $\Gamma(r_1, \dots, r_\ell, \psi, d)$*, proposition 1 does not allow to choose* $(r_1, \dots, r_\ell)$ *unless we know the value of* $d$*. We therefore simply consider* $(r_1, r_2, \cdots, r_\ell) = (1, 2, \dots, \ell)$*.*

Then, it can be clearly inferred from Proposition 1 that $\widehat{d}_N(a_N)$ converges to $d$ following a central limit theorem with convergence rate $\sqrt{\frac{N}{a_N}}$ when $a_N$ satisfies the condition $a_N\, N^{-1/(1+2d')} \underset{N\to\infty}{\longrightarrow} \infty$.
But $d'$ is actually unknown. Bardet *et al.* (2008)presented an automatic procedure for choosing an "optimal" scale $a_N$. We shall presently apply this procedure. Here a brief recall of its principle: for $\alpha \in (0,1)$, define

$$Q_N(\alpha,c,d) = \Big(Y_N(\alpha) - Z_{N^\alpha}\begin{pmatrix} c \\ 2d \end{pmatrix}\Big)' \cdot \Big(Y_N(\alpha) - Z_{N^\alpha}\begin{pmatrix} c \\ 2d \end{pmatrix}\Big), \quad \text{with} \quad Y_N(\alpha) = \big(\log T_N(iN^\alpha)\big)_{1\leq i \leq \ell}.$$

$Q_N(\alpha,c,d)$ corresponds to a squared distance between the $\ell$ points $\big(\log(iN^\alpha)\,,\log T_N(iN^\alpha)\big)_i$ and a line. It can be minimized first by defining for $\alpha \in (0,1)$

$$\widehat{Q}_N(\alpha) = Q_N\big(\alpha, \widehat{c}(N^\alpha), 2\widehat{d}(N^\alpha)\big) \quad \text{with} \quad \begin{pmatrix} \widehat{c}(N^\alpha) \\ 2\widehat{d}(N^\alpha) \end{pmatrix} = \big(Z'_{N^\alpha} Z_{N^\alpha}\big)^{-1} Z'_{N^\alpha} Y_N(\alpha);$$

and by defining $\widehat{\alpha}_N$ by:

$$\widehat{Q}_N(\widehat{\alpha}_N) = \min_{\alpha \in \mathcal{A}_N} \widehat{Q}_N(\alpha) \quad \text{where} \quad \mathcal{A}_N = \Big\{\frac{2}{\log N}\,, \frac{3}{\log N}\,, \ldots, \frac{\log[N/\ell]}{\log N}\Big\}.$$

**Remark 3.** *As outlined in Bardet et al's. (2008) definition of the set $\mathcal{A}_N$, $\log N$ can be replaced by any sequence negligible with respect to any power law of $N$. Hence, in numerical applications we will use* $10\log N$ *which significantly increases the precision of $\widehat{\alpha}_N$.*

Under the assumptions of Proposition 1, we obtain (see the proof in Bardet *et al.*, 2008),

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

We then define:

$$\widehat{\widehat{d}_N} := \widehat{d}(N^{\widehat{\alpha}_N}) \quad \text{and} \quad \widehat{\Gamma}_N := \Gamma(1,\ldots,\ell,\widehat{\widehat{d}_N},\psi). \tag{12}$$

It is clear that $\widehat{\widehat{d}_N} \underset{N\to\infty}{\overset{\mathcal{P}}{\longrightarrow}} d$ (for a convergence rate see also Bardet *et al.*, 2008) and therefore, from the expression of $\Gamma$ in (10) which is a continuous function of the variable $d$, we obtain $\widehat{\Gamma}_N \underset{N\to\infty}{\overset{\mathcal{P}}{\longrightarrow}} \Gamma(1,\ldots,\ell,d,\psi)$.
We can thus define a (pseudo)-generalized least square estimator (PGLSE) of $d$. After defining :

$$\widetilde{\alpha}_N := \widehat{\alpha}_N + \frac{6\widehat{\alpha}_N}{(\ell-2)(1-\widehat{\alpha}_N)}\,\frac{\log\log N}{\log N}.$$

In the sequel and for a for reason of technical feasibility (*i.e.* $\Pr(\widetilde{\alpha}_N \leq \alpha^*) \underset{N\to\infty}{\longrightarrow} 0$ which is not satisfied by $\widehat{\alpha}_N$ (see Bardet *et al.*, 2008), we consider $\widetilde{\alpha}_N$ rather than $\widehat{\alpha}_N$. Consequently, we use the usual expression of PGLSE, the adaptive estimators of $c$ and $d$ can be defined as follows:

$$\begin{pmatrix} \widetilde{c}_N \\ 2\widetilde{d}_N \end{pmatrix} := \big(Z'_{N^{\widetilde{\alpha}_N}}\, \widehat{\Gamma}_N^{-1}\, Z_{N^{\widetilde{\alpha}_N}}\big)^{-1} Z'_{N^{\widetilde{\alpha}_N}}\, \widehat{\Gamma}_N^{-1}\, Y_N(\widetilde{\alpha}_N). \tag{13}$$

The following theorem provides the asymptotic behavior of the estimator $\widetilde{d}_N$,

**Theorem 1.** *Under assumptions of Proposition 1,*

$$\sqrt{\frac{N}{N^{\widetilde{\alpha}_N}}}\big(\widetilde{d}_N - d\big) \xrightarrow[N\to\infty]{d} \mathcal{N}\big(0\,;\,\sigma_d^2(\ell)\big) \quad \textit{with} \quad \sigma_d^2(\ell) := \big(0\ \tfrac{1}{2}\big)\big(Z_1'\big(\Gamma(1,\ldots,\ell,d,\psi)\big)^{-1}Z_1\big)^{-1}\big(0\ \tfrac{1}{2}\big)' \tag{14}$$

$$\textit{and for all} \ \ \rho > \frac{2(1+3d')}{(\ell-2)d'}, \quad \frac{N^{\frac{d'}{1+2d'}}}{(\log N)^{\rho}} \times \big|\widetilde{d}_N - d\big| \xrightarrow[N\to\infty]{\mathcal{P}} 0. \tag{15}$$

**Remark 4.** *1. From Gauss-Markov Theorem it is clear that the asymptotic variance of $\widetilde{d}_N$ is smaller or equal to the one of $\widehat{\widehat{d}_N}$. Moreover $\widetilde{d}_N$ satisfies the CLT (14) which provides confidence intervals which can be easily computed.*

*2. In the Gaussian case, the adaptive estimator $\widetilde{d}_N$ converge to $d$, its rate of convergence being equal to the minimax rate of convergence $N^{\frac{d'}{1+2d'}}$ up to a logarithm factor (see Giraitis et al., 1997). Thus, this estimator is comparable to adaptive log-periodogram or local Whittle estimators (see respectively Moulines and Soulier, 2003, and Robinson, 1995).*

*3. Under additive assumptions on $\psi$ ($\psi$ is supposed to have its first $m$ vanishing moments), the estimator $\widetilde{d}_N$ can also be applied to a process $X$ with an additive polynomial trend of degree $\leq m-1$. Then the trend is being "vanished" by the wavelet function in the expression of the wavelet coefficient and the value of $\widetilde{d}_N$ is the same as the result obtained without this additive trend. No such robustness property can be obtained with the cited adaptive log-periodogram or local Whittle estimator (however to an adaptive version of the local Whittle estimator which prooved robust for polynomial trends refer to Andrews and Sun, 2004).*

Finally it is easy to deduce from the previous pseudo-generalized least square regression an adaptive goodness-of-fit test. It consists on a sum of the PGLS squared distances between the PGLS regression line and the points. To be precise, consider the statistic:

$$\widetilde{T}_N := \frac{N}{N^{\widetilde{\alpha}_N}}\Big(Y_N(\widetilde{\alpha}_N) - Z_{N^{\widetilde{\alpha}_N}}\begin{pmatrix}\widetilde{c}_N \\ 2\widetilde{d}_N\end{pmatrix}\Big)'\,\widehat{\Gamma}_N^{-1}\Big(Y_N(\widetilde{\alpha}_N) - Z_{N^{\widetilde{\alpha}_N}}\begin{pmatrix}\widetilde{c}_N \\ 2\widetilde{d}_N\end{pmatrix}\Big). \tag{16}$$

Then, using the previous results, we obtain:

**Theorem 2.** *Under assumptions of Proposition 1,*

$$\widetilde{T}_N \xrightarrow[N\to\infty]{d} \chi^2(\ell-2). \tag{17}$$

This (adaptive) goodness-of-fit test is therefore very simple to be computed and used. In the case where $d > 0$, which can be tested easily from Theorem 1, this test can also be seen as a test of long memory for linear processes.

# 4 Simulations

We then examined the numerical consistency and robustness of $\widetilde{d}_N$. We proceeded to Simulations and we compared $\widetilde{d}_N$ estimator-computed results with the more accurate semiparametric long-memory estimators. To conclude we examined the numerical properties of the test statistic $\widetilde{T}_N$.

**Remark 5.** *Note that all softwares (in Matlab language) used in this section are freely available access on* `http://samm.univ-paris1.fr/-Jean-Marc-Bardet`.

First of all we need to specify the the simulation conditions. The results are based on 100 generated independent samples of each process belonging to the following "benchmark". The concrete generation procedures of these processes are based on the circulant matrix method in case of Gaussian processes and the truncation of an infinite sum if the process is non-Gaussian (see Doukhan *et al.*, 2003). The simulations carried out for $d = 0$, 0.1, 0.2, 0.3 and 0.4, for $N = 10^3$ and $10^4$ as well as the following processes which satisfy Assumption A$(d, d')$:

1. the fractional Gaussian noise (fGn) of parameter $H = d + 1/2$ (for $d \in [0, 0.5)$) and $\sigma^2 = 1$. A fGn is such that Assumption A$(d, 2)$ holds even if in general studies of the fGn do not include the Gaussian linear process;

2. a FARIMA$[p, d, q]$ process with parameter $d$ such that $d \in [0, 0.5)$, $p, q \in \mathbb{N}$. A FARIMA$[p, d, q]$ process is such that Assumption A$(d, 2)$ holds if $(\xi_i)_i$ the innovation process is such that $E\xi_i = 0$, $\mathbb{E}\xi_i^4 < \infty$ and $\xi_i$ symmetric random variables.

3. The centered Gaussian stationary process $X^{(d,d')}$, with spectral density is

$$f_3(\lambda) = \frac{1}{\lambda^{2d}}(1 + \lambda^{d'}) \quad \text{for } \lambda \in [-\pi, 0) \cup (0, \pi], \tag{18}$$

with $d \in [0, 0.5)$ and $d' \in (0, \infty)$. $X^{(d,d')}$ being a Gaussian process with spectral density $f_3$, it is considered a linear process within the Wold decomposition Theorem as well , thus confirming Assumption $A(d, d')$ holds.

The "benchmark" referred to ,below include following particular processes for $d = 0$, 0.1, 0.2, 0.3, 0.4:

- $X_1$ : fGn processes with parameters $H = d + 1/2$;
- $X_2$ : FARIMA$[0, d, 0]$ processes with standard Gaussian innovations;
- $X_3$ : FARIMA$[0, d, 0]$ processes with innovations following a uniform $\mathcal{U}[-1, 1]$ distribution;

- $X_4$ : FARIMA$(0,d,0)$ processes with innovations satisfying a symmetric Burr distribution with cumulative distribution function $F(x)=1-\frac{1}{2}\frac{1}{1+x^2}$ for $x\geq 0$ and $F(x)=\frac{1}{2}\frac{1}{1+x^2}$ for $x\leq 0$ (and therefore $\mathbb{E}|X_i|^2=\infty$ but $\mathbb{E}|X_i|<\infty$);

- $X_5$ : FARIMA$(0,d,0)$ processes with innovations satisfying a symmetric Burr distribution with cumulative distribution function $F(x)=1-\frac{1}{2}\frac{1}{1+|x|^{3/2}}$ for $x\geq 0$ and $F(x)=\frac{1}{2}\frac{1}{1+|x|^{3/2}}$ for $x\leq 0$ (and therefore $\mathbb{E}|X_i|^2=\infty$ but $\mathbb{E}|X_i|<\infty$);

- $X_6$ : FARIMA$[1,d,1]$ processes with standard Gaussian innovations, MA coefficient $\phi=-0.3$ and AR coefficient $\phi=0.7$;

- $X_7$ : FARIMA$[1,d,1]$ processes with innovations following a uniform $\mathcal{U}[-1,1]$ distribution, MA coefficient $\phi=-0.3$ and AR coefficient $\phi=0.7$;

- $X_8$ : $X^{(d,d')}$ Gaussian processes with $d'=1$.

Note that the processes $X_4$ and $X_5$ do not satisfy the condition $\mathbb{E}\xi_0^4$ required in Theorems 1 and 2. However, considering the logarithm of wavelet coefficient sample variance and not only the wavelet coefficient sample variance, we should be able to prove the consistency of $\widetilde{d}_N$ under $\mathbb{E}\xi_0^r$ with $r\geq 2$.

## 4.1 Comparison of the wavelet-based estimator with other estimators

the wavelet-based estimator has been selected on the following base:

**Choice of the function $\psi$:** A wavelet function $\psi$ associated with a multi-resolution analysis being not mandatory, as mentioned above, we use function $\psi(x)=x^3(1-x)^3\big(x^3-\frac{3}{2}\,x^2+\frac{15}{22}\,x-\frac{1}{11}\big)\mathbb{I}_{x\in[0,1]}$ which satisfies Assumption $\Psi(2)$

**Choice of the parameter $\ell$:** This parameter largely determines the "beginning" of the linear part of the graph drawn by points $(\log(ia_N),\log T_N(ia_N))_{1\leq i\leq\ell}$ and hence the data-driven $\widehat{a}_N$.

We adopted on this point a two step procedure:

1. According to numerical study (not detailed here), $\ell=[2*log(N)]$ (therefore $\ell=13$ for $N=1000$ and $\ell=18$ for $N=10000$) seems an appropriate first step: the computation of $\widehat{\alpha}_n$.

2. Concerning computation of $\tilde{d}_N$, $\widehat{\Gamma}_N$ seems to be independant of $d$. Using classical approximations of the integrals defined in $\Gamma(1,\ldots,\ell,d,\psi)$, we compute $\sigma_d^2(\ell)=\big(0\ \frac{1}{2}\big)\big(Z_1'\big(\Gamma(1,\ldots,\ell,d,\psi)\big)^{-1}Z_1\big)^{-1}\big(0\ \frac{1}{2}\big)'$

taking into account several values of $d$ and $\ell$ . For the results of these numerical experiments refer to Figure 2. It can be inferred that any $d \in [0, 0.5)$, $\sigma_d^2(\ell)$ is almost independent on $d$ and decreases as $\ell$ increases. Chosing the second step $\ell = N^{1-\widetilde{\alpha}_N}(\log N)^{-1}$, we notice that the larger considered scale is $N(\log N)^{-1}$ (which is negligible with respect to $N$, confirming CLT 9).

Figure 1: *Graph of the approximated values of* $\sigma_d^2(\ell)$ *defined in (14) for* $d \in [0, 0.5]$ *and* $\ell = 10, 20, 50, 100, 200$ *and* $500$.

Applying $\widetilde{d}_N$ as well as 2 other semiparametric $d$-estimators (see Bardet *et al*, 2003 or 2008) to the above mentioned benchmark-processes, we obtain :

- $\widehat{d}_{MS}$ is the adaptive global log-periodogram estimator introduced by Moulines and Soulier (1998, 2003), also called FEXP estimator, with bias-variance balance parameter $\kappa = 2$;
- $\widehat{d}_R$ is the local Whittle estimator introduced by Robinson (1995). The trimming parameter is $m = N/30$.

For simulation results see Table 1.

*Conclusions from Table 1:* Compared to other estimators, $\widetilde{d}_N$ shows numerically convincing convergence rate. With both the "spectral" estimator $\widehat{d}_R$ and $\widehat{d}_{MS}$, the results are quiet stable and hardly sensible to $d$ and to the flatness of the spectral density of the process. However the spectral density of the process notably effects the convergence rate of $\widetilde{d}_N$. As compared to other estimators, $\widetilde{d}_N$ is a very accurate and even more efficient for "smooth" spectral densities (fGn and FARIMA$(0, d, 0)$), $\widetilde{d}_N$.

**Remark 6.** *A previous comparaison (Bardet et al. (2008)) of two adaptive wavelet-based estimators (respectively defined in Veitch et al., (2003) and in Bardet et al. (2008)) with* $\widehat{d}_{MS}$ *and* $\widehat{d}_R$ *(as well as with two further estimators as defined respectively in Giraitis et al., (2000), and Giraitis et al., (2006) neither of which*

$N = 10^3 \longrightarrow$

| Model | $\sqrt{MSE}$ | $d = 0$ | $d = 0.1$ | $d = 0.2$ | $d = 0.3$ | $d = 0.4$ |
|---|---|---|---|---|---|---|
| $X_1$ | $\sqrt{MSE}\ \widehat{d}_{MS}$ | 0.089 | 0.091 | 0.096 | 0.090 | 0.100 |
| | $\sqrt{MSE}\ \widehat{d}_R$ | 0.102 | 0.114 | 0.116 | 0.106 | 0.102 |
| | $\sqrt{MSE}\ \widetilde{d}_N$ | **0.047** | **0.046** | **0.042** | **0.052** | **0.047** |
| | $\widetilde{p}_n$ | 0.85 | 0.76 | 0.78 | 0.76 | 0.64 |
| $X_2$ | $\sqrt{MSE}\ \widehat{d}_{MS}$ | 0.091 | 0.094 | 0.086 | 0.091 | 0.099 |
| | $\sqrt{MSE}\ \widehat{d}_R$ | 0.107 | 0.105 | 0.112 | 0.110 | 0.097 |
| | $\sqrt{MSE}\ \widetilde{d}_N$ | **0.048** | **0.050** | **0.053** | **0.061** | **0.074** |
| | $\widetilde{p}_n$ | 0.82 | 0.82 | 0.75 | 0.73 | 0.67 |
| $X_3$ | $\sqrt{MSE}\ \widehat{d}_{MS}$ | 0.092 | 0.094 | 0.080 | 0.099 | 0.096 |
| | $\sqrt{MSE}\ \widehat{d}_R$ | 0.113 | 0.113 | 0.100 | 0.112 | 0.095 |
| | $\sqrt{MSE}\ \widetilde{d}_N$ | **0.052** | **0.071** | **0.063** | **0.077** | **0.092** |
| | $\widetilde{p}_n$ | 0.84 | 0.72 | 0.75 | 0.67 | 0.51 |
| $X_4$ | $\sqrt{MSE}\ \widehat{d}_{MS}$ | 0.088 | 0.079 | 0.079 | 0.093 | 0.104 |
| | $\sqrt{MSE}\ \widehat{d}_R$ | 0.096 | 0.100 | 0.103 | 0.097 | 0.095 |
| | $\sqrt{MSE}\ \widetilde{d}_N$ | **0.051** | **0.066** | **0.056** | **0.061** | **0.064** |
| | $\widetilde{p}_n$ | 0.84 | 0.78 | 0.78 | 0.75 | 0.66 |
| $X_5$ | $\sqrt{MSE}\ \widehat{d}_{MS}$ | **0.069** | **0.067** | **0.077** | 0.121 | 0.143 |
| | $\sqrt{MSE}\ \widehat{d}_R$ | 0.072 | 0.078 | 0.093 | **0.087** | **0.074** |
| | $\sqrt{MSE}\ \widetilde{d}_N$ | 0.073 | 0.069 | 0.083 | **0.087** | 0.120 |
| | $\widetilde{p}_n$ | 0.73 | 0.69 | 0.68 | 0.74 | 0.64 |
| $X_6$ | $\sqrt{MSE}\ \widehat{d}_{MS}$ | **0.096** | **0.091** | **0.090** | **0.086** | **0.093** |
| | $\sqrt{MSE}\ \widehat{d}_R$ | 0.111 | 0.102 | 0.100 | 0.101 | 0.101 |
| | $\sqrt{MSE}\ \widetilde{d}_N$ | 0.153 | 0.146 | 0.144 | 0.158 | 0.147 |
| | $\widetilde{p}_n$ | 0.52 | 0.47 | 0.48 | 0.39 | 0.50 |
| $X_7$ | $\sqrt{MSE}\ \widehat{d}_{MS}$ | **0.085** | **0.096** | **0.086** | **0.093** | 0.098 |
| | $\sqrt{MSE}\ \widehat{d}_R$ | 0.106 | 0.116 | 0.097 | 0.099 | **0.092** |
| | $\sqrt{MSE}\ \widetilde{d}_N$ | 0.155 | 0.150 | 0.56 | 0.147 | 0.157 |
| | $\widetilde{p}_n$ | 0.60 | 0.55 | 0.49 | 0.52 | 0.41 |
| $X_8$ | $\sqrt{MSE}\ \widehat{d}_{MS}$ | **0.097** | **0.104** | **0.097** | **0.094** | **0.101** |
| | $\sqrt{MSE}\ \widehat{d}_R$ | 0.120 | 0.116 | 0.117 | 0.113 | 0.110 |
| | $\sqrt{MSE}\ \widetilde{d}_N$ | 0.179 | 0.189 | 0.177 | 0.175 | 0.176 |
| | $\widetilde{p}_n$ | 0.75 | 0.75 | 0.68 | 0.66 | 0.67 |

$N = 10^4 \longrightarrow$

| Model | $\sqrt{MSE}$ | $d = 0$ | $d = 0.1$ | $d = 0.2$ | $d = 0.3$ | $d = 0.4$ |
|---|---|---|---|---|---|---|
| $X_1$ | $\sqrt{MSE}\ \widehat{d}_{MS}$ | 0.032 | 0.029 | 0.031 | 0.031 | 0.036 |
| | $\sqrt{MSE}\ \widehat{d}_R$ | 0.028 | 0.028 | **0.029** | 0.029 | 0.032 |
| | $\sqrt{MSE}\ \widetilde{d}_N$ | **0.016** | **0.027** | 0.034 | **0.025** | **0.022** |
| | $\widetilde{p}_n$ | 0.97 | 0.93 | 0.97 | 0.94 | 0.97 |
| $X_2$ | $\sqrt{MSE}\ \widehat{d}_{MS}$ | 0.034 | 0.030 | 0.029 | 0.032 | 0.028 |
| | $\sqrt{MSE}\ \widehat{d}_R$ | 0.027 | 0.027 | 0.029 | 0.028 | **0.023** |
| | $\sqrt{MSE}\ \widetilde{d}_N$ | **0.026** | **0.019** | **0.019** | **0.019** | 0.025 |
| | $\widetilde{p}_n$ | 0.95 | 0.97 | 0.98 | 0.96 | 0.94 |
| $X_3$ | $\sqrt{MSE}\ \widehat{d}_{MS}$ | 0.034 | 0.034 | 0.033 | 0.030 | 0.031 |
| | $\sqrt{MSE}\ \widehat{d}_R$ | 0.029 | 0.028 | 0.028 | 0.028 | **0.029** |
| | $\sqrt{MSE}\ \widetilde{d}_N$ | **0.027** | **0.017** | **0.016** | **0.022** | 0.030 |
| | $\widetilde{p}_n$ | 0.93 | 0.96 | 0.97 | 0.93 | 0.92 |
| $X_4$ | $\sqrt{MSE}\ \widehat{d}_{MS}$ | 0.029 | 0.060 | 0.036 | 0.031 | 0.031 |
| | $\sqrt{MSE}\ \widehat{d}_R$ | 0.025 | 0.027 | 0.029 | 0.031 | 0.029 |
| | $\sqrt{MSE}\ \widetilde{d}_N$ | **0.016** | **0.020** | **0.021** | **0.015** | **0.023** |
| | $\widetilde{p}_n$ | 0.95 | 0.91 | 0.97 | 0.92 | 0.91 |
| $X_5$ | $\sqrt{MSE}\ \widehat{d}_{MS}$ | 0.093 | **0.046** | 0.039 | 0.073 | 0.047 |
| | $\sqrt{MSE}\ \widehat{d}_R$ | **0.040** | **0.046** | 0.035 | 0.032 | **0.024** |
| | $\sqrt{MSE}\ \widetilde{d}_N$ | 0.056 | 0.071 | **0.027** | **0.025** | **0.024** |
| | $\widetilde{p}_n$ | 0.85 | 0.88 | 0.93 | 0.86 | 0.85 |
| $X_6$ | $\sqrt{MSE}\ \widehat{d}_{MS}$ | 0.031 | 0.032 | 0.033 | 0.032 | 0.029 |
| | $\sqrt{MSE}\ \widehat{d}_R$ | **0.029** | **0.028** | **0.028** | **0.028** | **0.028** |
| | $\sqrt{MSE}\ \widetilde{d}_N$ | 0.045 | 0.044 | 0.046 | 0.044 | 0.041 |
| | $\widetilde{p}_n$ | 0.96 | 0.93 | 0.89 | 0.93 | 0.90 |
| $X_7$ | $\sqrt{MSE}\ \widehat{d}_{MS}$ | 0.030 | 0.031 | 0.037 | 0.030 | 0.029 |
| | $\sqrt{MSE}\ \widehat{d}_R$ | **0.027** | **0.027** | **0.032** | **0.028** | **0.027** |
| | $\sqrt{MSE}\ \widetilde{d}_N$ | 0.049 | 0.044 | 0.050 | 0.048 | 0.046 |
| | $\widetilde{p}_n$ | 0.94 | 0.91 | 0.88 | 0.87 | 0.86 |
| $X_8$ | $\sqrt{MSE}\ \widehat{d}_{MS}$ | **0.038** | 0.040 | **0.040** | **0.035** | 0.037 |
| | $\sqrt{MSE}\ \widehat{d}_R$ | 0.039 | **0.038** | **0.040** | 0.036 | **0.035** |
| | $\sqrt{MSE}\ \widetilde{d}_N$ | 0.085 | 0.083 | 0.086 | 0.087 | 0.085 |
| | $\widetilde{p}_n$ | 0.92 | 0.94 | 0.94 | 0.95 | 0.93 |

Table 1: *Comparison of the different long-memory parameter estimators for benchmark processes. For each process and value of* $d$ *and* $N$, $\sqrt{MSE}$ *takes into account* 100 *independently generated samples. The frequency of acceptation of the adaptive goodness-of-fit test is* $\widetilde{p}_n = \frac{1}{n}\,\#\big(\widetilde{T}_N < q_{\chi^2(\ell-2)}(0.95)\big)$.

| | Model | $\sqrt{MSE}$ | $d=0$ | $d=0.1$ | $d=0.2$ | $d=0.3$ | $d=0.4$ |
|---|---|---|---|---|---|---|---|
| $N=10^3$ $\longrightarrow$ | GARMA$(0,d,0)$ | $\sqrt{MSE}\ \widehat{d}_{MS}$ | 0.089 | 0.091 | 0.123 | 0.132 | 0.166 |
| | | $\sqrt{MSE}\ \widehat{d}_R$ | 0.112 | 0.111 | 0.119 | **0.106** | **0.106** |
| | | $\sqrt{MSE}\ \widetilde{d}_N$ | **0.041** | **0.076** | **0.114** | 0.142 | 0.180 |
| | | $\widetilde{p}_n$ | 0.82 | 0.78 | 0.63 | 0.59 | 0.46 |
| | Trend | $\sqrt{MSE}\ \widehat{d}_{MS}$ | 0.548 | 0.411 | 0.292 | 0.190 | 0.142 |
| | | $\sqrt{MSE}\ \widehat{d}_R$ | 0.499 | 0.394 | 0.279 | 0.167 | 0.091 |
| | | $\sqrt{MSE}\ \widetilde{d}_N$ | **0.044** | **0.052** | **0.056** | **0.060** | **0.065** |
| | | $\widetilde{p}_n$ | 0.83 | 0.81 | 0.80 | 0.73 | 0.64 |
| | Trend + Seasonality | $\sqrt{MSE}\ \widehat{d}_{MS}$ | 0.479 | 0.347 | 0.233 | 0.142 | 0.112 |
| | | $\sqrt{MSE}\ \widehat{d}_R$ | 0.499 | 0.393 | 0.279 | 0.167 | **0.091** |
| | | $\sqrt{MSE}\ \widetilde{d}_N$ | **0.216** | **0.215** | **0.215** | **0.217** | 0.185 |
| | | $\widetilde{p}_n$ | 0.35 | 0.26 | 0.18 | 0.21 | 0.18 |

| | Model | $\sqrt{MSE}$ | $d=0$ | $d=0.1$ | $d=0.2$ | $d=0.3$ | $d=0.4$ |
|---|---|---|---|---|---|---|---|
| $N=10^4$ $\longrightarrow$ | GARMA$(0,d,0)$ | $\sqrt{MSE}\ \widehat{d}_{MS}$ | 0.031 | 0.035 | 0.039 | 0.049 | 0.062 |
| | | $\sqrt{MSE}\ \widehat{d}_R$ | 0.028 | **0.031** | **0.030** | **0.030** | **0.034** |
| | | $\sqrt{MSE}\ \widetilde{d}_N$ | **0.023** | 0.053 | 0.052 | 0.058 | 0.060 |
| | | $\widetilde{p}_n$ | 0.96 | 0.94 | 0.93 | 0.91 | 0.88 |
| | Trend | $\sqrt{MSE}\ \widehat{d}_{MS}$ | 0.452 | 0.286 | 0.167 | 0.096 | 0.056 |
| | | $\sqrt{MSE}\ \widehat{d}_R$ | 0.433 | 0.308 | 0.191 | 0.100 | 0.051 |
| | | $\sqrt{MSE}\ \widetilde{d}_N$ | **0.014** | **0.016** | **0.016** | **0.021** | **0.028** |
| | | $\widetilde{p}_n$ | 0.99 | 0.97 | 0.97 | 0.95 | 0.93 |
| | Trend + Seasonality | $\sqrt{MSE}\ \widehat{d}_{MS}$ | 0.471 | 0.307 | 0.196 | 0.123 | 0.076 |
| | | $\sqrt{MSE}\ \widehat{d}_R$ | 0.432 | 0.305 | 0.191 | 0.100 | 0.052 |
| | | $\sqrt{MSE}\ \widetilde{d}_N$ | **0.044** | **0.069** | **0.047** | **0.042** | **0.045** |
| | | $\widetilde{p}_n$ | 0.83 | 0.81 | 0.76 | 0.78 | 0.82 |

Table 2: *Robustness of the different long-memory parameter estimators. For each process and value of d and N, $\sqrt{MSE}$ takes into account* 100 *independent generated samples. The frequency of acceptation of the adaptive goodness-of-fit test is* $\widetilde{p}_n = \frac{1}{n}\,\#\big(\widetilde{T}_N < q_{\chi^2(\ell-2)}(0.95)\big)$.

*display good numerical properties of consistenciy.) shows that $\sqrt{MSE}$ of $\widetilde{d}_N$ obtained in Table 1 is generally smaller to $\sqrt{MSE}$ of Bardet et al.'s (2008)-based estimator. Because we opted for definition (7) instead of (8) and PGLS regression instead of LS regression.*

**Comparison of the robustness of the different semiparametric estimators:**

To conclude, take three different processes not satisfying Assumption $A(d,d')$ as follows:

- A Gaussian stationary process with a spectral density $f(\lambda) = \big||\lambda| - \pi/2\big|^{-2d}$ for all $\lambda \in [-\pi,\pi] \setminus \{-\pi/2, \pi/2\}$. The local behavior of $f$ in 0 is $f(|\lambda|) \sim (\pi/2)^{-2d}\,|\lambda|^{-2d}$ with $d=0$. It does not satisfy Assumption $A(0,2)$.
- A Gaussian FARIMA$(0,d,0)$ with an additive linear trend ($X_t = FARIMA_t + (1-2t/n)$ for $t = 1,\cdots,n$ and therefore mean value$(X_1,\cdots,X_n) \simeq 0$);
- A Gaussian FARIMA$(0,d,0)$ with an additive linear trend and an additive sinusoidal seasonal component of period $T=12$ ($X_t = FARIMA_t + (1-2t/n) + \sin(\pi\, t/6)$ for $t = 1,\cdots,n$ hence mean value$(X_1,\cdots,X_n) \simeq 0$).

For results of these simulations see Table 2.

*Conclusions from Table 2:* The main advantage of $\widetilde{d}_N$ with respect to $\widehat{d}_{MS}$ and $\widehat{d}_R$ as listed in this table: is the robust with respectness to smooth trends (or seasonality). Note that the sample mean value of $\widehat{d}_{MS}$ and $\widehat{d}_R$ for processes with trend or with trend and seasonality is almost 0.5 for any choice of $d$.

## 4.2 Consistency and robustness of the adaptive goodness-of-fit test:

Tables 1 and 2 provide informations concerning the adaptive goodness-of-fit test. The consistency properties of this test are clearly satisfactory when $N$ is large enough ($N = 1000$ seems to be too small to correctly using this goodness-of-fit test).

In order to appreciate the tendancy of the test statistic under $H_1$. We take a process which satisfying neither the stationarity condition nor relation (1) (verified by the spectral density). We have 3 particular cases :

1. a process $X$ denoted MFARIMA and defined as a succession of two independent Gaussian FARIMA processes. More precisely, we consider $X_t = FARIMA(0, 0.1, 0)$ for $t = 1, \cdots, n/2$ and $X_t = FARIMA(0, 0.4, 0)$ for $t = n/2 + 1, \cdots, n$.

2. a process $X$ denoted MGN and defined by the increments of a multifractional Brownian motion (introduced in Peltier and Lévy-Vehel, 1995). Using the harmonizable representation, define $Y = (Y_t)_t$ by
$$Y_t := C(t) \int_{\mathbb{R}} \frac{e^{itx} - 1}{|x|^{H(t)+1/2}} W(dx)$$
where $W(dx)$ is a complex-valued Gaussian noise with variance $dx$ and $H(\cdot)$ as well as $C(\cdot)$ are functions (the case $H(\cdot) = H$ with $H \in (0, 1)$ is the case of fBm). Consider $H(t) = 0.5+0.4\sin(t/10)$ and $C(t) = 1$. Then $X_t = Y_{t+1} - Y_t$ for $t \in \mathbb{Z}$. The process $X$ is not a stationary process, it rather behaves "locally" as a fGn with a parameter $H(t)$ (therefore depending on $t$).

3. a process $X$ denoted MFGN and defined by the increments of a multiscale fractional Brownian motion (introduced in Bardet and Bertrand, 2007). Let $Z = (Z_t)_t$ be such that
$$Z_t := \int_{\mathbb{R}} \sigma(x) \frac{e^{itx} - 1}{|x|^{H(x)+1/2}} W(dx)$$
where $W(dx)$ is a complex-valued Gaussian noise with variance $dx$, $H(\cdot)$ and $\sigma(\cdot)$ are piecewise constant functions. Consider function $H(x) = 0.9$ for $0.001 \le x \le 0.04$ and $H(x) = 0.1$ for $0.04 \le x \le 3$. Define $X_t := Z_{t+1} - Z_t$ for $t \in \mathbb{Z}$ then $X = (X_t)_{t \in \mathbb{Z}}$ is a Gaussian stationary process which can be written as

| Model | $N=10^3$ | $N=10^4$ |
|---|---|---|
| MFARIMA | $\widetilde{p}_n=0.58$ | $\widetilde{p}_n=0.87$ |
| MGN | $\widetilde{p}_n=0.18$ | $\widetilde{p}_n=0.08$ |
| MFGN | $\widetilde{p}_n=0.02$ | $\widetilde{p}_n=0.04$ |

Table 3: *Robustness of the adaptive goodness-of-fit test. The frequency of acceptation of the adaptive goodness-of-fit test is* $\widetilde{p}_n=\frac{1}{n}\#\big(\widetilde{T}_N<q_{\chi^2(\ell-2)}(0.95)\big)$.

a Gaussian linear process (Wold decomposition Theorem) behaving as a fGn of parameter 0.9 for low frequencies (large time) and a fGn of parameter 0.1 for high frequencies (small time).

We used the test statistic to 100 independent replications of these processes. The results figure in Table 3. The goodness-of-fit test is rejected for processes MGN and MFGN. Whereas for the process MFARIMA which actually does not satisfy the Assumption of the Theorem 2 it is not rejected. It is due to the fact the test calculates the average behavior of the sample whereas in case of change (for example MFARIMA) it calculates the average of LRD parameter

(an average of 0.30 for $\widetilde{d_N}$ and a standard deviation 0.03 are obtained).

# 5 Proofs

We shall proceed to applications of lemma.

**Lemma 1.** *If $g$ is a function satisfying Assumption $\Psi(k)$ with $k\geq 1$, then for all $\lambda\in\mathbb{R}$,*

$$\Big|\frac{1}{a}\sum_{j=1}^{a}g\big(\frac{j}{a}\big)e^{-i\lambda\,\frac{j}{a}}-\int_0^1 g(t)e^{-i\lambda\,t}dt\Big|\leq C_g(k)\,\min\Big(\frac{1+|\lambda|^k}{a^k}\,,\,1\Big)\quad with\quad C_g(k)=2\sum_{p=0}^{k}\Big(\begin{array}{c}k\\p\end{array}\Big)\sup_{x\in[0,1]}|g^{(p)}(x)|. \tag{19}$$

*Proof of Lemma 1.* 1/ If $h$ is a $\mathcal{C}^k(\mathbb{R})$ function such as $h(x)=0$ for $x\notin[0,1]$ with $k\geq 1$, then for all $a>0$:

$$\Big|\frac{1}{a}\sum_{j=1}^{a}h\big(\frac{j}{a}\big)-\int_0^1 h(t)dt\Big|\leq\sup_{x\in[0,1]}|h^{(k)}(x)|\,\frac{1}{a^k}. \tag{20}$$

This proof is established by induction on $k$. If $k=1$, the classical approximation of an integral by a Riemann sum implies

$$\Big|\frac{1}{a}\sum_{j=1}^{a}h\big(\frac{j}{a}\big)-\int_0^1 h(t)dt\Big|\leq\big(\frac{1}{2}\sup_{x\in[0,1]}|h'(x)|\big)\,\frac{1}{a}\leq\sup_{x\in[0,1]}|h'(x)|\,\frac{1}{a}.$$

Assuming that property (20) is true for any $k\leq n$ with $n\in\mathbb{N}^*$. We can to prove that (20) is also true for $k=n+1$. Supposing that $h$ satisfies Assumption $\Psi(n+1)$. We then obtain, with the usual Taylor expansion $\big|h(t)-h(u)-\sum_{k=1}^n\frac{(t-u)^k}{k!}h^{(k)}(u)\big|\leq\frac{|t-u|^{n+1}}{(n+1)!}\sup_{x\in[0,1]}|h^{(n+1)}(x)|$ for $(t,u)\in[0,1]^2$,

$$\begin{aligned}\Big|\frac{1}{a}\sum_{j=1}^{a}h\big(\frac{j}{a}\big)-\int_0^1 h(t)dt\Big| &\leq \Big|\sum_{j=1}^{a}\int_{(j-1)/a}^{j/a}\sum_{k=1}^{n}\frac{(j/a-t)^k}{k!}h^{(k)}(j/a)dt\Big|+\big(\frac{1}{(n+2)!}\sup_{x\in[0,1]}|h^{(n+1)}(x)|\big)\frac{1}{a^{n+1}}\\ &\leq \sum_{k=1}^{n}\frac{1}{a^k(k+1)!}\Big|\frac{1}{a}\sum_{j=1}^{a}h^{(k)}(j/a)dt\Big|+\big(\frac{1}{(n+2)!}\sup_{x\in[0,1]}|h^{(n+1)}(x)|\big)\frac{1}{a^{n+1}}.\end{aligned}$$

If we use (20) for $h^{(k)}$ and $k = 1, \dots, n$, we have

$$\Big|\frac{1}{a}\sum_{j=1}^{a} h^{(k)}(j/a)dt - \int_0^1 h^{(k)}(t)dt\Big| \le \frac{1}{(n-k+1)!}\sup_{x\in[0,1]}|h^{(n+1)}(x)|\frac{1}{a^{n+1-k}}$$

since $h^{(k)}$ satisfies Assumption $\Psi(n+1-k)$. Given $\int_0^1 h^{(k)}(t)dt = \big[\frac{1}{(k+1)!}h^{(k+1)}(t)\big]_0^1 = 0$. We have,

$$\begin{aligned}\Big|\frac{1}{a}\sum_{j=1}^{a} h\big(\frac{j}{a}\big) - \int_0^1 h(t)dt\Big| &\le \Big(\sum_{k=1}^{n}\frac{1}{(k+1)!}\frac{1}{(n-k+1)!} + \frac{1}{(n+2)!}\Big)\sup_{x\in[0,1]}|h^{(n+1)}(x)|\,\frac{1}{a^{n+1}} \\ &\le (e-2)\sup_{x\in[0,1]}|h^{(n+1)}(x)|\,\frac{1}{a^{n+1}},\end{aligned}$$

and thus (20) is verified for $k = n+1$ and therefore for any $k \in \mathbb{N}^*$.

2/ Now, we apply (20) for $h(t) = g(t)e^{-it\lambda}$ when $\lambda \in [a, a]$. Since $|h^{(k)}(t)| \le \sum_{p=0}^{k}\left(\begin{array}{c} k \\ p \end{array}\right)|\lambda|^p|g^{(k-p)}(t)|$, and for all $\lambda \in [a, a]$, $\sup_{x\in[0,1]}|h^{(k)}(x)| \le \max(1, |\lambda|^k)\sum_{p=0}^{k}\left(\begin{array}{c} k \\ p \end{array}\right)\sup_{x\in[0,1]}|g^{(p)}(x)|$ and (19) holds.

If $|\lambda| > a$, it is obvious that

$$\Big|\frac{1}{a}\sum_{j=1}^{a} g\big(\frac{j}{a}\big)e^{-i\lambda\,\frac{j}{a}} - \int_0^1 g(t)e^{-i\lambda\,t}dt\Big| \le 2\sup_{x\in[0,1]}|g(x)|$$

Conscequently (19) holds. Moreover, if $g$ is not the null function, we can not expect a really smaller bound. Indeed, if we denote $\lambda'$ such as $\int_0^1 g(t)e^{-i\lambda' t}dt \ne 0$ (if $\lambda'$ does not exist, $g(x) = 0$ for all $x \in \mathbb{R}$). Then, for $a > \lambda'$ and for $\lambda = \lambda' + 2n\pi a$ with $n \in \mathbb{Z}^*$, then $\frac{1}{a}\sum_{j=1}^{a} g(j/a)e^{-i\lambda j/a} = \frac{1}{a}\sum_{j=1}^{a} g(j/a)e^{-i\lambda' j/a} = \int_0^1 g(t)e^{-i\lambda' t} + O(a^{-k})$ when $a \to \infty$ from the above case $|\lambda'| \le a$. But we also have $\int_0^1 g(t)e^{-i\lambda t} = O(|\lambda|^{-k}) = O(a^{-k})$ from $k$ integrations by parts since $g$ satisfies Assumption $\Psi(k)$. Therefore, for any $\lambda = \lambda' + 2n\pi a$ with $n \in \mathbb{Z}^*$, we have:

$$\Big|\frac{1}{a}\sum_{j=1}^{a} g\big(\frac{j}{a}\big)e^{-i\lambda\,\frac{j}{a}} - \int_0^1 g(t)e^{-i\lambda\,t}dt\Big| = \Big|\int_0^1 g(t)e^{-i\lambda' t}\Big| + O\big(a^{-k}\big)$$

Which means that no better bound than $O(1)$ when $\lambda \in \mathbb{R}$ can be obtained. □

**Lemma 2.** *If $g$ is a function satisfying Assumption $\Psi(k)$ with $k \ge 0$, then for all $a \ge 1$ and $\lambda \in [-a\pi, 0) \cup (0, a\pi]$,*

$$\Big|\frac{1}{a}\sum_{j=1}^{a} g\big(\frac{j}{a}\big)e^{-i\lambda\,\frac{j}{a}}\Big| \le D_g(k)\,\frac{1}{|\lambda|^k} \quad \textit{with} \quad D_g(k) = 10^k\sup_{x\in[0,1]}|g^{(k)}(x)|. \tag{21}$$

*Proof of Lemma 2.* This proof is also established by induction on $k$. If $k = 0$, it is obvious that:

$$\Big|\frac{1}{a}\sum_{j=1}^{a} g\big(\frac{j}{a}\big)e^{-i\lambda\,\frac{j}{a}}\Big| \le \sup_{x\in[0,1]}|g(x)|\big),$$

thus satisfying (21). Assume (21) is true for any $k \le n$ with $n \in \mathbb{N}^*$. We can prove that (21) is also true for $k = n+1$. Assume $g$ satisfies Assumption $\Psi(n+1)$. With $S_j(a, \lambda) := \displaystyle\sum_{\ell=0}^{j} e^{-i\lambda\ell/a} = \frac{1}{2i\sin(\lambda/2a)}\big(e^{i\lambda/2a} -$

$e^{-i\lambda/2a}e^{-ij\lambda/a}\big)$ for $j\in\{0,1,\ldots,a\}$, we obtain:

$$\begin{aligned}\Big|\frac{1}{a}\sum_{j=1}^{a} g\big(\frac{j}{a}\big)e^{-i\lambda\,\frac{j}{a}}\Big| &= \Big|\frac{1}{a}\sum_{j=1}^{a} g\big(\frac{j}{a}\big)\big(S_j(a,\lambda)-S_{j-1}(a,\lambda)\big)\Big| \\ &\le I_a(\lambda)+\frac{1}{a}\big|g\big(\frac{1}{a}\big)\big| \quad \text{with} \quad I_a(\lambda):=\Big|\frac{1}{a}\sum_{j=1}^{a-1}\big(g\big(\frac{j}{a}\big)-g\big(\frac{j+1}{a}\big)\big)S_j(a,\lambda)\Big|. \qquad (22)\end{aligned}$$

But since $g$ satisfies Assumption $\Psi(n+1)$ and $a\ge 1$, we have :

$$\frac{1}{a}\big|g\big(\frac{1}{a}\big)\big|\le \sup_{x\in[0,1]}|g^{(n+1)}(x)|\,\frac{1}{a^{n+1}(n+1)!}. \qquad (23)$$

With the usual Taylor expansion $\big|g\big(\frac{j+1}{a}\big)-g\big(\frac{j}{a}\big)-\sum_{k=1}^{n}\frac{1}{a^k k!}g^{(k)}\big(\frac{j}{a}\big)\big|\le\frac{1}{a^{n+1}(n+1)!}\sup_{x\in[0,1]}|g^{(n+1)}(x)|$ for $j\in\{0,1,\ldots,a-1\}$, we obtain:

$$I_a(\lambda)\ \le\ \sum_{k=1}^{n}\frac{1}{a^k k!}\Big|\frac{1}{a}\sum_{j=1}^{a-1}g^{(k)}\big(\frac{j}{a}\big)S_j(a,\lambda)\Big|+\frac{1}{a^{n+1}(n+1)!}\sup_{x\in[0,1]}|g^{(n+1)}(x)|.$$

From the definition of $S_j(a,\lambda)$ and with the inequality $\frac{2}{\pi}u\le\sin(u)\le u$ for $u\in[0,\pi/2]$, we have for $\lambda\in[-a\pi,0)\cup(0,a\pi]$ and $k\in\{1,\ldots,n\}$:

$$\begin{aligned}\Big|\frac{1}{a}\sum_{j=1}^{a-1}g^{(k)}\big(\frac{j}{a}\big)S_j(a,\lambda)\Big| &\le \frac{1}{2|\sin(\lambda/2a)|}\Big(\Big|\frac{1}{a}\sum_{j=1}^{a-1}g^{(k)}\big(\frac{j}{a}\big)\Big|+\Big|\frac{1}{a}\sum_{j=1}^{a-1}g^{(k)}\big(\frac{j}{a}\big)e^{-i\lambda\frac{j}{a}}\Big|\Big) \\ &\le \frac{\pi a}{2|\lambda|}\Big(\frac{1}{a^{n+1-k}(n+1-k)!}\sup_{x\in[0,1]}|g^{(n+1)}(x)|+D_{g^{(k)}}(n+1-k)\frac{1}{|\lambda|^{n+1-k}}\Big),\end{aligned}$$

using (20) for bounding $\frac{1}{a}\sum_{j=1}^{a-1}g^{(k)}\big(\frac{j}{a}\big)$ and the induction hypothesis for bounding $\frac{1}{a}\sum_{j=1}^{a-1}g^{(k)}\big(\frac{j}{a}\big)e^{-i\lambda\frac{j}{a}}$.

Hence, with (23),

$$\begin{aligned}I_a(\lambda)+\frac{1}{a}\big|g\big(\frac{1}{a}\big)\big| &\le \frac{1}{a^{n+1}}\sup_{x\in[0,1]}|g^{(n+1)}(x)|\sum_{k=0}^{n+1}\frac{1}{(n+1-k)!\,k!}+\frac{\pi a}{2|\lambda|}\sup_{x\in[0,1]}|g^{(n+1)}(x)|\sum_{k=1}^{n}\frac{10^{n+1-k}}{a^k k!}\frac{1}{|\lambda|^{n+1-k}} \\ &\le \frac{(2\pi)^{n+1}}{(n+1)!\,|\lambda|^{n+1}}\sup_{x\in[0,1]}|g^{(n+1)}(x)|+\frac{10^{n+1}}{|\lambda|^{n+1}}\sup_{x\in[0,1]}|g^{(n+1)}(x)|\sum_{k=1}^{n}\frac{1}{k!}\big(\frac{\pi}{10}\big)^k \qquad (24) \\ &\le \frac{10^{n+1}}{|\lambda|^{n+1}}\sup_{x\in[0,1]}|g^{(n+1)}(x)|\sum_{k=1}^{n+1}\frac{1}{k!}\big(\frac{\pi}{5}\big)^k \\ &\le \frac{10^{n+1}}{|\lambda|^{n+1}}\sup_{x\in[0,1]}|g^{(n+1)}(x)|\,\big(e^{\pi/5}-1\big), \qquad (25)\end{aligned}$$

since $a^{-k}\le\pi^k\,|\lambda|^{-k}$ for all $\lambda\in[-a\pi,0)\cup(0,a\pi]$ and $k\in\{0,1,\ldots,n+1\}$. Thus since $e^{\pi/5}-1<1$ and from (22) and (25), we deduce that (21) is true for $k=n+1$ and therefore for any $k\in\mathbb{N}$. □

*Proof of Property 1.* Since $(X_t)_{t\in\mathbb{Z}}$ being a stationary centered linear process, $e(a,b)=\sum_{j=1}^{a}\Big(\frac{1}{\sqrt{a}}\,\psi\big(\frac{j}{a}\big)\Big)X_{b+j}$ for any $(a,b)\in\mathbb{N}^*\times\mathbb{Z}$ from (4) and $\sum_{j=1}^{a}\frac{1}{\sqrt{a}}\big|\psi\big(\frac{j}{a}\big)\big|<\infty$, it is obvious that for $a\in\mathbb{N}^*$, $(e(a,b))_{b\in\mathbb{Z}}$ is a stationary centered linear process.

With computations similar to those performed in Bardet *et al.* (2008) [Proof of Property 1], we obtain with

$f$ the spectral density of $X$ and for $a \in \mathbb{N}^*$,

$$\mathbb{E}(e^2(a,0)) \;=\; \int_{-a\pi}^{a\pi} f\big(\frac{u}{a}\big) \times \Big|\frac{1}{a}\sum_{j=1}^{a} \psi\big(\frac{j}{a}\big) e^{-i\frac{j}{a}u}\Big|^2\, du.$$

Now, since $\psi$ satisfies Assumption $\Psi(k)$, from Lemma 1, for $a$ large enough and $u \in [-\sqrt{a}, \sqrt{a}]$, we obtain from (3):

$$\begin{aligned}\Big|\big|\frac{1}{a}\sum_{j=1}^{a} \psi\big(\frac{j}{a}\big) e^{-i\frac{j}{a}u}\big|^2 - |\widehat{\psi}(u)|^2\Big| &\le 2C_\psi(k)\frac{|u|^k}{a^k}\,|\widehat{\psi}(u)| + C_\psi^2(k)\frac{|u|^{2k}}{a^{2k}} \\ &\le \Big(2C_\psi(k)\sup_{x\in[0,1]}|\psi^{(k)}(x)| + C_\psi^2(k)\Big)\,\frac{1}{a^k}, \qquad (26)\end{aligned}$$

Moreover, for $|u| \in [\sqrt{a}\,,\, a\pi]$, from Lemma 2 and $a \in \mathbb{N}^*$, we have:

$$\Big|\big|\frac{1}{a}\sum_{j=1}^{a} \psi\big(\frac{j}{a}\big) e^{-i\frac{j}{a}u}\big|^2 \;\le\; D_\psi^2(k)\frac{1}{|u|^{2k}}, \qquad (27)$$

Given the existence of

$c_f > 0$ satisfying $f(\lambda) \le c_f|\lambda|^{-2d}$ for all $\lambda \in [-\pi,\pi]$, together with (26) and (27) we obtain with $F_\psi(k) = 2C_\psi(k)\sup_{x\in[0,1]}|\psi^{(k)}(x)| + C_\psi^2(k)$ and for all $d < 1/2$,

$$\begin{aligned}\Big|\mathbb{E}(e^2(a,0)) - \int_{-\sqrt{a}}^{\sqrt{a}} f\big(\frac{u}{a}\big) \times |\widehat{\psi}(u)|^2\,du\Big| &\le \frac{F_\psi(k)}{a^k}\int_{-\sqrt{a}}^{\sqrt{a}} f\big(\frac{u}{a}\big)du + 2D_\psi^2(k)\int_{\sqrt{a}}^{a\pi}\frac{1}{|u|^{2k}}f\big(\frac{u}{a}\big)\,du \\ &\le a^{2d}\Big(\frac{2c_f\,F_\psi(k)}{1-2d} + \frac{2D_\psi^2(k)}{2k+2d-1}\Big)\frac{1}{a^{k+d-1/2}}. \qquad (28)\end{aligned}$$

Using again (3), for $a$ large enough, we have :

$$\begin{aligned}\Big|\int_{-\sqrt{a}}^{\sqrt{a}} f\big(\frac{u}{a}\big)\,|\widehat{\psi}(u)|^2\,du - \int_{-\infty}^{\infty} f\big(\frac{u}{a}\big)\,|\widehat{\psi}(u)|^2\,du\Big| &\le \big(2c_f\sup_{x\in[0,1]}|\psi^{(k)}(x)|\big)a^{2d}\int_{\sqrt{a}}^{\infty}\frac{1}{u^{2d+2k}}\,du \\ &\le a^{2d}\Big(\frac{2c_f\sup_{x\in[0,1]}|\psi^{(k)}(x)|}{2k+2d-1}\Big)\frac{1}{a^{k+d-1/2}}. \qquad (29)\end{aligned}$$

So, from Assumption A$(d,d')$, we obtain the following expansion:

$$\begin{aligned}\int_{-\infty}^{\infty} f\big(\frac{u}{a}\big)\,|\widehat{\psi}(u)|^2\,du &= 2\pi\int_{-\infty}^{\infty}\Big(c_d\big|\frac{u}{a}\big|^{-2d} + c_{d'}\big|\frac{u}{a}\big|^{d'-2d} + \big|\frac{u}{a}\big|^{d'-2d}\varepsilon\big(\frac{u}{a}\big)\Big)\,|\widehat{\psi}(u)|^2\,du \\ &= 2\pi\,c_d\,K_{(\psi,2d)}\,a^{2d} + 2\pi\,c_{d'}\,K_{(\psi,2d-d')}\,a^{2d-d'} + o(a^{2d-d'}) \qquad (30)\end{aligned}$$

Definition (6) of $K_{(\psi,\alpha)}$ ($\lim_{\lambda\to 0}\varepsilon(\lambda) = 0$) as well as Lebesgue Theorem and (28), (29) and (30), we find that $C$ exists only depending on $\psi$ and $k$ such as for $a$ large enough, we have:

$$\Big|\mathbb{E}(e^2(a,0)) - 2\pi\,c_d\,K_{(\psi,2d)}\,a^{2d} - 2\pi\,c_{d'}\,K_{(\psi,2d-d')}\,a^{2d-d'}\Big| \;\le\; a^{2d}\,\big(C\,a^{-k-d+1/2} + o(a^{2d-d'})\big). \qquad (31)$$

When $k \ge d' - d + 1/2$ implying $k + d - 1/2 > d'$, then (5) holds. $\square$

*Proof of Theorem 1.* We decompose this proof into 4 steps. First define the normalized wavelet coefficients of $X$ by:

$$\widetilde{e}_N(a,b) := \frac{e(a,b)}{\sqrt{\mathbb{E}(e^2(a,0))}} \quad \text{for } a \in \mathbb{N}^* \text{ and } b \in \mathbb{Z}, \tag{32}$$

and the normalized sample variance of wavelet coefficients by:

$$\widetilde{T}_N(a) := \frac{1}{N-a}\sum_{k=1}^{N-a} \widetilde{e}^2(a,k). \tag{33}$$

**Step 1** We prove that $N\text{Cov}\,(\widetilde{T}_N(r\,a_N), \widetilde{T}_N(r'\,a_N))$ converges to the asymptotic covariance matrix $\Gamma(r_1,\dots,r_\ell,\psi,d)$ defined in (10). First for $\lambda \in \mathbb{R}$, denote

$$S_a(\lambda) := \frac{1}{a}\sum_{t=1}^{a} \psi(\frac{t}{a}) e^{i\lambda t/a}.$$

Then for $a \in \mathbb{N}^*$ and $b = 1, \cdots, N-a$, since $\psi$ is a $[0,1]$-supported function and $\widehat{\alpha} \in \mathbb{L}^2([-\pi,\pi])$ inducing $\alpha(k) = \int_{-\pi}^{\pi} \widehat{\alpha}(\lambda) e^{ik\lambda} d\lambda$, we have:

$$\begin{aligned}\sum_{t=1}^{N} \alpha(t-s)\psi(\frac{t-b}{a}) &= \sum_{t=0}^{a} \psi(\frac{t}{a}) \int_{-\pi}^{\pi} \widehat{\alpha}(\lambda) e^{i\lambda(t-s+b)} d\lambda \\ &= \int_{-\pi}^{\pi} a S_a(a\lambda)\widehat{\alpha}(\lambda) e^{i(b-s)\lambda} d\lambda \\ &= \int_{-a\pi}^{a\pi} S_a(\lambda)\widehat{\alpha}(\frac{\lambda}{a}) e^{i(b-s)\frac{\lambda}{a}} d\lambda. \end{aligned} \tag{34}$$

But,

$$\begin{aligned}\text{Cov}\,(\widetilde{T}_N(a), \widetilde{T}_N(a')) &= \frac{1}{N-a}\frac{1}{N-a'}\sum_{b=1}^{N-a}\sum_{b'=1}^{N-a'} \text{Cov}\,(\widetilde{e}^2(a,b), \widetilde{e}^2(a',b')) \\ &= \frac{(\mathbb{E}(e^2(a,0))\mathbb{E}(e^2(a',0)))^{-1}}{4\pi^2(N-a)(N-a')}\sum_{b=1}^{N-a}\sum_{b'=1}^{N-a'} \text{Cov}\,(e^2(a,b), e^2(a',b')). \end{aligned} \tag{35}$$

and,

$$\begin{aligned}\text{Cov}\,(e^2_{(a,b)}, e^2_{(a',b')}) &= \frac{1}{a\,a'}\sum_{t_1,t_2,t_3,t_4=1}^{N}\sum_{s_1,s_2,s_3,s_4\in\mathbb{Z}} \Big(\prod_{i=1}^{2}\alpha(t_i-s_i)\psi(\frac{t_i-b}{a})\Big)\Big(\prod_{i=1}^{2}\alpha(t_i-s_i)\psi(\frac{t_i-b'}{a'})\Big)\text{Cov}\,\big(\xi_{s_1}\xi_{s_2}, \xi_{s_3}\xi_{s_4}\big) \\ &= C_1 + C_2, \end{aligned} \tag{36}$$

since there are only two nonvanishing cases , i.e. $s_1 = s_2 = s_3 = s_4$ (Case 1 => $C_1$), $s_1 = s_3 \neq s_2 = s_4$ and $s_1 = s_4 \neq s_2 = s_3$ (Case 2 => $C_2$).

*Case 1:* in such a case, $\text{Cov}\,\big(\xi_{s_1}\xi_{s_2}, \xi_{s_3}\xi_{s_4}\big) = \mu_4 - 1$ and

$$C_1 = \frac{\mu_4-1}{a\,a'}\sum_{s\in\mathbb{Z}}\Big|\sum_{t=1}^{N}\alpha(t-s)\psi(\frac{t-b}{a})\Big|^2\Big|\sum_{t=1}^{N}\alpha(t-s)\psi(\frac{t-b'}{a'})\Big|^2$$

$$C_1 = (\mu_4-1)\,a\,a' \lim_{M\to\infty}\int_{[-\pi,\pi]^4} d\lambda d\lambda' d\mu d\mu' e^{i[b(\lambda-\lambda')+b'(\mu-\mu')]}$$
$$\times \sum_{s=-M}^{M} e^{is[(\lambda-\lambda')+(\mu-\mu')]} S_a(a\lambda)\widehat{\alpha}(\lambda)\overline{S_a(a\lambda')\widehat{\alpha}(\lambda')} S_{a'}(a'\mu)\widehat{\alpha}(\mu)\overline{S_{a'}(a'\mu')\widehat{\alpha}(\mu')}$$

using (34)( $\overline{z}$ denoting the conjugate of $z \in \mathbb{C}$). From the usual asymptotic behavior of Dirichlet kernel, for $g$ a $2\pi$-periodic function such as $g \in \mathcal{C}^1((-\pi,\pi))$, we have $\lim_{M\to\infty} \int_{-\pi}^{\pi} D_M(z) g(x+z) dz = g(x)$ uniformly in $x$ with

$$D_M(z) := \frac{1}{2\pi} \sum_{k=-M}^{M} e^{ikz} = \frac{1}{2\pi} \frac{\sin\big((2M+1)z/2\big)}{\sin\big(z/2\big)}. \tag{37}$$

Thus with $h : \mathbb{R}^4 \mapsto \mathbb{R}$ a continuously differentiable function $2\pi$-periodic for each component,

$$\lim_{M\to\infty} \int_{[-\pi,\pi]^4} 2\pi\, D_M((\lambda-\lambda')+(\mu-\mu'))h(\lambda,\lambda',\mu,\mu')d\lambda d\lambda' d\mu d\mu' = 2\pi \int_{[-\pi,\pi]^3} h(\lambda'-\mu+\mu',\lambda',\mu,\mu')d\lambda' d\mu d\mu';$$

Therefore, we have:

$$C_1 = 2\pi\,(\mu_4-1)\,a\,a' \int_{[-\pi,\pi]^3} d\lambda' d\mu d\mu' e^{i(\mu-\mu')(b'-b)}$$
$$\times\, S_a(a(\lambda'-\mu+\mu'))\widehat{\alpha}(\lambda'-\mu+\mu')\overline{S_a(a\lambda')\widehat{\alpha}(\lambda')} S_{a'}(a'\mu)\widehat{\alpha}(\mu)\overline{S_{a'}(a'\mu')\widehat{\alpha}(\mu')}. \tag{38}$$

* *Case 2:* in such a case, with $s_1 \neq s_2$, Cov $\big(\xi_{s_1}\xi_{s_2}, \xi_{s_1}\xi_{s_2}\big) = 1$ using the asymptotic behaviors of two Dirichlet kernels, we have:

$$C_2 = \frac{2}{a\,a'} \sum_{(s,s')\in\mathbb{Z}^2, s\neq s'} \sum_{t_1=1}^{N} \alpha(t_1-s)\psi(\frac{t_1-b}{a}) \sum_{t_2=1}^{N} \alpha(t_2-s)\psi(\frac{t_2-b'}{a'}) \sum_{t_3=1}^{N} \alpha(t_3-s')\psi(\frac{t_3-b}{a}) \sum_{t_4=1}^{N} \alpha(t_4-s')\psi(\frac{t_4-b'}{a'})$$
$$= -\frac{2C_1}{\mu_4-1} + \frac{1}{a\,a'} \sum_{(s,s')\in\mathbb{Z}^2} \sum_{t_1=1}^{N} \alpha(t_1-s)\psi(\frac{t_1-b}{a}) \sum_{t_2=1}^{N} \alpha(t_2-s)\psi(\frac{t_2-b'}{a'}) \sum_{t_3=1}^{N} \alpha(t_3-s')\psi(\frac{t_3-b}{a}) \sum_{t_4=1}^{N} \alpha(t_4-s')\psi(\frac{t_4-b'}{a'})$$
$$C_2 = -\frac{2C_1}{\mu_4-1} + 2\,a\,a' \lim_{M\to\infty} \lim_{M'\to\infty} \int_{[-\pi,\pi]^4} d\lambda d\lambda' d\mu d\mu' e^{i[b(\lambda-\mu)-b'(\lambda'-\mu')]}$$
$$\times \sum_{s=-M}^{M} \sum_{s=-M'}^{M'} e^{is(\lambda'-\lambda)+is'(\mu'-\mu)} S_a(a\lambda)\widehat{\alpha}(\lambda)\overline{S_{a'}(a'\lambda')\widehat{\alpha}(\lambda')} S_a(a\mu)\widehat{\alpha}(\mu)\overline{S_{a'}(a'\mu')\widehat{\alpha}(\mu')}$$
$$= -\frac{2C_1}{\mu_4-1} + 8\pi^2\,a\,a' \int_{[-\pi,\pi]^2} e^{i(\lambda-\mu)(b-b')} S_a(a\lambda)\overline{S_{a'}(a'\lambda)} S_a(a\mu)\overline{S_{a'}(a'\mu)} \times \big|\widehat{\alpha}(\lambda)\big|^2 \big|\widehat{\alpha}(\mu)\big|^2 d\lambda d\mu,$$

Compute $\sum_{b=1}^{N-a} \sum_{b'=1}^{N-a'} (C_1+C_2)$. For both ($C_1$ and $C_2$), a Dirichlet kernel function is confirmed as follows:

$$F_N(a,a',v) := \sum_{b=1}^{N-a} \sum_{b'=1}^{N-a'} e^{i\,v\,(b-b')} = e^{iv(a-a')/2} \frac{\sin((N-a)v/2)\,\sin((N-a')v/2)}{\sin^2(v/2)}. \tag{39}$$

For a continuous function $h : [-\pi,\pi] \mapsto \mathbb{R}$,

$$\lim_{N\to\infty} \frac{1}{N} \int_{-\pi}^{\pi} h(v) F_N(a,a',v) dv = \lim_{N\to\infty} \frac{1}{N^2} \int_{-\pi N}^{\pi N} h(\frac{v}{N}) F_N(a,a',\frac{v}{N}) dv = 4h(0) \int_{-\infty}^{\infty} \frac{\sin^2(v/2)}{v^2} dv = 2\pi h(0),$$

The Lebesgue Theorem a, $a/N \to 0$ $(N \to 0)$and (38) give us:

$$N \, \frac{1}{N-a} \frac{1}{N-a'} \sum_{b=1}^{N-a} \sum_{b'=1}^{N-a'} C_1 \sim 4\pi^2\,(\mu_4 - 1) a a' \int_{[-\pi,\pi]^2} d\lambda' d\mu' |S_a(a\lambda')|^2 \, |S_{a'}(a'\mu')|^2 \, |\widehat{\alpha}(\lambda')|^2 |\widehat{\alpha}(\mu')|^2$$

$$\sim 4\pi^2\,(\mu_4 - 1) \int_{-a\pi}^{a\pi} |S_a(\lambda)|^2 \, |\widehat{\alpha}(\lambda/a)|^2 d\lambda \int_{-a'\pi}^{a'\pi} |S_a(\mu)|^2 \, |\widehat{\alpha}(\mu/a')|^2 d\mu$$

$$\Longrightarrow N \, \frac{(\mathbb{E}(e^2(a,0)))^{-1} \mathbb{E}(e^2(a',0)))^{-1}}{4\pi^2(N-a)(N-a')} \sum_{b=1}^{N-a} \sum_{b'=1}^{N-a'} C_1 \underset{N\to\infty}{\longrightarrow} (\mu_4 - 1) \tag{40}$$

and therefore $$\frac{N}{a_N} \, \frac{(ra_N r' a_N)^{-2d} (c_d K_{(\psi,2d)})^{-2}}{4\pi^2(N - ra_N)(N - r'a_N)} \sum_{b=1}^{N-ra_N} \sum_{b'=1}^{N-r'a_N} C_1 \underset{N\to\infty}{\longrightarrow} 0, \tag{41}$$

with $a = ra_N$ and $a' = r'a_N$, using 1 since $a_N \to \infty$.

Moreover, taking again $a_N \to \infty$ and $N/a_N \to \infty$, we have:

$$N \, \frac{1}{N-a} \frac{1}{N-a'} \sum_{b=1}^{N-a} \sum_{b'=1}^{N-a'} C_2 \sim 16\pi^3 a a' \int_{-\pi}^{\pi} \big|S_a(a\lambda)\big|^2 \big|S_{a'}(a'\lambda)\big|^2 \big|\widehat{\alpha}(\lambda)\big|^4 d\lambda - \frac{2N}{\mu_4 - 1} \frac{1}{N-a} \frac{1}{N-a'} \sum_{b=1}^{N-a} \sum_{b'=1}^{N-a'} C_1$$

$$\sim 16\pi^3 r r' a_N \int_{-a_N\pi}^{a_N\pi} \big|S_{ra_N}(r\lambda)\big|^2 \big|S_{r'a_N}(r'\lambda)\big|^2 \big|\widehat{\alpha}(\lambda/a_N)\big|^4 d\lambda - \frac{2N}{\mu_4 - 1} \frac{1}{N - ra_N} \frac{1}{N - r'a_N} \sum_{b=1}^{N-ra_N} \sum_{b'=1}^{N-r'a_N} C_1$$

$$\Longrightarrow \frac{N}{a_N} \, \frac{(r\,r'\,a_N^2)^{-2d} (c_d K_{(\psi,2d)})^{-2}}{4\pi^2(N - ra_N)(N - r'a_N)} \sum_{b=1}^{N-ra_N} \sum_{b'=1}^{N-r'a_N} C_2 \underset{N\to\infty}{\longrightarrow} 4\pi \frac{(rr')^{1-2d}}{K^2_{(\psi,2d)}} \int_{-\infty}^{\infty} \frac{\big|\widehat{\psi}(r\lambda)\big|^2 \big|\widehat{\psi}(r'\lambda)\big|^2}{\lambda^{4d}} d\lambda,$$

Since $a_N \to \infty$ and $N/a_N \to \infty$, using Property 1 and (41), we have:

$$\frac{N}{a_N} \operatorname{Cov}\big(\widetilde{T}_N(r\,a_N), \widetilde{T}_N(r'\,a_N)\big) \underset{N\to\infty}{\longrightarrow} 4\pi \frac{(rr')^{1-2d}}{K^2_{(\psi,2d)}} \int_{-\infty}^{\infty} \frac{\big|\widehat{\psi}(r\lambda)\big|^2 \big|\widehat{\psi}(r'\lambda)\big|^2}{\lambda^{4d}} d\lambda. \tag{42}$$

Note that if $r = r'$ then $\dfrac{N}{r\,a_N} \operatorname{Var}\big(\widetilde{T}_N(r\,a_N)\big) \underset{N\to\infty}{\longrightarrow} \sigma^2_\psi(d) = 64\pi^5 \, \dfrac{K_{(\psi*\psi,4d)}}{K^2_{(\psi,2d)}}$ depending only on $\psi$ and $d$.

**Step 2** Consequently if the distribution of the innovations $(\xi_t)_t$ is such that it exists $r > 0$ satisfying $\mathbb{E}\big(e^{r\xi_0}\big) \leq \infty$ (the so-called the Cramèr condition), then for any $a \in \mathbb{N}^*$, $(\widetilde{T}_N(r_i\,a_N))_{1\leq i\leq \ell} = \Big(\frac{1}{N - r_i a_N} \sum_{k=1}^{N-r_i a_N} \widetilde{e}^2(r_i a_N, k)\Big)_{1\leq i\leq \ell}$ satisfies a central limit theorem.

Such theorem holds if it can be proved that $\sqrt{\frac{N}{a_N}} \sum_{i=1}^{\ell} \frac{u_i}{N - r_i a_N} \sum_{k=1}^{N-r_i a_N} \widetilde{e}^2(r_i a_N, k)$ asymptotically follows a Gaussian distribution for any vector $(u_i)_{1\leq i\leq \ell} \in \mathbb{R}^\ell$.

This result is based on an adaptation demonstration of Giraitis (1985)( Appell polynomials decomposition allows to prove central limit theorems for function of linear process). $X$ being a two-sided linear process, martingale type results as in Wu (2002) or Furmanczyk (2007) cannot be used. Moreover, $(a_N)_N$ being a sequence depending on $N$, the central limit theorem for triangular arrays has yet to be proved. As far as we are concerned, the paper of Roueff and Taqqu (2009)(dealing with central limit theorem for triangular arrays of decimated linear process) can be applied to establish a multidimensional central limit for the variogram of wavelet coefficients associated to a multi-resolution analysis, however, it cannot be used in our case.Because the present variogramm is defined as in (8) with coefficients taken every $n/n_j$ ($\simeq a_N$ with our notation) and mean value of $n_j$ ($N/a_N$ with our notation) coefficients (with a convergence rate $\sqrt{n_j}$). Hence, we consider in the present case wavelet coefficient variogram (7) being an average of $N-a_N$ terms with a convergence rate is $N/a_N$. and then adapt it to the method and results of Giraitis (1985).

Consider the case $\ell=1$. For $a>0$, $(\widetilde{e}(a,b))_{1\leq b\leq N-a}$ is a stationary linear process satisfying the assumptions of the paper of Giraitis (refered as to $X_t$). Supposing $H_2(x)=x^2-1$ the second-order Hermite polynomial , we will prove that:

$$\Big(\frac{N}{a_N}\Big)^{1/2}\frac{1}{N-a_N}\sum_{b=1}^{N-a_N}\big(\widetilde{e}^2(a_N,b)-1\big)\simeq\Big(\frac{1}{Na_N}\Big)^{-1/2}\sum_{b=1}^{N-a_N}H_2(\widetilde{e}(a_N,b))\ \xrightarrow[N\to\infty]{\mathcal{L}}\ \mathcal{N}\big(0,\sigma^2_\psi(d)\big).$$

The distribution of $\xi_0$ being supposed to satisfy the Cramèr condition and refereing to the proof of Proposition 6 (Giraitis, 1985), we define $S_N^{(n)}=\sum_{b=1}^{N-a_N}A_n^{(a_N)}(\widetilde{e}(a_N,b))$ where $A_n^{(a_N)}$ is the Appell polynomial of degree $n$ corresponding to the probability distribution of $\widetilde{e}(a_N,\cdot)$. We can than prove that the cumulants of order $k\geq 3$ are such as

$$\chi\big(S_N^{(n(1))},\dots,S_N^{(n(k))}\big)=o\big((Na_N)^{k/2}\big)\tag{43}$$

for any $n(1),\cdots,n(k)\geq 2$ (the computation of the cumulants of order 2 is induced by Step 1 of this proof). Indeed, $\chi\big(S_N^{(n(1))},\dots,S_N^{(n(k))}\big)=\sum_{\gamma\in\Gamma_0(T)}d_\gamma I_\gamma(N)$ where $\Gamma_0(T)$ is the set of possible diagrams (for the definition of $I_\gamma(N)$ see (34) of Giraitis (1985)).

In the case of Gaussian diagrams, $I_\gamma(N)=o\big((Na_N)^{k/2}\big)$ is the result of gaussian case and the second order moments.

If $\gamma$ , however is a non-Gaussian diagram, *mutatis mutandis*, we use the notation and proof of Lemma 2 of Giraitis (1985). From Step 1, we obtain:

$$\widetilde{e}(a,b)=\sum_{s\in\mathbb{Z}}\beta_a(b-s)\,\xi_s\quad\text{with}\quad\beta_a(s)=\frac{\sqrt{a}}{\sqrt{\mathbb{E}e^2(a,b)}}\int_{-\pi}^{\pi}S_a(a\lambda)\widehat{\alpha}(\lambda)e^{i\lambda s}d\lambda.\tag{44}$$

Then for $u\in[-\pi,\pi]$,

$$\begin{aligned}\widehat{\beta}_a(u)&=\frac{1}{2\pi}\sum_{s=-\infty}^{\infty}\beta_a(s)e^{-isu}\\&=\frac{\sqrt{a}}{2\pi\sqrt{\mathbb{E}e^2(a,b)}}\lim_{m\to\infty}\int_{-\pi}^{\pi}\sum_{s=-m}^{m}S_a(a\lambda)\widehat{\alpha}(\lambda)e^{is(\lambda-u)}d\lambda\\&=\frac{\sqrt{a}}{\sqrt{\mathbb{E}e^2(a,b)}}S_a(au)\widehat{\alpha}(u),\end{aligned}$$

with the asymptotic behavior of Dirichlet kernel. Now, in the case a/ of Lemma 2 of Giraitis (1985), take diagram $V_1=\{(1,1),(2,1),(3,1)\}$ and assume that for rows $L_j$ of array $T$, $j=1,\cdots,k$ ($k\geq 3$), $|V_1\cap L_j|\geq 1$

for at least 3 different rows $L_j$. If we then replicate inequality (39), assume hyperplane $x_{V_1}$, a part of the integral (34) provides:

$$\Big|\int_{\{x_{11}+x_{21}+x_{31}=0\}\cap[-\pi,\pi]^3} dx_{11}dx_{21}dx_{31}\prod_{j=1}^{3} D_N((x_{j1}+\cdots+x_{jn(j)})\widehat{\beta}_a(x_{j1})\Big| \leq C\,\alpha_1(u_1)\alpha_2(u_2)\alpha_3(u_3),$$

with $u_i = x_{i2}+\cdots+x_{in(i)}$ and the same expressions of $\alpha_i$ provided in Giraitis (1985). We finally have to bound $\alpha_i(u)$. Taking the same approximations as in the proof of Property 1, for $a_N$ and $N$ large enough, we have:

$$\begin{aligned}
\alpha_1(u) &= \int_{-\pi}^{\pi}\big|\widehat{\beta}_{a_N}(u)\,D_N(x+u)\big|dx \sim \sqrt{2\pi}\,\frac{1}{\sqrt{a_N}}\int_{-a_N\pi}^{a_N\pi}\big|\frac{\widehat{\psi}(x)}{|x|^d}\big|\,\big|D_N\big(\frac{x}{a_N}+u\big)\big|du \\
&\leq 2\sqrt{a_N}\sup_{x\in\mathbb{R}}\Big\{\frac{|\widehat{\psi}(x)|}{|x|^d}\Big\}\int_{-\pi}^{\pi}|D_N(x+u)|dx \\
&\leq 2C\sup_{x\in\mathbb{R}}\Big\{\frac{|\widehat{\psi}(x)|}{|x|^d}\Big\}\sqrt{a_N}\,\log N,
\end{aligned}$$

since $C>0$ exists such as $\int_{-\pi}^{\pi}|D_N(x+u)|dx \leq C\log N$ for any $u\in[-\pi,\pi]$. upposinf S $i=2,3$, $a_N$ and $N$ large enough, we have:

$$\begin{aligned}
\alpha_i^2(u) &= \|\widehat{\beta}_{a_N}(\cdot)\,D_N(u+\cdot)\|_2^2 \\
&\leq 2\int_{-a_N\pi}^{a_N\pi}\frac{|\widehat{\psi}(x)|^2}{|x|^{2d}}D_N^2\big(\frac{x}{a_N}+u\big)du \\
&\leq 2C\sup_{x\in\mathbb{R}}\Big\{\frac{|\widehat{\psi}(x)|^2}{|x|^{2d}}\Big\}\,a_N\int_{-\pi}^{\pi}|D_N^2(x+u)|dx \\
&\leq C'\sup_{x\in\mathbb{R}}\Big\{\frac{|\widehat{\psi}(x)|^2}{|x|^{2d}}\Big\}\,Na_N.
\end{aligned}$$

Then $\alpha_1(u_1)\alpha_2(u_2)\alpha_3(u_3) = o((Na_N)^{3/2})$.

For the $k-3$ other terms, a result corresponding to Lemma 1 of Giraitis (1985) can also be obtained. If, for $a_N$ and $N$ large enough,

$$\begin{aligned}
\|g_{N,j}\|_2^2 &= \int_{[-\pi,\pi]^{n(j)}} dx\,D_N^2(x_1+\cdots+x_{n(j)})\prod_{i=1}^{n(j)}|\widehat{\beta}_{a_N}(x_i)|^2 \\
&\leq C\int_{[-a_N\pi,a_N\pi]^{n(j)}} dx\,D_N^2\big(\frac{1}{a_N}(x_1+\cdots+x_{n(j)})\big)\prod_{i=1}^{n(j)}\frac{|\widehat{\psi}(x_i)|^2}{|x_i|^{2d}} \\
&\leq C\,\big|\sup_{x\in\mathbb{R}}\big\{\frac{|\widehat{\psi}(x)|^2}{|x|^{2d}}\big\}\big|^{n(j)}\,a_N\,\|D_N(\,\cdot\,)\|_2^2 \\
&\leq C'\,Na_N
\end{aligned}$$

with $C'\geq 0$ independent on $N$ and $a_N$. We Thus obtain $\|g_{N,j}\|_2 \leq C\,(Na_N)^{1/2}$ with $C\geq 0$. Furthermore, $C'\geq 0$ exists such as $\|g'_{N,j}\|_2 \leq C\,(Na_N)^{1/2}$ for $j\geq 2$ while $\|g'_{N,1}\|_2 = O(\sqrt{a_N}\,\log N) = o((Na_N)^{1/2})$. Consequently, if $\gamma$ such as $|V_1\cap L_j|\geq 1$ for at least 3 different rows $L_j$, and more generally with $|V_1|\geq 3$, we have:

$$I_\gamma(N) = o\big((Na_N)^{k/2}\big). \qquad (45)$$

For further $\gamma$, we need to bound the function $h(u_1,u_2)$ as defined in Giraitis (1985, p. 32) as follows (with $x=x_{11}+x_{12}$) and with $u_1+u_2\neq 0$:

$$\begin{aligned} h(u_1,u_2) &= \Big(\int_{-\pi}^{\pi}\big|\widehat{\beta}_{a_N}(-x)\,D_N(u_1+x)\,D_N(u_2-x)\big|dx\Big)\,\Big(\int_{-\pi}^{\pi}\big|\widehat{\beta}_{a_N}(x)\big|^2dx\Big) \\ &\leq \big|\sup_{x\in\mathbb{R}}\big\{\frac{|\widehat{\psi}(x)|^2}{|x|^{2d}}\big\}\big|\,a_N\,\Big(\int_{-\pi}^{\pi}\big|D_N\big(u_1+x\big)\,D_N\big(u_2-x\big)\big|dx\Big)\Big(2\pi\int_{-\infty}^{\infty}\frac{|\widehat{\psi}(x)|^2}{|x|^{2d}}\,dx\Big). \end{aligned}$$

But

$$\begin{aligned} \int_{-\pi}^{\pi}\big|D_N\big(u_1+x\big)\,D_N\big(u_2-x\big)\big|dx &\leq 2\int_{-2\pi N}^{2\pi N}\Big|\frac{\sin(x)}{x}\,\frac{\sin(\frac{N}{2}(u_1+u_2)-x)}{\sin(\frac{1}{2}(u_1+u_2)-\frac{x}{N})}\Big|dx \\ &\leq \begin{cases} C\,\log N\big|\sin(\frac{1}{2}(u_1+u_2))\big|^{-1} & \text{if } |u_1+u_2|\geq (N\log N)^{-1} \\ C\,N & \text{if } |u_1+u_2|<(N\log N)^{-1} \end{cases}. \end{aligned}$$

Therefore,

$$\begin{aligned} \|h(u_1,u_2)\|_2^2=\int_{[-\pi,\pi]^2}h^2(u_1,u_2)du_1du_2 &\leq C\,a_N^2\,\Big(\log^2 N\int_{(N\log N)^{-1}}^{\pi}(\sin x)^{-2}\,dx+N^2\int_0^{(N\log N)^{-1}}dx\Big) \\ &\leq C\,a_N^2\,\big(N\log^3 N+N\log N\big), \end{aligned}$$

and hence $\|h(u_1,u_2)\|_2=o(Na_N)$. Finally, (45) holds for all $\gamma$ and it implies (43).

If $\ell>1$, the same proof can be replicated with the linearity properties of cumulants. Thus, $(\widetilde{T}_N(r_i\,a_N))_{1\leq i\leq\ell}$ satisfies the following central limit:

$$\sqrt{\frac{N}{a_N}}\,\big(\widetilde{T}_N(r_i\,a_N)-1\big)_{1\leq i\leq\ell}\xrightarrow[N\to\infty]{d}\mathcal{N}\big(0\,,\,\Gamma(r_1,\ldots,r_\ell,\psi,d)\big), \tag{46}$$

with $\Gamma(r_1,\ldots,r_\ell,\psi,d)=(\gamma_{ij})_{1\leq i,j\leq\ell}$ given in (10).

**Step 3** With the truncation procedure, we can now we extend the central limit obtained in Step 2 (for linear processes with an innovation distribution satisfying a Cramèr condition ($\mathbb{E}\big(e^{r\xi_0}\big)<\infty$)) to the weaker condition $\mathbb{E}\xi_0^4<\infty$. Take $\mathbb{E}\xi_0^4<\infty$. Let $M>0$ and define $\xi_t^-=\xi_t\,\mathbb{I}_{|\xi|\leq M}$ and $\xi_t^+=\xi_t\,\mathbb{I}_{|\xi|>M}$, $\widetilde{e}^-(a,b)=\sum_{s\in\mathbb{Z}}\beta_a(b-s)\,\xi_s^-$ and $\widetilde{e}^+(a,b)=\sum_{s\in\mathbb{Z}}\beta_a(b-s)\,\xi_s^+$ using (44). We have $\widetilde{e}(a,b)=\widetilde{e}^+(a,b)+\widetilde{e}^-(a,b)$. To confirm (46), take :

$$\widetilde{T}_N(r_i\,a_N)-1 = \frac{1}{N-r_ia_N}\Big(\sum_{b=1}^{N-r_ia_N}\big(\widetilde{e}^-(r_ia_N,b)\big)^2-1\Big)-2\widetilde{e}^+(r_ia_N,b)\widetilde{e}^-(r_ia_N,b)+\big(\widetilde{e}^+(r_ia_N,b)\big)^2\Big) \tag{47}$$

We prove that $\big(\widetilde{T}_N^-(r_i\,a_N)-1\big)_{1\leq i\leq\ell}=\big(\frac{1}{N-r_ia_N}\sum_{b=1}^{N-r_ia_N}\big(\widetilde{e}^-(r_ia_N,b)\big)^2-1\big)_{1\leq i\leq\ell}$ satisfies (46). Indeed, $(\widetilde{e}^-(r_ia_N,b))$ is a linear process with innovations $(\xi_t^-)$ satisfying the Cramèr condition and it is obvious that $\Big(\frac{\mathbb{E}\big(\widetilde{e}(r_ia_N,b)\big)^2}{\mathbb{E}\big(\widetilde{e}^-(r_ia_N,b)\big)^2}\Big)^{1/2}\widetilde{e}^-(r_ia_N,b)_{b,i}$ has exactly the same distribution as $\widetilde{e}(r_ia_N,b)_{b,i}$. Therefore We yet have to prove that $\sqrt{\frac{N}{a_N}}\Big(\frac{\mathbb{E}\big(\widetilde{e}(r_i a_N,b)\big)^2}{\mathbb{E}\big(\widetilde{e}^-(r_i a_N,b)\big)^2}-1\Big)$ converges to 0. If $\mathbb{E}\big(\widetilde{e}(r_i a_N,b)\big)^2=\big(\sum_{s\in\mathbb{Z}}\beta_a^2(s)\big)\,\mathbb{E}(\xi_0)^2=1$ and $\mathbb{E}\xi_0^2=1$ (from Property 1), then

$$\Big|\frac{\mathbb{E}\big(\widetilde{e}^-(r_i a_N,b)\big)^2}{\mathbb{E}\big(\widetilde{e}(r_i a_N,b)\big)^2}-1\Big| \;\leq\; 2\Big(\mathbb{E}\big(\widetilde{e}^+(r_i a_N,b)\big)^2\Big)^{1/2}+\mathbb{E}\big(\widetilde{e}^+(r_i a_N,b)\big)^2.$$

Assuming that the distribution of $\xi_0$ is symmetric, we then obtain $\mathbb{E}\big(\widetilde{e}^+(r_i a_N,b)\big)^2=\big(\sum_{s\in\mathbb{Z}}\beta_a^2(s)\big)\,\mathbb{E}(\xi_0^+)^2=\mathbb{E}(\xi_0^+)^2$, with Hölder's and Markov's inequalities, however we have:

$$\mathbb{E}(\xi_0^+)^2\leq(\mathbb{E}\xi_0^4)^{1/2}(\Pr(|\xi_0|>M))^{1/2}\leq(\mathbb{E}\xi_0^4)\,M^{-2}.$$

Hence, there exists $C>0$ independent of $M$ and $N$,

$$\sqrt{\frac{N}{a_N}}\,\Big|\frac{\mathbb{E}\big(\widetilde{e}^-(r_i a_N,b)\big)^2}{\mathbb{E}\big(\widetilde{e}(r_i a_N,b)\big)^2}-1\Big|\leq\frac{C}{M}\sqrt{N}a_N \underset{N\to\infty}{\longrightarrow} 0$$

when $M=N$ (for instance). Therefore $\big(\widetilde{T}_N^-(r_i\,a_N)-1\big)_{1\leq i\leq\ell}$ satisfies the CLT (46).

From (47), it remains to prove:

$$\sqrt{\frac{N}{a_N}}\,\frac{1}{N-r_i a_N}\Big(\sum_{b=1}^{N-r_i a_N}-2\widetilde{e}^+(r_i a_N,b)\widetilde{e}^-(r_i a_N,b)+\big(\widetilde{e}^+(r_i a_N,b)\big)^2\Big)\underset{N\to\infty}{\overset{\mathcal{P}}{\longrightarrow}}0.$$

Wich based on Markov's and Hölder inequalities , is verified when $\sqrt{\frac{N}{a_N}}\,\Big(\mathbb{E}\big(\widetilde{e}^+(r_i a_N,b)\big)^2+2\sqrt{\mathbb{E}\big(\widetilde{e}^+(r_i a_N,b)\big)^2}\Big)\underset{N\to\infty}{\longrightarrow}0$ with $\mathbb{E}\big(\widetilde{e}^+(r_i a_N,b)\big)^2=1$. Using $\mathbb{E}\big(\widetilde{e}^+(r_i a_N,b)\big)^2\leq(\mathbb{E}\xi_0^4)\,M^{-2}$ obtained above, we can infere that this statement holds when $M=N$ (for instance). Consequently, from (47), CLT (46) holds even if the distribution of $\xi_0$ is only symmetric and such that $\mathbb{E}\xi_0^4<\infty$.

**Step 4** It remains to apply the Delta-method to (46) with function $(x_1,\ldots,x_\ell)\mapsto(\log x_1,\ldots,\log x_\ell)$:

$$\sqrt{\frac{N}{a_N}}\,\big(\log\big(T_N(r_i\,a_N)\big)-\log(\mathbb{E}e^2(a_N,1))\big)_{1\leq i\leq\ell}\underset{N\to\infty}{\overset{d}{\longrightarrow}}\mathcal{N}\big(0\,,\,\Gamma(r_1,\ldots,r_\ell,\psi,d)\big),$$

With $\mathbb{E}e^2(a_N,1)$ provided in Property 1, we obtain

$$\log\mathbb{E}e^2(a_N,1)=2d\,\log(a_N)+\log\big(\frac{c_d K_{(\psi,2d)}}{2\pi}\big)+\frac{c_{d'}K_{(\psi,2d-d')}}{2\pi\,a_N^{d'}}\big(1+o(1)\big)$$

Therefore, when $\sqrt{\frac{N}{a_N}}\frac{1}{a_N^{d'}}\underset{N\to\infty}{\longrightarrow}0$, *i.e.* $N^{\frac{1}{1+2d'}}=o(a_N)$, CLT (9) holds. □

*Proof of Theorem 1.* We use Theorem 1 of Bardet *et al.* (2008) which proved that CLT (9) remains valid when $a_N$ is replaced by $N^{\widetilde{\alpha}_N}$. Since $\widetilde{d}_N=\widetilde{M}_N\,Y_N(\widetilde{\alpha}_N)$ with $\widetilde{M}_N=\big(0\;\;1/2\big)\big(Z_1'\widehat{\Gamma}_N^{-1}Z_1\big)^{-1}Z_1'\widehat{\Gamma}_N^{-1}$ we deduce that $\sqrt{N/N^{\widetilde{\alpha}_N}}\big(\widetilde{d}_N-d\big)$ is asymptotically Gaussian with asymptotic variance limit in probability of $\widetilde{M}_N\,\Gamma(1,\ldots,\ell,d,\psi)\,\widetilde{M}_N'$, that is $\sigma^2$.

Relation (15) is also an obvious consequence of Theorem 1 of Bardet *et al.* (2008). □

*Proof of Theorem 2.* The theory of linear models can be applied as follows: $Z_{N^{\widetilde{\alpha}_N}} \begin{pmatrix} \widetilde{c}_N \\ 2\widetilde{d}_N \end{pmatrix}$ is an orthogonal projector of $Y_N(\widetilde{\alpha}_N)$ on a subspace of dimension 2, therefore $Y_N(\widetilde{\alpha}_N) - Z_{N^{\widetilde{\alpha}_N}} \begin{pmatrix} \widetilde{c}_N \\ 2\widetilde{d}_N \end{pmatrix}$ is an orthogonal projector of $Y_N(\widetilde{\alpha}_N)$ on a subspace of dimension $\ell - 2$. Moreover, using CLT (9) where $a_N$ is replaced by $N^{\widetilde{\alpha}_N}$, we deduce that $\sqrt{N/N^{\widetilde{\alpha}_N}}\widehat{\Gamma}_N^{-1} Y_N(\widetilde{\alpha}_N)$ asymptotically follows a Gaussian distribution with asymptotic covariance matrix $I_\ell$ (identity matrix). Hence, from the usual Cochran Theorem, we deduce (17). □

# References


[1] Abry, P., Flandrin, P., Taqqu, M.S. and Veitch, D. (2003). Self-similarity and long-range dependence through the wavelet lens, In Doukhan, P., Oppenheim, G. and Taqqu, M.S. (Editors), *Theory and applications of long-range dependence*, 527–556, Birkhäuser, Boston.

[2] Andrews, D.W.K. and Sun, Y. (2004). Adaptive local polynomial Whittle estimation of long-range dependence, *Econometrica*, **72**, 569–614.

[3] Bardet, J.M. and Bertrand, P. (2007) Definition, properties and wavelet analysis of multiscale fractional Brownian motion, *Fractals* **15**, 73–87.

[4] Bardet, J.M., Bibi, H. and Jouini, A. (2008). Adaptive wavelet-based estimator of the memory parameter for stationary Gaussian processes, *Bernoulli*, 14, 691-724.

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

[6] Bardet, J.M., Lang, G., Oppenheim, G., Philippe, A., Stoev, S. and Taqqu, M.S. (2003). Semiparametric estimation of the long-range dependence parameter: a survey. In *Theory and applications of long-range dependence*, Birkhäuser Boston, 557-577.

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

[8] Furmanczyk, K. (2007). Some remarks on the central limit theorem for functionals of linear processes under short-range dependence. *Probab. Math. Statist.*, 27, 235–245.

[9] Giraitis, L. (1985). The central limit theorem for functionals of a linear process. *Lith. Mat. J.*, 25, 43–57.

[10] Giraitis, L., Robinson P.M., and Samarov, A. (1997). Rate optimal semiparametric estimation of the memory parameter of the Gaussian time series with long range dependence. *J. Time Ser. Anal.*, 18, 49–61.

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

[12] Moulines, E., Roueff, F. and Taqqu, M.S. (2007). On the spectral density of the wavelet coefficients of long-memory time series with application to the log-regression estimation of the memory parameter, *J. Time Ser. Anal.*, 28, 155–187.

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

[14] Peltier, R. and Lévy Véhel, J. (1995) Multifractional Brownian Motion: definition and preliminary results. Preprint INRIA, available on `http://hal.inria.fr/docs/00/07/40/45/PDF/RR-2645.pdf`.

[15] Robinson, P.M. (1995). Gaussian semiparametric estimation of long range dependence, *Ann. Statist.*, 23, 1630–1661.

[16] Roueff, F. and Taqqu, M.S. (2009a). Asymptotic normality of wavelet estimators of the memory parameter for linear processes. *J. Time Ser. Anal.*, 30, 534–558.

[17] Roueff, F. and Taqqu, M.S. (2009b). Central limit theorems for arrays of decimated linear processes. *Stochastic Process. Appl.*, 119, 3006–3041

[18] Veitch, D., Abry, P. and Taqqu, M.S. (2003). On the Automatic Selection of the Onset of Scaling, *Fractals*, 11, 377–390.

[19] Wu, W.B. (2002). Central limit theorems for functionals of linear processes and their applications. *Statist. Sinica*, 12, 635–649.