Electronic Journal of Statistics
Vol. 6 (2012) 2383–2419
ISSN: 1935-7524
DOI: 10.1214/12-EJS754

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

Jean-Marc Bardet and Hatem Bibi

SAMM, Université Paris 1 Panthéon-Sorbonne, France
e-mail: Jean-Marc.Bardet@univ-paris1.fr; Hatem.Bibi@malix.univ-paris1.fr

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. As such this 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 has been modified and the 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.

AMS 2000 subject classifications: Primary 62M07, 62M09; secondary 62M10, 62M15, 60F05.
Keywords and phrases: Long range dependence, linear processes, wavelet estimator, semiparametric estimator, adaptive estimator, adaptive goodness-of-fit test.

Received April 2012.

Contents

1. Introduction

Presently, long memory processes have become a widely-studied subject area and find frequent applications (see for instance [9]). The best known long-memory stationary time series are the fractional Gaussian noises (FGN) with Hurst parameter H and the FARIMA(p, d, q) processes. For both these time series, the spectral density f at 0 follows a 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 is frequently considered as a definition of a stationary long-memory (or long-range dependent) process where d is the long memory parameter.

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

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

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

- *$(\xi_s)_{s\in\mathbb{Z}}$ is a sequence of independent identically distributed random variables following a symmetric distribution, i.e. $\Pr(\xi_0 > M) = \Pr(\xi_0 < -M)$ for all $M \in \mathbb{R}$, and satisfying $\mathbb{E}\xi_0 = 0$, $Var\,\xi_0 = 1$ 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*

$$|\widehat{\alpha}(\lambda)|^2 = \frac{1}{|\lambda|^{2d}}\left(c_d + c_{d'}|\lambda|^{d'}(1+\varepsilon(\lambda))\right) \quad \textit{for any} \quad \lambda \in [-\pi, 0) \cup (0, \pi], \quad (1)$$

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

Consequently, if X satisfies Assumption A(d, d'), the spectral density f of X is such that for any $\lambda \in [-\pi, 0) \cup (0, \pi]$,

$$f(\lambda) = 2\pi\,|\widehat{\alpha}(\lambda)|^2 = \frac{2\pi}{|\lambda|^{2d}}\left(c_d + c_{d'}|\lambda|^{d'}(1 + \varepsilon(\lambda))\right), \quad (2)$$

with $\varepsilon(\lambda) \to 0$ when $\lambda \to 0$. Thus, if $d \in (0, 1/2)$, the process X is a long-memory process, and if $d \le 0$, it is a short-memory process (see [9]). Finally, an estimator of d will be to be *adaptive* when this estimator has an expression that is valid for all processes in the semi-parametric class A(d, d'), *i.e.* when this estimator does not depend on d and d'.

After preliminary studies devoted to self-similar processes, [1] was the first to propose the use of a wavelet-based estimator for estimating the parameter d of a long memory process by computing the log-log regression slope for different scales of wavelet coefficient sample variances. [6] provided proofs of the consistency of such an estimator in a Gaussian semiparametric frame. [19] improved these results and established a central limit theorem (CLT in the sequel) for the estimator of d which they proved rate optimal for the minimax criterion. Finally, [25] 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. However, these results 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. A procedure based on a chi-square test was introduced in [28] but despite convincing numerical results, it lacks proofs of its consistency. Whereas, [4] proved the consistency of a procedure for choosing optimal scales based on the detection of the "most linear part" of the log-variogram graph. Moreover, the considered wavelet function is not necessarily associated with a multi-resolution analysis: although the computation cost is more important, this offers a larger wavelet function choice and scales are not limited to powers of 2.

The present paper is an extension of this previous study of [4]. Improvements concern three following central issues:

1. The semiparametric Gaussian framework of [4] is extended to the semiparametric framework Assumption $\mathrm{A}(d,d')$ for linear processes. The same automatic procedure of the optimal scale selection can also be used and thus we obtain adaptive estimators.
2. As in [4], 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 a very simple asymptotic covariance matrix (see (9) for its definition) 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 in the sequel) of d, satisfying a CLT with an asymptotic variance which is smaller than the adaptive 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. Since the asymptotic covariance matrix is easily approximated, the test is very simple 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: "the process is such that Assumption $\mathrm{A}(d,d')$ holds", and reasonable properties of robustness under H_1: "the process is such that Assumption $\mathrm{A}(d,d')$ does not hold".

In the light of these results, this paper is a conclusion to the study of [4], and the adaptive PGLS estimator and test are interesting extensions of [25].

The present paper is organized into four sections as follows. Assumptions, definitions and a first multidimensional central limit theorem are the subject matter of Section 2. Section 3 is devoted to the construction and consistency of the adaptive PGLS estimator and goodness-of-fit test. Section 4 features a Monte Carlo simulations-based demonstration of the convergence of the adaptive estimator, followed by comparisons with other efficient semiparametric estimators and investigations into the consistency and robustness properties of the adaptive goodness-of-fit test. Proofs figure in Section 5.

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

Let $\psi : \mathbb{R} \to \mathbb{R}$ be a function (called the wavelet function) and $k \in \mathbb{N}^*$ (with the usual notation $\mathbb{N}^* = \mathbb{N} \setminus \{0\}$). We shall consider the following assumption on ψ:

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

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

Straightforward implications of Assumption $\Psi(k)$ are:

- $\psi^{(j)}(0) = \psi^{(j)}(1) = 0$ for any $0 \le j \le k$, where $\psi^{(j)}$ is the j-th derivative of ψ.
- If $\widehat{\psi}(u)$ is the Fourier transform of ψ, *i.e.*

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

then $\widehat{\psi}(u) = C\, u(1 + \varepsilon(u))$ when $u \to 0$ with C a complex number not depending on u and $\varepsilon(u) \to 0$ when $u \to 0$ since $\widehat{\psi}(0) = 0$ from the assumption $\int_0^1 \psi(t)\, dt = 0$.
- Moreover,

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

If $Y = (Y_t)_{t\in\mathbb{R}}$ is a continuous-time process, with a.s. continuous trajectories, 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

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

for $(a, b) \in \mathbb{N}^* \times \mathbb{Z}$. Note that if a path $(X_1, \ldots, X_N)$ is observed, for $a \in \mathbb{N}^*$ and $b = 1, \ldots, N - a$ we can also write $e(a, b) = \frac{1}{\sqrt{a}} \sum_{t=1}^N \psi(\frac{t-b}{a}) X_t$, which is more directly implied by the definition of $d(a, b)$.

In the sequel, we will 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* ψ *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 when $a \to \infty$,

$$\mathbb{E}(e^2(a,0)) = 2\pi\big(c_d K_{(\psi,2d)} a^{2d} + c_{d'} K_{(\psi,2d-d')} a^{2d-d'}\big) + o\big(a^{2d-d'}\big),$$
$$\textit{with} \qquad K_{(\psi,\alpha)} := \int_{-\infty}^{\infty} |\widehat{\psi}(u)|^2\, |u|^{-\alpha} du > 0 \quad \textit{for all } \alpha < 1. \tag{5}$$

Refer to Section 5 for all the proofs of this paper.

Remark 1. The condition $k > d' - d + 1/2$ of Property 1 could be surprising since the larger d', *i.e.* the smoother the spectral density, the larger k is required. Heuristically, the reason of such a condition is the following: the larger the second order term $|\lambda|^{d'-2d}$ of the spectral density (*i.e.* the smoother the spectral density), the larger k should be (*i.e.* the smoother the wavelet function should be) to detect this second order term, which is required for the adpative procedure (see Section 3). Moreover, from the proof of Property 1, if $k \le d' - d + 1/2$ then the second order expansion term $a^{2d-d'}$ in (5) is replaced by $O\big(a^{2d} a^{1/2-d-k}\big)$. Then the forthcoming central limit theorem (8) still holds but under the condition $a_N\, N^{-1/(2d+2k)} \to \infty$ replacing the condition $a_N\, N^{-1/(1+2d')} \to \infty$: this induces a slower convergence rate of this CLT than under the condition $k > d' - d + 1/2$.

Let $(X_1, \dots, X_N)$ be an observed path of X satisfying Assumption 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 $2d$. Hence, for $a \in \{1, \dots, N-1\}$, consider the sample variance of the wavelet coefficients,

$$T_N(a) := \frac{1}{N-a} \sum_{b=1}^{N-a} e^2(a,b). \tag{6}$$

Remark 2. In [6, 4, 19] or [25], the considered sample variance of wavelet coefficients is

$$V_N(a) := \frac{1}{[N/a]} \sum_{b=1}^{[N/a]} e^2(a,ab) \tag{7}$$

(with $a = 2^j$ in case of multiresolution analysis). Definition (6) has both a drawback and two advantages with respect to the usual definition (7). On the one hand, $T_N(a)$ is not adapted to the fast Mallat's algorithm and therefore its use is more time consuming than the one of $V_N(a)$. On the other hand its advantage is twofold: if γ and γ' respectively denote the asymptotic variances of $\sqrt{N/a}\, T_N(a)$ and $\sqrt{N/a}\, V_N(a)$ when a, $N \to \infty$, then the expression of γ is clearly simpler than the one of γ' since

$$\begin{cases} \gamma &= 4\pi \dfrac{1}{K^2_{(\psi,2d)}} \displaystyle\int_{-\infty}^{\infty} \frac{\big|\widehat{\psi}(\lambda)\big|^4}{|\lambda|^{4d}} d\lambda \quad \text{(see (9) below)} \\ \gamma' &= \dfrac{2}{K^2_{(\psi,2d)}} \displaystyle\sum_{m=-\infty}^{\infty} \Big(\int_{-\infty}^{\infty} \frac{\big|\widehat{\psi}(u)\big|^2}{|u|^{2d}} \cos(u\, m)\, du \Big)^2 \quad \text{(see [4])}\ , \end{cases}$$

and this will have consequences to the computation of PGLS estimators below. Furthermore, as inferred from numerical approximations not reported here, for our choice of ψ (see Section 4), $\gamma(d) \in [\gamma'(d)/2, \gamma'(d)]$ (following d). This confers the same advantage to the variance of the wavelet-based estimators of d computed from $(T_N(ar_i))_i$ with respect to the one computed from $(V_N(ar_i))_i$ (see below).

The following proposition specifies a multidimensional central limit theorem for a vector $(\log T_N(r_i a_N))_i$, which provides the first step towards obtaining a CLT for the estimator of d computed from an ordinary least square regression:

Proposition 1. *Define $\ell \in \mathbb{N} \setminus \{0,1\}$ and $(r_1, \ldots, r_\ell) \in (\mathbb{N}^*)^\ell$ with $0 < r_1 < r_2 < \cdots < r_\ell$. Under Assumption $A(d, d')$ with $d < 1/2$ and $d' > 0$, if ψ satisfies Assumption $\Psi(k)$ with $k \geq d' - d + 1/2$ and if $(a_n)_{n \in \mathbb{N}}$ is such as $N/a_N \to \infty$ and $a_N\, N^{-1/(1+2d')} \to \infty$ when $N \to \infty$, then*

$$\Big(\sqrt{\frac{N}{r_i a_N}} \big(\log T_N(r_i a_N) - 2d \log(r_i a_N) - \log(2\pi\, c_d\, K_{(\psi,2d)}) \big) \Big)_{1 \leq i \leq \ell} \xrightarrow[N\to\infty]{d} \mathcal{N}_\ell\big(0\,;\, \Gamma(r_1, \ldots, r_\ell, \psi, d)\big), \quad (8)$$

with $\Gamma(r_1, \ldots, r_\ell, \psi, d) = (\gamma(r_i, r_j))_{1 \leq i,j \leq \ell}$ the asymptotic covariance matrix such as

$$\gamma(r_i, r_j) = 4\pi\, \frac{(r_i r_j)^{1/2-2d}}{K^2_{(\psi,2d)}} \int_{-\infty}^{\infty} \frac{\big|\widehat{\psi}(r_i\lambda)\big|^2 \big|\widehat{\psi}(r_j\lambda)\big|^2}{|\lambda|^{4d}} d\lambda. \quad (9)$$

Remark 3. This result can be compared with the one of [25]. In this paper, two assumptions are required concerning the process: Assumption 1 concerning the behavior of the spectral density is a little weaker (and therefore more general) than condition (2), while Assumption 2 is very close to our conditions on $(\xi_s)_{s\in\mathbb{Z}}$ in Assumption $\mathrm{A}(d,d')$ except that we require the symmetry of the distribution of this process. However, our proof of Proposition 1 could also be obtained under Assumption 1 of [25]: the second order expansion in (2) is only required for obtaining the adaptive estimation of the minimal scale (see forthcoming Section 3). The condition on the symmetry of the distribution of $(\xi_s)_{s\in\mathbb{Z}}$ provides a very simple expression (9) of the asymptotic covariance while its expression $(V_{ij}(d,\psi)$ in (66)) in [25] is notably more complicated (and therefore almost impossible to be approximated). Finally, our condition $a_N\, N^{-1/(1+2d')} \to \infty$ exactly corresponds to the condition (63) of Theorem 4 in [25] since $a_N = 2^{L(n)}$. Note that if $a_N\, N^{-1/(1+2d')} \to \ell_0$ with $0 \leq \ell_0 < \infty$, then

$$\sqrt{\frac{N}{a_N}} \Big[\log \mathbb{E}\big(e^2(a_N,0)\big) - \big(2d\, \log(a_N) + \log\big(2\pi\, c_d K_{(\psi,2d)}\big)\big) \Big]$$

does not tend to 0 (see (57)) and thus CLT (8) does not hold (if $0 < \ell_0 < \infty$ then the limit distribution in CLT (8) is a non-centered Gaussian distribution).

Since it is not easy to minimize (following the order relation induced by definite positive matrices) $\Gamma(r_1,\ldots,r_\ell,\psi,d)$ in terms of $(r_1,\ldots,r_\ell)$ and for simplifying the following results, we chose now only to consider the case $(r_1,r_2,\ldots,r_\ell)=(1,2,\ldots,\ell)$. It is also possible to consider the case when $\ell\to\infty$.

Proposition 2. *Under Assumption $A(d,d')$ with $d<1/2$ and $d'>0$, if ψ satisfies Assumption $\Psi(k)$ with $k\geq d'-d+1/2$ and if $(a_n)_{n\in\mathbb{N}}$ and $(\ell_n)_{n\in\mathbb{N}}$ are two sequences of positive integer numbers such as when $N\to\infty$*

$$a_N/N\to 0,\quad a_N\,N^{-1/(1+2d')}\to\infty,\quad \ell_N\to\infty\quad \textit{and}\qquad \ell_N\,a_N/N\to 0,$$

then

$$\Big(\sqrt{\frac{N}{r_i a_N}}\big(\log T_N(r_i a_N)-2d\log(r_i a_N)-\log(2\pi\,c_d\,K_{(\psi,2d)})\big)\Big)_{1\leq i\leq \ell_N} \xrightarrow[N\to\infty]{f.d.d.} (Z_n)_{n\in\mathbb{N}^*} \qquad (10)$$

where $\xrightarrow[N\to\infty]{f.d.d.}$ denotes the convergence for finite dimensional distributions and $(Z_n)_{n\in\mathbb{N}^*}$ is a sequence of centered Gaussian random variables with covariance $Cov(Z_i,Z_j)=\gamma(i,j)$ where γ is defined in (9).

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 [25]). Proposition 1 confirms the consistency of estimator $\widehat{d}_N$ of d. Hence, we define

$$\widehat{d}_N(a_N):=\big(0\ \frac{1}{2}\big)\,(Z'_{a_N}\,Z_{a_N})^{-1}Z'_{a_N}\big(\log T_N(r_i a_N)\big)_{1\leq i\leq \ell}$$
$$\text{with}\quad Z_{a_N}=\left(\begin{array}{cccc}1 & 1 & \cdots & 1\\ \log(a_N) & \log(2a_N) & \cdots & \log(\ell a_N)\end{array}\right)'. \qquad (11)$$

where A' denotes the transpose of a matrix A. Then, it can be clearly inferred from Proposition 1 that $\widehat{d}_N(a_N)$ converges to d following a CLT with convergence rate $\sqrt{N/a_N}$ when a_N satisfies the condition $a_N\,N^{-1/(1+2d')}\to\infty$ $(N\to\infty)$.

But d' is actually unknown. [4] 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

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

$Q_N(\alpha,c,d)$ corresponds to a squared distance between the line of slope $2d$ and intercept c and the ℓ points $\big(\log(iN^\alpha)\,,\,\log T_N(iN^\alpha)\big)_i$. It can be minimized in terms of α, c and d first by defining for $\alpha\in(0,1)$

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

and then by defining $\widehat\alpha_N$ by:

$$\widehat Q_N(\widehat\alpha_N):=\min_{\alpha\in\mathcal A_N}\widehat Q_N(\alpha)\quad\text{where}\quad \mathcal A_N:=\Big\{\frac{2}{\log N}\,,\,\frac{3}{\log N}\,,\dots,\,\frac{\log[N/\ell]}{\log N}\Big\}.$$

Remark 4. As outlined in [4] in the 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 [4]),

$$\widehat\alpha_N=\frac{\log\widehat a_N}{\log N}\ \xrightarrow[N\to\infty]{\mathcal P}\ \alpha^*=\frac{1}{1+2d'}.$$

We then define:

$$\widehat{\widehat d_N}:=\widehat d(N^{\widehat\alpha_N})\quad\text{and}\quad \widehat\Gamma_N:=\Gamma(1,\dots,\ell,\widehat{\widehat d_N},\psi).\tag{12}$$

It is clear that $\widehat{\widehat d_N}\ \xrightarrow[N\to\infty]{\mathcal P}\ d$ (for a convergence rate see also [4]) and from the expression of Γ in (9) which is a continuous function of the variable d, $\widehat\Gamma_N\ \xrightarrow[N\to\infty]{\mathcal P}\ \Gamma(1,\dots,\ell,d,\psi)$. We will prefer to consider:

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

rather than $\widehat\alpha_N$ for technical reasons (*i.e.* $\Pr(\widetilde\alpha_N\le\alpha^*)\ \xrightarrow[N\to\infty]{}\ 0$ which is not satisfied by $\widehat\alpha_N$, see [4]). Consequently, with 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 the assumptions of Proposition 1,*

$$\sqrt{\frac{N}{N^{\widetilde\alpha_N}}}\big(\widetilde d_N-d\big)\ \xrightarrow[N\to\infty]{d}\ \mathcal N\big(0\,;\,\sigma_d^2(\ell)\big)$$
$$\textit{with}\quad \sigma_d^2(\ell):=\big(0\ \ \tfrac12\big)\Big(Z_1'\big(\Gamma(1,\dots,\ell,d,\psi)\big)^{-1}Z_1\Big)^{-1}\big(0\ \ \tfrac12\big)'\tag{14}$$
$$\textit{and for all}\quad \rho>\frac{2(1+3d')}{(\ell-2)d'},\qquad \frac{N^{\frac{d'}{1+2d'}}}{(\log N)^\rho}\times\big|\widetilde d_N-d\big|\ \xrightarrow[N\to\infty]{\mathcal P}\ 0.\tag{15}$$

We can note the following points:

1. From Gauss-Markov Theorem, 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 $\widetilde{d}_N$ converges to d with a rate of convergence being equal to the minimax rate of convergence $N^{\frac{d'}{1+2d'}}$ up to a logarithm factor (see [13]). Thus, this estimator is comparable to adaptive log-periodogram or local Whittle estimators (see [21]).
3. Under additive assumptions on ψ (ψ 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 adaptive versions of the local Whittle and FEXP estimators robust for polynomial trends were respectively defined in [2] and [16]).

Remark 5. As we studied the case $\ell \to \infty$ in Proposition 2, we have expected to extend Theorem 1 to the case $\ell \to \infty$. For establishing such a result, it is sufficient to prove that $\lim_{\ell\to\infty} \big(Z_1'\big(\Gamma(1,\dots,\ell,d,\psi)\big)^{-1} Z_1\big)^{-1}$ exists. But unfortunately we did not succeed to theoretically prove that this limit exists or to prove that the sequence $(\sigma_d^2(\ell))_\ell$ decreases. However, since $\Gamma(1,\dots,\ell,d,\psi)$ is defined from the general term (9) which can be easily approximated using classical approximations of the integrals once ψ is chosen, we realized numerical experiments for exhibiting the dependence of the asymptotic variance $\sigma_d^2(\ell)$ with d and ℓ. For the results of these numerical experiments refer to Figure 1. It can be

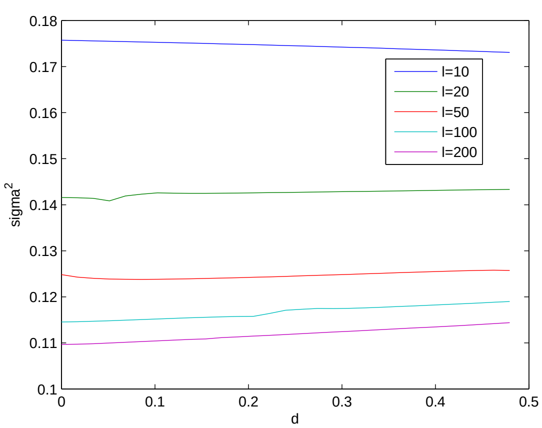

FIG 1. *Graph of the approximated values of* $\sigma_d^2(\ell)$ *defined in (14) for* $d \in [0, 0.5]$ *and* $\ell = 10, 20, 50, 100$ *and* 200.

inferred that for any $d \in [0, 0.5)$, $\sigma_d^2(\ell)$ is almost independent on d and decreases as ℓ increases. Then it could be interesting to select the "largest" possible value of ℓ (since we necessarily have $\ell \leq N/N^{\alpha^*}$) for improving the estimation of d. This will induce both the choices of ℓ realized in the forthcoming Section 4 devoted to simulations.

Finally an adaptive goodness-of-fit test can be deduced from the previous PGLS regression. It consists of a sum of the PGLS squared distances between the PGLS regression line and the points. To be precise, consider the statistic:

$$\widetilde{W}_N := \frac{N}{N^{\widetilde{\alpha}_N}} \Big(Y_N(\widetilde{\alpha}_N) - Z_{N^{\widetilde{\alpha}_N}} \begin{pmatrix} \widetilde{c}_N \\ 2\widetilde{d}_N \end{pmatrix} \Big)' \, \widehat{\Gamma}_N^{-1} \Big(Y_N(\widetilde{\alpha}_N) - Z_{N^{\widetilde{\alpha}_N}} \begin{pmatrix} \widetilde{c}_N \\ 2\widetilde{d}_N \end{pmatrix} \Big). \tag{16}$$

This test statistic can be used as a goodness-of-fit test for deciding between hypothesis H_0: "the process is a linear LRD process" and H_1: "the process is not a LRD process". Then, using the previous results, we obtain:

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

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

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 using 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 realized simulations and we compared the values of $\widetilde{d}_N$ with those of the more accurate semiparametric long-memory estimators. To conclude we examined the numerical properties of the test statistic $\widetilde{W}_N$.

Remark 6. Note that all softwares (in Matlab language) used in this section are freely available on:
`http://samm.univ-paris1.fr/-Jean-Marc-Bardet`.

The wavelet-based estimator has been computed using the following parameters:

Choice of the function ψ: A wavelet function ψ associated with a multiresolution analysis being not mandatory, as mentioned above, we use function $\psi(x) = x^4(1-x)^4\big(x - \frac{1}{2}\big)\mathbb{I}_{x\in[0,1]}$ which satisfies Assumption $\Psi(3)$ (and therefore in any cases $3 = k > d' - d + 1/2$ which is required for theoretical limit theorems).

Choice of the parameter ℓ: This parameter is important since it determines the "beginning" of the linear part of the graph drawn by the points $(\log(ia_N), \log T_N(ia_N))_{1\leq i\leq \ell}$ and hence the data-driven estimator $\widetilde{a}_N$, but also the chosen number of scales used for the regression. Hence, we adopted on this point a two step procedure:

1. According to numerical studies (not detailed here), the choice of $\ell = [5 + 2 * \log(N/100)]$ (therefore $\ell = 6$, 9 and 12 respectively for $N = 250$, 1000 and 5000) seems to be an appropriated choice for a first step: the computation of $\widetilde{\alpha}_n$. This choice is almost not influenced by the values of d or the roughness of the spectral density of X. Note also that a choice as $\ell = 10$ for all N provides only slightly worse results to those obtained with $\ell = [5 + 2 * \log(N/100)]$ but still reasonable. We also use this value of ℓ for computing the goodness-of-fit $\widetilde{W}_N$.
2. Concerning the computation of $\widetilde{d}_N$, we can follow the arguments of Remark 5 based on Figure 1. Then we selected for the second step a large value of ℓ close to the largest possible value, *i.e.* $\ell = N^{1-\widetilde{\alpha}_N}(\log N)^{-1}$ since necessarily $\ell \leq N/N^{\widetilde{\alpha}_N}$ by construction and for checking the conditions of the CLT (10), we should have $N^{\widetilde{\alpha}} \times \ell = o(N)$).

First of all we need to specify the simulation conditions. The results are obtained from generated independent samples of each process belonging to the following "benchmark". The practical 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 [9]). The simulations are carried out for $d = 0$, 0.1, 0.2, 0.3 and 0.4, for $N = 250$, 1000 and 5000 for all 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$(d+1/2)$ is such that Assumption A$(d, 2)$ holds (even if a FGN is rarely presented as a 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 ξ_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, thus confirming Assumption $A(d, d')$ holds.

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

We computed $\widetilde{d}_N$ following the two steps procedure to a benchmark referred to below including the 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$(1, d, 1)$ processes with standard Gaussian innovations, MA coefficient $\phi = -0.3$ and AR coefficient $\theta = 0.7$;
- X_5 : FARIMA$(1, d, 1)$ processes with innovations following a uniform distribution $\mathcal{U}[-1, 1]$, MA coefficient $\phi = 0.8$ and AR coefficient $\theta = -0.6$;
- X_6 : $X^{(d,d')}$ Gaussian processes with $d' = 1$.

We also compared the results obtained with $\widetilde{d}_N$ with those obtained with both the other following semiparametric d-estimators known for their accuracies (see [7] or 2008):

- $\widehat{d}_{MS}$ is the adaptive global log-periodogram estimator, also called FEXP estimator, defined by [17] and based on results obtained in [23] and [20]. The review article [21] is an excellent survey on semiparametric spectral estimators of the long memory parameter (local log-periodogram, local Whittle or FEXP estimators). We thank Eric Moulines for having provided a Matlab sofware of this adaptive estimator (as it was recommended by E. Moulines, the bias-variance balance parameter κ was set to 2, as in [7].
- $\widehat{d}_R$ is the local Whittle estimator introduced by [18] and asymptotically studied by [24]. We used the typical non-data-based choice of bandwidth $m = [N^{4/5}]$ (see [15]). Note that this choice is not an adaptive choice but an optimal choice under a regularity condition of the spectral density f (corresponding to $d' = 2$ which is the case of FARIMA processes; for $d' > 0$ known a choice $[N^{\frac{2d'}{1+2d'}}]$ can be done).

For simulation results see Table 1.

Conclusions from Table 1: Compared to other estimators, $\widetilde{d}_N$ numerically shows a convincing convergence rate. With a bandwidth parameter $m = [N^{4/5}]$ which is optimal (up to a constant) for FGN and FARIMA processes, the local Whittle estimator $\widehat{d}_R$ provides excellent and best results for the processes X_1, X_2 and X_3 having very flat spectral densities, but less good results for the processes $X4-6$ having more rough spectral densities. The adaptive FEXP $\widehat{d}_{MS}$ presents opposite properties: its results are almost constant whatever is the roughness of spectral density. Finally the adaptive wavelet based estimator $\widetilde{d}_N$ provides intermediary results between those of $\widehat{d}_R$ and $\widehat{d}_{MS}$: good but not excellent for processes having smooth spectral densities $(X1-3)$ and not very good but not bad results for processes having rough spectral densities $(X4-6)$.

Remark 7. A previous comparison [4] of two adaptive wavelet-based estimators (respectively defined in [28] and in [4]) with $\widehat{d}_{MS}$ and $\widehat{d}_R$ (as well as with two further estimators as defined respectively in [14], and [8] neither of which display good numerical properties of consistency) shows that $\sqrt{MSE}$ of $\widetilde{d}_N$ obtained in Table 1 is generally smaller to $\sqrt{MSE}$ of [4]-based estimator (especially for processes $X4$ and $X6$) because we opted for definition (6) instead of (7) and PGLS regression instead of LS regression.

Robustness of the estimator $\widetilde{d}_N$: To study the robustness of the estimator $\widetilde{d}_N$ and also of the goodness-of-fit test $\widetilde{W}_N$, we chose different processes not

TABLE 1

Comparison of long-memory parameter estimators for benchmark processes. $\sqrt{MSE}$ is computed from 1000, 500 and 200 independent samples (for $N = 250$, 1000 and 5000 resp.). The frequency of the goodness-of-fit test acceptation is $\widetilde{p}_n = \frac{1}{n}\,\#\big(\widetilde{W}_N < q_{\chi^2(\ell-2)}(0.95)\big)$

$N = 250$ ⟶

Model	$\sqrt{MSE}$	$d=0$	$d=0.1$	$d=0.2$	$d=0.3$	$d=0.4$
X_1	$\sqrt{MSE}\ \widehat{d}_{MS}$	0.176	0.163	0.157	0.173	0.182
	$\sqrt{MSE}\ \widehat{d}_R$	**0.063**	**0.060**	**0.062**	**0.067**	**0.071**
	$\sqrt{MSE}\ \widetilde{d}_N$	0.098	0.095	0.089	0.086	0.089
	$\widetilde{p}_n$	0.95	0.94	0.94	0.91	0.91
X_2	$\sqrt{MSE}\ \widehat{d}_{MS}$	0.165	0.165	0.166	0.169	0.185
	$\sqrt{MSE}\ \widehat{d}_R$	**0.061**	**0.065**	**0.064**	**0.064**	**0.065**
	$\sqrt{MSE}\ \widetilde{d}_N$	0.098	0.107	0.104	0.105	0.106
	$\widetilde{p}_n$	0.95	0.93	0.93	0.94	0.88
X_3	$\sqrt{MSE}\ \widehat{d}_{MS}$	0.167	0.168	0.171	0.168	0.172
	$\sqrt{MSE}\ \widehat{d}_R$	**0.063**	**0.061**	**0.065**	**0.066**	**0.065**
	$\sqrt{MSE}\ \widetilde{d}_N$	0.099	0.102	0.105	0.099	0.101
	$\widetilde{p}_n$	0.95	0.96	0.95	0.92	0.93
X_4	$\sqrt{MSE}\ \widehat{d}_{MS}$	**0.179**	**0.175**	**0.178**	**0.175**	**0.175**
	$\sqrt{MSE}\ \widehat{d}_R$	0.333	0.330	0.325	0.324	0.317
	$\sqrt{MSE}\ \widetilde{d}_N$	0.219	0.204	0.194	0.190	0.183
	$\widetilde{p}_n$	0.85	0.84	0.82	0.79	0.77
X_5	$\sqrt{MSE}\ \widehat{d}_{MS}$	**0.170**	**0.171**	**0.156**	**0.173**	**0.180**
	$\sqrt{MSE}\ \widehat{d}_R$	0.478	0.518	0.532	0.538	0.539
	$\sqrt{MSE}\ \widetilde{d}_N$	0.359	0.381	0.383	0.381	0.371
	$\widetilde{p}_n$	0.86	0.85	0.86	0.81	0.80
X_6	$\sqrt{MSE}\ \widehat{d}_{MS}$	**0.175**	0.183	**0.177**	0.177	0.178
	$\sqrt{MSE}\ \widehat{d}_R$	0.173	**0.176**	**0.177**	**0.172**	**0.172**
	$\sqrt{MSE}\ \widetilde{d}_N$	0.223	0.225	0.218	0.209	0.202
	$\widetilde{p}_n$	0.94	0.93	0.92	0.92	0.91

$N = 1000$ ⟶

Model	$\sqrt{MSE}$	$d=0$	$d=0.1$	$d=0.2$	$d=0.3$	$d=0.4$
X_1	$\sqrt{MSE}\ \widehat{d}_{MS}$	0.091	0.088	0.089	0.090	0.097
	$\sqrt{MSE}\ \widehat{d}_R$	**0.033**	**0.034**	**0.035**	**0.038**	**0.043**
	$\sqrt{MSE}\ \widetilde{d}_N$	0.062	0.053	0.050	0.049	0.045
	$\widetilde{p}_n$	0.89	0.88	0.90	0.94	0.91
X_2	$\sqrt{MSE}\ \widehat{d}_{MS}$	0.089	0.089	0.089	0.091	0.090
	$\sqrt{MSE}\ \widehat{d}_R$	**0.034**	**0.033**	**0.036**	**0.035**	**0.035**
	$\sqrt{MSE}\ \widetilde{d}_N$	0.059	0.060	0.061	0.065	0.058
	$\widetilde{p}_n$	0.93	0.90	0.90	0.90	0.89
X_3	$\sqrt{MSE}\ \widehat{d}_{MS}$	0.086	0.094	0.093	0.096	0.098
	$\sqrt{MSE}\ \widehat{d}_R$	**0.034**	**0.033**	**0.033**	**0.035**	**0.035**
	$\sqrt{MSE}\ \widetilde{d}_N$	0.065	0.064	0.062	0.062	0.059
	$\widetilde{p}_n$	0.89	0.88	0.90	0.88	0.87
X_4	$\sqrt{MSE}\ \widehat{d}_{MS}$	**0.091**	**0.090**	**0.095**	**0.094**	**0.095**
	$\sqrt{MSE}\ \widehat{d}_R$	0.204	0.201	0.195	0.194	0.193
	$\sqrt{MSE}\ \widetilde{d}_N$	0.119	0.117	0.111	0.114	0.109
	$\widetilde{p}_n$	0.83	0.80	0.82	0.79	0.77
X_5	$\sqrt{MSE}\ \widehat{d}_{MS}$	**0.092**	**0.089**	**0.083**	**0.094**	**0.097**
	$\sqrt{MSE}\ \widehat{d}_R$	0.355	0.361	0.366	0.365	0.364
	$\sqrt{MSE}\ \widetilde{d}_N$	0.193	0.192	0.183	0.173	0.172
	$\widetilde{p}_n$	0.85	0.83	0.82	0.78	0.85
X_6	$\sqrt{MSE}\ \widehat{d}_{MS}$	**0.096**	**0.096**	**0.099**	**0.100**	**0.109**
	$\sqrt{MSE}\ \widehat{d}_R$	0.133	0.129	0.131	0.133	0.130
	$\sqrt{MSE}\ \widetilde{d}_N$	0.151	0.154	0.151	0.152	0.140
	$\widetilde{p}_n$	0.86	0.84	0.86	0.82	0.84

$N = 5000$ ⟶

Model	$\sqrt{MSE}$	$d=0$	$d=0.1$	$d=0.2$	$d=0.3$	$d=0.4$
X_1	$\sqrt{MSE}\ \widehat{d}_{MS}$	0.043	0.046	0.045	0.041	0.043
	$\sqrt{MSE}\ \widehat{d}_R$	**0.017**	**0.019**	**0.019**	**0.020**	**0.020**
	$\sqrt{MSE}\ \widetilde{d}_N$	0.025	0.022	0.025	0.028	0.027
	$\widetilde{p}_n$	0.92	0.95	0.95	0.96	0.98
X_2	$\sqrt{MSE}\ \widehat{d}_{MS}$	0.045	0.045	0.040	0.044	0.048
	$\sqrt{MSE}\ \widehat{d}_R$	**0.018**	**0.016**	**0.016**	**0.017**	**0.019**
	$\sqrt{MSE}\ \widetilde{d}_N$	0.031	0.032	0.030	0.029	0.029
	$\widetilde{p}_n$	0.94	0.94	0.93	0.93	0.93
X_3	$\sqrt{MSE}\ \widehat{d}_{MS}$	0.044	0.044	0.046	0.044	0.049
	$\sqrt{MSE}\ \widehat{d}_R$	**0.019**	**0.017**	**0.017**	**0.019**	**0.018**
	$\sqrt{MSE}\ \widetilde{d}_N$	0.030	0.038	0.026	0.031	0.032
	$\widetilde{p}_n$	0.92	0.95	0.92	0.92	0.93
X_4	$\sqrt{MSE}\ \widehat{d}_{MS}$	**0.046**	**0.043**	**0.043**	**0.050**	**0.045**
	$\sqrt{MSE}\ \widehat{d}_R$	0.113	0.110	0.110	0.109	0.108
	$\sqrt{MSE}\ \widetilde{d}_N$	0.059	0.054	0.058	0.062	0.049
	$\widetilde{p}_n$	0.93	0.87	0.87	0.84	0.88
X_5	$\sqrt{MSE}\ \widehat{d}_{MS}$	**0.042**	**0.048**	**0.046**	**0.043**	**0.054**
	$\sqrt{MSE}\ \widehat{d}_R$	0.242	0.244	0.245	0.246	0.243
	$\sqrt{MSE}\ \widetilde{d}_N$	0.087	0.081	0.074	0.076	0.074
	$\widetilde{p}_n$	0.78	0.80	0.79	0.80	0.77
X_6	$\sqrt{MSE}\ \widehat{d}_{MS}$	**0.049**	**0.050**	**0.047**	**0.051**	**0.048**
	$\sqrt{MSE}\ \widehat{d}_R$	0.100	0.100	0.099	0.099	0.096
	$\sqrt{MSE}\ \widetilde{d}_N$	0.090	0.087	0.083	0.084	0.077
	$\widetilde{p}_n$	0.82	0.81	0.76	0.82	0.86

satisfying Assumption $A(d,d')$ (except sometimes for $d=0$) and defined as follows:

1. a 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 $\mathbb{E}|X_i|<\infty$). Note that X_4 does 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 think that one should be able to prove the consistency of $\widetilde{d}_N$ under the condition $\mathbb{E}\xi_0^r$ with $r>0$.
2. a Gaussian stationary process with a spectral density $f(\lambda)=\big||\lambda|-\pi/2\big|^{-2\delta}$ for all $\lambda\in[-\pi,\pi]\setminus\{-\pi/2,\pi/2\}$ so called a GARMA$(0,\delta,0)$ process. The local behavior of f in 0 is $f(|\lambda|)\sim(\pi/2)^{-2\delta}$, therefore $d=0$, but it does not satisfy Assumption $A(0,d')$ (here d' should be 2) since $f(\lambda)\to\infty$ when $\lambda\to\pi/2$, except when $\delta=0$.
3. a Gaussian FARIMA$(0,d,0)$ with an additive linear trend such as $X_t=FARIMA_t+(1-2t/N)$ for $t=1,\ldots,N$ and therefore the mean value of $(X_1,\ldots,X_N)\simeq 0$. The process (X_t) is not a stationary process but $\widetilde{d}_N$ is not sensible to polynomial trends.
4. a Gaussian FARIMA$(0,d,0)$ with an additive linear trend and an additive sinusoidal seasonal component of period $T=12$ such as $X_t=FARIMA_t+(1-2t/N)+\sin(\pi\,t/6)$ for $t=1,\ldots,N$ hence the mean value of $(X_1,\ldots,X_N)\simeq 0$. The process (X_t) is not a stationary process.
5. a Gaussian process (X_t) denoted by [FGN(0.1), FGN$(d+1/2)$] and composed by a FGN with parameter $H=0.1$ for $1\leq t\leq N/2$ and a FGN with parameter $H=d+1/2$ for $1+N/2\leq t\leq N$. This is a non stationary process since there is a change of the model.
6. a process X denoted MGN(d) defined by the increments of a multifractional Brownian motion (introduced in [22]). Using the harmonizable representation, define $Y=(Y_t)_t$ by

$$Y_t:=C(t)\int_{\mathbb{R}}\frac{e^{itx}-1}{|x|^{H(t)+1/2}}dW(x)$$

where $H(\cdot)$ as well as $C(\cdot)$ are functions (the case $H(\cdot)=H$ with $H\in(0,1)$ is the case of fBm) and the complex isotropic random measure dW satisfies $dW=dW1+i\,dW2$ with $dW1$ and $dW2$ two independent real-valued Brownian measures (see more details on this part in section 7.2.2 of [27]). When $g=g_1+i\,g_2$ and $h=h_1+i\,h_2$ where g_1, h_1 and g_2, h_2 are respectively even and odd real-valued functions such that $\int_{\mathbb{R}}(g_i^2(x))dx<\infty$ and $\int_{\mathbb{R}}(h_i^2(x))dx<\infty$ $(i=1,2)$, then $\mathbb{E}\big[\big(\int_{\mathbb{R}}g(\xi)\,dW(\xi)\big)\big(\int_{\mathbb{R}}h(\xi)\,dW(\xi)\big)\big]=\int_{\mathbb{R}}g(x)\overline{h(x)}\,dx$. Here we chose $H(t)=0.5+d\,\sin(t/10)$ and $C(t)=1$. Then with $X_t=Y_{t+1}-Y_t$ 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).

7. a process X denoted MFGN(d) and defined by the increments of a multiscale fractional Brownian motion (introduced in [3]). Let $Z = (Z_t)_t$ be such that
$$Z_t := \int_{\mathbb{R}} \sigma(x) \frac{e^{itx} - 1}{|x|^{H(x)+1/2}} dW(x)$$
with dW previously defined, $H(\cdot)$ and $\sigma(\cdot)$ being piecewise constant functions. We chose $\sigma(x) = \mathbb{I}_{0.001 \leq |x| \leq 0.1}$ and $H(x) = d + 1/2$ for $0.001 \leq |x| \leq 0.04$ and $H(x) = 0.1$ for $0.04 \leq |x| \leq 3$ (such a choice was done for modeling heartbeat signals in the paper [5]). 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 a Gaussian linear process (Wold decomposition Theorem) and behaving as a FGN of parameter $d + 1/2$ for low frequencies (large time) and as a FGN of parameter 0.1 for high frequencies (small time).

For results of these simulations see Table 2.

Conclusions from Table 2: An advantage of $\widetilde{d}_N$ is its robustness with respect to smooth trends (or seasonality). Note that there exist versions of $\widehat{d}_{MS}$ and $\widehat{d}_R$ which are also efficient for processes with smooth trend or with smooth trend and seasonality (see [2] and [16]), but $\widetilde{d}_N$ has the advantage to keep the same expression for such processes or for stationary processes.

4.2. Consistency and robustness of the adaptive goodness-of-fit test

Tables 1 and 2 also provide informations concerning the adaptive goodness-of-fit test $\widetilde{W}_N$.

Conclusions from Table 1: Considering Table 1, *i.e.* processes satisfying the theoretical assumptions required in the article, the consistency properties of this test are clearly satisfactory and almost not depend on d. However note that smoother the spectral density better the goodness-of-fit test results. Note also that results of the goodness-of-fit test do not improve with N because we chose a parameter $\ell = [5 + 2 * \log(N/100)]$ increasing with N and therefore the test is more difficult to be accepted when N increases.

Conclusions from Table 2: Considering Table 2, *i.e.* processes which do not satisfy the theoretical assumptions required in the article, the adaptive goodness-of-fit test $\widetilde{W}_N$ has good robustness properties when N is large enough. For processes with heavy tail distributions or smooth trends, or GARMA processes, the test can be reasonably used when N is larger than $N \geq 1000$ (for those processes, it should be perhaps possible to prove the consistency of the test). But for the 3 last processes ([FGN(0.1), FGN($d+1/2$)], MGN(d) or MFGN(d)) there is clearly no consistency of the test statistic to a χ^2 distribution (except for $d = 0$ for MGNd): we are clearly under the test hypothesis H_1 and for $N \geq 5000$ the results of the test are almost convincing.

TABLE 2

Robustness of $\widetilde{d}_N$: $\sqrt{MSE}$ is computed from 1000, 500, 200 and 100 independent samples (for $N = 250$, 1000, 5000 and 20000 resp.). The frequency of the goodness-of-fit test acceptation is $\widetilde{p}_n = \frac{1}{n}\#\big(\widetilde{W}_N < q_{\chi^2(\ell-2)}(0.95)\big)$

$N = 250 \longrightarrow$

Model	$\sqrt{MSE}$	d ou $\delta = 0$	d ou $\delta = 0.1$	d ou $\delta = 0.2$	d ou $\delta = 0.3$	d ou $\delta = 0.4$
FARIMA$(0, d, 0)$	$\sqrt{MSE}\ \widetilde{d}_N$	0.116	0.113	0.113	0.115	0.119
with Burr innovations	$\widetilde{p}_n$	0.94	0.93	0.93	0.93	0.90
GARMA$(0, \delta, 0)$	$\sqrt{MSE}\ \widetilde{d}_N$	0.102	0.128	0.178	0.241	0.268
	$\widetilde{p}_n$	0.94	0.94	0.84	0.86	0.87
Trend + FARIMA$(0, d, 0)$	$\sqrt{MSE}\ \widetilde{d}_N$	0.110	0.114	0.102	0.102	0.107
	$\widetilde{p}_n$	0.93	0.90	0.92	0.92	0.90
Trend + Seasonality	$\sqrt{MSE}\ \widetilde{d}_N$	0.190	0.202	0.225	0.250	0.290
+ FARIMA$(0, d, 0)$	$\widetilde{p}_n$	0.07	0.12	0.16	0.16	0.17
[FGN(0.1), FGN$(d + 1/2)$]	$\sqrt{MSE}\ \widetilde{d}_N$	-	-	-	-	-
	$\widetilde{p}_n$	0.87	0.84	0.80	0.79	0.79
MGN(d)	$\sqrt{MSE}\ \widetilde{d}_N$	-	-	-	-	-
	$\widetilde{p}_n$	0.93	0.91	0.83	0.78	0.75
MFGN(d)	$\sqrt{MSE}\ \widetilde{d}_N$	-	-	-	-	-
	$\widetilde{p}_n$	0.95	0.95	0.91	0.88	0.81

$N = 1000 \longrightarrow$

Model	$\sqrt{MSE}$	d ou $\delta = 0$	d ou $\delta = 0.1$	d ou $\delta = 0.2$	d ou $\delta = 0.3$	d ou $\delta = 0.4$
FARIMA$(0, d, 0)$	$\sqrt{MSE}\ \widetilde{d}_N$	0.069	0.070	0.074	0.070	0.064
with Burr innovations	$\widetilde{p}_n$	0.92	0.95	0.88	0.90	0.88
GARMA$(0, \delta, 0)$	$\sqrt{MSE}\ \widetilde{d}_N$	0.058	0.090	0.111	0.112	0.138
	$\widetilde{p}_n$	0.91	0.82	0.85	0.88	0.88
Trend + FARIMA$(0, d, 0)$	$\sqrt{MSE}\ \widetilde{d}_N$	0.059	0.058	0.064	0.060	0.067
	$\widetilde{p}_n$	0.84	0.86	0.90	0.87	0.84
Trend + Seasonality	$\sqrt{MSE}\ \widetilde{d}_N$	0.245	0.187	0.155	0.147	0.141
+ FARIMA$(0, d, 0)$	$\widetilde{p}_n$	0.44	0.64	0.69	0.66	0.69
[FGN(0.1), FGN$(d + 1/2)$]	$\sqrt{MSE}\ \widetilde{d}_N$	-	-	-	-	-
	$\widetilde{p}_n$	0.62	0.57	0.54	0.53	0.53
MGN(d)	$\sqrt{MSE}\ \widetilde{d}_N$	-	-	-	-	-
	$\widetilde{p}_n$	0.92	0.74	0.58	0.46	0.40
MFGN(d)	$\sqrt{MSE}\ \widetilde{d}_N$	-	-	-	-	-
	$\widetilde{p}_n$	0.67	0.53	0.38	0.19	0.09

$N = 5000 \longrightarrow$

Model	$\sqrt{MSE}$	d ou $\delta = 0$	d ou $\delta = 0.1$	d ou $\delta = 0.2$	d ou $\delta = 0.3$	d ou $\delta = 0.4$
FARIMA$(0, d, 0)$	$\sqrt{MSE}\ \widetilde{d}_N$	0.047	0.040	0.051	0.044	0.027
with Burr innovations	$\widetilde{p}_n$	0.90	0.93	0.89	0.85	0.93
GARMA$(0, \delta, 0)$	$\sqrt{MSE}\ \widetilde{d}_N$	0.035	0.058	0.053	0.064	0.094
	$\widetilde{p}_n$	0.95	0.82	0.90	0.86	0.78
Trend + FARIMA$(0, d, 0)$	$\sqrt{MSE}\ \widetilde{d}_N$	0.037	0.029	0.025	0.027	0.028
	$\widetilde{p}_n$	0.79	0.88	0.89	0.94	0.95
Trend + Seasonality	$\sqrt{MSE}\ \widetilde{d}_N$	0.079	0.071	0.056	0.056	0.064
+ FARIMA$(0, d, 0)$	$\widetilde{p}_n$	0.37	0.65	0.73	0.77	0.78
[FGN(0.1), FGN$(d + 1/2)$]	$\sqrt{MSE}\ \widetilde{d}_N$	-	-	-	-	-
	$\widetilde{p}_n$	0.40	0.36	0.35	0.34	0.30
MGN(d)	$\sqrt{MSE}\ \widetilde{d}_N$	-	-	-	-	-
	$\widetilde{p}_n$	0.91	0.65	0.39	0.20	0.12
MFGN(d)	$\sqrt{MSE}\ \widetilde{d}_N$	-	-	-	-	-
	$\widetilde{p}_n$	0.05	0.04	0.02	0.02	0.05

$N = 20000 \longrightarrow$

Model	$\sqrt{MSE}$	d ou $\delta = 0$	d ou $\delta = 0.1$	d ou $\delta = 0.2$	d ou $\delta = 0.3$	d ou $\delta = 0.4$
FARIMA$(0, d, 0)$	$\sqrt{MSE}\ \widetilde{d}_N$	0.047	0.040	0.051	0.044	0.027
with Burr innovations	$\widetilde{p}_n$	0.90	0.93	0.89	0.85	0.93
GARMA$(0, \delta, 0)$	$\sqrt{MSE}\ \widetilde{d}_N$	0.013	0.030	0.033	0.039	0.040
	$\widetilde{p}_n$	0.98	0.91	0.92	0.90	0.93
Trend + FARIMA$(0, d, 0)$	$\sqrt{MSE}\ \widetilde{d}_N$	0.027	0.016	0.012	0.018	0.015
	$\widetilde{p}_n$	0.82	0.84	0.92	0.96	0.94
Trend + Seasonality	$\sqrt{MSE}\ \widetilde{d}_N$	0.079	0.041	0.040	0.031	0.048
+ FARIMA$(0, d, 0)$	$\widetilde{p}_n$	0.33	0.67	0.83	0.81	0.83
[FGN(0.1), FGN$(d + 1/2)$]	$\sqrt{MSE}\ \widetilde{d}_N$	-	-	-	-	-
	$\widetilde{p}_n$	0.35	0.33	0.29	0.28	0.27
MGN(d)	$\sqrt{MSE}\ \widetilde{d}_N$	-	-	-	-	-
	$\widetilde{p}_n$	0.93	0.51	0.26	0.15	0.06
MFGN(d)	$\sqrt{MSE}\ \widetilde{d}_N$	-	-	-	-	-
	$\widetilde{p}_n$	0.05	0.06	0.04	0.07	0.04

5. Proofs

First, we will use many times the following lemmas:

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

$$\Big|\frac{1}{a}\sum_{j=1}^{a} g\big(\frac{j}{a}\big)e^{-i\lambda\,\frac{j}{a}} - \int_0^1 g(t)e^{-i\lambda\, t}dt\Big| \leq C_g(k)\,\min\Big(\frac{1+|\lambda|^k}{a^k}\,,\,1\Big)$$

$$\textit{with} \quad C_g(k) = 2\sum_{p=0}^{k}\binom{k}{p}\sup_{x\in[0,1]}|g^{(p)}(x)|. \tag{19}$$

Proof of Lemma 1. 1/ We first prove that 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

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

Now assume that the relationship (20) is true for any $k \leq n$ with $n \in \mathbb{N}^*$. We are going to prove that (20) is also true for $k = n+1$. Indeed, assume that h satisfies Assumption $\Psi(n+1)$. Then, 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 $(t,u) \in [0,1]^2$,

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

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

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

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

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

and thus (20) is true 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 we have $|h^{(k)}(t)|\le\sum_{p=0}^{k}\binom{k}{p}|\lambda|^p|g^{(k-p)}(t)|$, and for all $\lambda\in[a,a]$,

$$\sup_{x\in[0,1]}|h^{(k)}(x)|\le\max(1,|\lambda|^k)\sum_{p=0}^{k}\binom{k}{p}\sup_{x\in[0,1]}|g^{(p)}(x)|$$

and (19) follows.

Now when $|\lambda|>a$, it is clear that

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

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

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

that induces that we cannot expect a better bound than $O(1)$ when $\lambda\in\mathbb{R}$. □

Lemma 2. *If g is a function satisfying Assumption $\Psi(k)$ with $k\ge 0$, then for all $a\in\mathbb{N}^*$ 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:

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

and (21) is satisfied. Now assume that property (21) is true for any $k \leq n$ with $n \in \mathbb{N}^*$. We are going to prove that (21) is also true for $k=n+1$. Indeed, assume that g satisfies Assumption $\Psi(n+1)$. Then, with $S_j(a,\lambda) := \sum_{\ell=0}^{j} e^{-i\lambda\ell/a} = \frac{1}{2i\sin(\lambda/2a)}\big(e^{i\lambda/2a}-e^{-i\lambda/2a}e^{-ij\lambda/a}\big)$ for $j\in\{0,1,\dots,a\}$,

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

But since g satisfies Assumption $\Psi(n+1)$ and $a\geq 1$,

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

Now, with the usual Taylor expansion

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

for $j\in\{0,1,\dots,a-1\}$. Therefore,

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

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

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

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),

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

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. First, since $(X_t)_{t\in\mathbb{Z}}$ is a stationary centered linear process, $e(a,b)=\sum_{j=1}^{a}\big(\frac{1}{\sqrt{a}}\,\psi(\frac{j}{a})\big)X_{b+j}$ 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 clear that for $a\in\mathbb{N}^*$, $(e(a,b))_{b\in\mathbb{Z}}$ is a stationary centered linear process.

Now following similar computations to those performed in [4] [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 ψ satisfies Assumption $\Psi(k)$ and therefore (3), from Lemma 1, for $u\in[-\sqrt{a},\sqrt{a}]$ and a large enough,

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

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

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

Now, using (26) and (27), since there exists $c_f > 0$ satisfying $f(\lambda) \leq c_f |\lambda|^{-2d}$ for all $\lambda \in [-\pi, \pi]$, we deduce with $F_\psi(k) = 2C_\psi(k) \sup_{x\in[0,1]} |\psi^{(k)}(x)| + C^2_\psi(k)$ and for all $d < 1/2$,

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

Now, using again (3), for a large enough,

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

Finally, from Assumption $\mathrm{A}(d,d')$ and using the definition (5) of $K_{(\psi,\alpha)}$, we obtain the following expansion:

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

because $\lim_{\lambda\to 0} \varepsilon(\lambda) = 0$ and applying Lebesgue Theorem. Then, using (28), (29) and (30), we obtain that there exists C only depending on ψ and k such as for a large enough,

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

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

Proof of Proposition 1. We decompose this proof in 4 steps. First define the normalized wavelet coefficients of X by:

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

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 in this part that $\big(\frac{N}{\sqrt{r_i r_j a_N}} \operatorname{Cov}(\widetilde{T}_N(r_i a_N), \widetilde{T}_N(r_j a_N))\big)_{1\leq i,j\leq \ell}$ converges to the asymptotic covariance matrix $\Gamma(r_1,\ldots,r_\ell,\psi,d)$ defined in (9). 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 $b = 1,\ldots,N-a$, since ψ is $[0,1]$-supported function and $\widehat{\alpha} \in \mathbb{L}^2([-\pi,\pi])$ inducing $\alpha(k) = \int_{-\pi}^{\pi} \widehat{\alpha}(\lambda) e^{ik\lambda} d\lambda$,

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

But, for $a, a' \in \mathbb{N}^*$,

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

Now,

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

since there are only two nonvanishing cases: $s_1 = s_2 = s_3 = s_4$ (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, $\mathrm{Cov}\big(\xi_{s_1}\xi_{s_2},\xi_{s_3}\xi_{s_4}\big)=\mu_4-1$ 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 the relation (34), with $\overline{z}$ denoting the conjugate of $z\in\mathbb{C}$. From the usual asymptotic behavior of Dirichlet kernel (see [10]), for f a 2π-periodic $\mathbb{L}^p$-function with $p>1$, $\lim\limits_{M\to\infty}\int_{-\pi}^{\pi}D_M(z)f(x+z)dz=f(x)$ a.e. 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 $g:\mathbb{R}^4\mapsto\mathbb{R}$ a $\mathbb{L}^p$-function 2π-periodic for each component ($p>1$),

$$\begin{aligned}\lim_{M\to\infty}&\int_{[-\pi,\pi]^4}D_M((\lambda-\lambda')+(\mu-\mu'))g(\lambda,\lambda',\mu,\mu')d\lambda d\lambda' d\mu d\mu'\\ &=\lim_{M\to\infty}\int_{[-\pi,\pi]^4}D_M(\lambda)g(\lambda+\lambda'-\mu+\mu',\lambda',\mu,\mu')d\lambda d\lambda' d\mu d\mu'\\ &=\int_{[-\pi,\pi]^3}g(\lambda'-\mu+\mu',\lambda',\mu,\mu')d\lambda' d\mu d\mu'.\end{aligned}$$

Therefore,

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

* *Case 2:* in such a case, with $s_1\neq s_2$, $\mathrm{Cov}\big(\xi_{s_1}\xi_{s_2},\xi_{s_1}\xi_{s_2}\big)=1$ and

$$\begin{aligned}C_2&=\frac{2}{a\,a'}\sum_{(s,s')\in\mathbb{Z}^2,s\neq s'}\sum_{t_1=1}^{N}\alpha(t_1-s)\psi(\frac{t_1-b}{a})\sum_{t_2=1}^{N}\alpha(t_2-s)\psi(\frac{t_2-b'}{a'})\\ &\qquad\times\sum_{t_3=1}^{N}\alpha(t_3-s')\psi(\frac{t_3-b}{a})\sum_{t_4=1}^{N}\alpha(t_4-s')\psi(\frac{t_4-b'}{a'})\\ &=-\frac{2C_1}{\mu_4-1}+\frac{1}{a\,a'}\sum_{(s,s')\in\mathbb{Z}^2}\sum_{t_1=1}^{N}\alpha(t_1-s)\psi(\frac{t_1-b}{a})\sum_{t_2=1}^{N}\alpha(t_2-s)\psi(\frac{t_2-b'}{a'})\\ &\qquad\times\sum_{t_3=1}^{N}\alpha(t_3-s')\psi(\frac{t_3-b}{a})\sum_{t_4=1}^{N}\alpha(t_4-s')\psi(\frac{t_4-b'}{a'})\end{aligned}$$

$$
\begin{aligned}
&= -\frac{2C_1}{\mu_4-1} + 2\,a\,a' \lim_{M\to\infty}\lim_{M'\to\infty}\int_{[-\pi,\pi]^4} d\lambda d\lambda' d\mu d\mu' e^{i[b(\lambda-\mu)-b'(\lambda'-\mu')]}\\
&\qquad \times \sum_{s=-M}^{M}\sum_{s=-M'}^{M'} e^{is(\lambda'-\lambda)+is'(\mu'-\mu)} S_a(a\lambda)\widehat{\alpha}(\lambda)\overline{S_{a'}(a'\lambda')\widehat{\alpha}(\lambda')}\\
&\qquad \times S_a(a\mu)\widehat{\alpha}(\mu)\overline{S_{a'}(a'\mu')\widehat{\alpha}(\mu')}\\
&= -\frac{2C_1}{\mu_4-1} + 8\pi^2\,a\,a' \int_{[-\pi,\pi]^2} e^{i(\lambda-\mu)(b-b')} S_a(a\lambda)\overline{S_{a'}(a'\lambda)}\\
&\qquad \times S_a(a\mu)\overline{S_{a'}(a'\mu)} \times \left|\widehat{\alpha}(\lambda)\right|^2 \left|\widehat{\alpha}(\mu)\right|^2 d\lambda d\mu,
\end{aligned}
$$

using the asymptotic behaviors of two Dirichlet kernels.

Now we are going back to (35). First, from (38),

$$
\begin{aligned}
&\frac{N}{(N-a)(N-a')}\sum_{b=1}^{N-a}\sum_{b'=1}^{N-a'} C_1\\
&= 2\pi\,(\mu_4-1)\,\frac{aa'\,N}{(N-a)(N-a')}\int_{[-\pi,\pi]^3} d\lambda' d\mu d\mu' F_N(a,a',\mu-\mu')\\
&\quad \times S_a(a(\lambda'-\mu+\mu'))\widehat{\alpha}(\lambda'-\mu+\mu')\overline{S_a(a\lambda')\widehat{\alpha}(\lambda')}S_{a'}(a'\mu)\widehat{\alpha}(\mu)\overline{S_{a'}(a'\mu')\widehat{\alpha}(\mu')}\\
&= 2\pi\,(\mu_4-1)\,\frac{aa``N}{(N-a)(N-a')}\int_{[-\pi,\pi]^3} d\lambda' d\mu d\mu' F_N(a,a',\mu)\\
&\quad \times S_a(a(\lambda'-\mu))\widehat{\alpha}(\lambda'-\mu)\overline{S_a(a\lambda')\widehat{\alpha}(\lambda')}S_{a'}(a'(\mu+\mu'))\widehat{\alpha}(\mu+\mu')\overline{S_{a'}(a'\mu')\widehat{\alpha}(\mu')}\\
&= 2\pi\,(\mu_4-1)\,\frac{1}{(N-a)(N-a')}\int_{-\pi a}^{\pi a}\int_{-\pi a'}^{\pi a'}\int_{-\pi N}^{\pi N} dx\,dy\,dz\,F_N(a,a',\frac{x}{N})\\
&\quad \times S_a(y-\frac{a}{N}x)\widehat{\alpha}(\frac{y}{a}-\frac{x}{N})\overline{S_a(y)}\overline{\widehat{\alpha}(\frac{y}{a})}S_{a'}(z+\frac{a'}{N}x)\widehat{\alpha}(\frac{z}{a'}+\frac{x}{N})\overline{S_{a'}(z)}\overline{\widehat{\alpha}(\frac{z}{a'})}
\end{aligned}
$$

where we used the 2π-periodicity of functions and we denoted the Fejer-type kernel:

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

Consider the function:

$$
\begin{aligned}
h_N(x,y,z) &= \mathbb{I}_{\{|x|\le\pi N,\,|y|\le\pi a,\,|z|\le\pi a'\}}\,\frac{F_N(a,a',\frac{x}{N})}{(N-a)(N-a')}\,S_a(y-\frac{a}{N}x)\overline{S_a(y)}\\
&\quad \times S_{a'}(z+\frac{a'}{N}x)\overline{S_{a'}(z)}\Big(|a|^{2d}\,\widehat{\alpha}(\frac{y}{a}-\frac{x}{N})\overline{\widehat{\alpha}(\frac{y}{a})}\Big)\,\Big(|a'|^{2d}\,\widehat{\alpha}(\frac{z}{a'}+\frac{x}{N})\overline{\widehat{\alpha}(\frac{z}{a'})}\Big).
\end{aligned}
$$

We are going to apply the Lebesgue Theorem to $\int_{\mathbb{R}^3} h_N(x,y,z)\,dx\,dy\,dz$ when $\min(a,a',N)\to\infty$ and $\max(a,a')/N \underset{N\to\infty}{\longrightarrow} 0$. Hence, for all $(x,y,z)\in(\mathbb{R}^*)^3$, from the usual expansion $\sin x \sim x$ $(x \to 0)$, from (2) and Lemmas 1 and 2,

$$h_N(x,y,z) \underset{N\to\infty}{\longrightarrow} 4\,\frac{\sin^2(x/2)}{x^2}\left|\widehat{\psi}(y)\right|^2\left|\widehat{\psi}(z)\right|^2 c_d\left|y\right|^{-2d} c_d\left|z\right|^{-2d}.$$

Now, we are going to bound $|h_N(x,y,z)|$.

Since $\frac{\sin(x)}{N\sin(x/N)} \leq \frac{3}{1+x}$ for all $x \in]0,\pi/2]$ and $N \geq 1$, and therefore,

$$\Big|4\,\frac{F_N(a,a',\frac{x}{N})}{(N-a)(N-a')}\Big| \leq \frac{9}{(1+|x/2|)^2}.$$

Moreover, always with (2) and Lemmas 1 (for $|\lambda| \leq \sqrt{a}$) and 2 (for $\sqrt{a} \leq |\lambda| \leq \pi a$), we can state that there exists a positive real number $F_\psi > 0$ such that for all $\lambda \in [-\pi a, \pi a]$,

$$\left|S_a(\lambda)\right| \leq F_\psi\,\frac{1}{1+|\lambda|^k}.$$

Therefore, using the behavior of the spectral density of x implying the inequality $|\widehat{\alpha}(\lambda)| \leq c\,|\lambda|^{-d}$ for all $\lambda \in [-\pi,0)\cup(0,\pi]$ (with $c>0$), there exists $C>0$ such as for all $(x,y,z) \in (\mathbb{R}^*)^3$,

$$\begin{aligned}\left|h_N(x,y,z)\right| \leq C\,\frac{1}{(1+|x/2|)^2}\,&\frac{(|y|\,|y-\frac{ax}{N}|)^{-d}}{(1+|y|^k)(1+|y-\frac{ax}{N}|^k)}\\ &\times\frac{(|z|\,|z+\frac{a'x}{N}|)^{-d}}{(1+|z|^k)(1+|z+\frac{a'x}{N}|^k)}\,\mathbb{I}_{\{|x|\leq \pi N,\,|y|\leq \pi a,\,|z|\leq \pi a'\}}.\end{aligned} \tag{40}$$

Finally, we have $|y-\frac{ax}{N}|^{-d} \leq |y|^{-d}$ when $y \in [-\pi a, \frac{ax}{N}]$ and $|y-\frac{ax}{2N}|^{-d} \geq |y|^{-d}$ when $y \in [\frac{ax}{2N},\pi a]$. The same partition can be done with z and thus we can write:

$$h_N(x,y,z) = h_N^{(1)}(x,y,z) + h_N^{(2)}(x,y,z) + h_N^{(3)}(x,y,z) + h_N^{(4)}(x,y,z),$$

where the respective integration domains of $h_N^{(1)}(x,y,z)$, $h_N^{(2)}(x,y,z)$ and $h_N^{(4)}(x,y,z)$ are determined by $\mathbb{I}_{\{|x|\leq\pi N,\,-\pi a\leq y\leq\frac{ax}{2N},\,-\pi a'\leq z\leq -\frac{a'x}{2N}\}}$, $\mathbb{I}_{\{|x|\leq\pi N,\,\frac{ax}{2N}\leq y\leq \pi a,\,-\pi a'\leq z\leq -\frac{a'x}{2N}\}}$ and $\mathbb{I}_{\{|x|\leq\pi N,\,\frac{ax}{2N}\leq y\leq \pi a,\,-\frac{a'x}{2N}\leq z\leq \pi a'\}}$. Then, using (40), each function $h_N^{(i)}$, $i=1,\ldots,4$, can be bounded. For instance,

$$\begin{aligned}|h_N^{(1)}(x,y,z)| &\leq C\,\frac{1}{(1+|x/2|)^2}\,\frac{|y|^{-2d}}{1+|y|^k}\,\frac{|z|^{-2d}}{1+|z|^k} \quad \text{for all } (x,y,z)\in(\mathbb{R}^*)^3\\ |h_N^{(2)}(x,y,z)| &\leq C\,\frac{1}{(1+|x/2|)^2}\,\frac{|y-\frac{ax}{2N}|^{-2d}}{1+|y-\frac{ax}{2N}|^k}\,\frac{|z|^{-2d}}{1+|z|^k} \quad \text{for all } (x,y,z)\in(\mathbb{R}^*)^3.\end{aligned}$$

The Lebesgue Theorem can be directly applied for $h_N^{(1)}(x,y,z)$ and after the change of variable $y \to y - \frac{ax}{2N}$ for $h_N^{(2)}(x,y,z)$ (which change nothing from symmetry property). As a consequence,

$$\frac{1}{(aa')^{2d}}\,\frac{N}{(N-a)(N-a')}\sum_{b=1}^{N-a}\sum_{b'=1}^{N-a'}C_1$$
$$\underset{N\to\infty}{\longrightarrow}\;8\pi\,(\mu_4-1)\,c_d^2\int_{\mathbb{R}^3}\frac{\sin^2(x/2)}{x^2}\,\big|\widehat{\psi}(y)\big|^2\,\big|\widehat{\psi}(z)\big|^2\,\big|y\big|^{-2d}\,\big|z\big|^{-2d}dxdydz.$$

Since $\int_{\mathbb{R}}\frac{\sin^2(x/2)}{x^2}\,dx=\frac{\pi}{2}$ and $(\mathbb{E}(e^2(a,0))\sim 2\pi\,c_dK_{(\psi,2d)}\,a^{2d}$ $(a\to\infty)$ from (5), we deduce with (35) that

$$N\,\frac{(\mathbb{E}(e^2(a,0))\mathbb{E}(e^2(a',0)))^{-1}}{(N-a)(N-a')}\sum_{b=1}^{N-a}\sum_{b'=1}^{N-a'}C_1\;\underset{N\to\infty}{\longrightarrow}\;\mu_4-1$$
$$\Longrightarrow\qquad\frac{N}{\sqrt{aa'}}\,\frac{(\mathbb{E}(e^2(a,0))\mathbb{E}(e^2(a',0)))^{-1}}{(N-a)(N-a')}\sum_{b=1}^{N-a}\sum_{b'=1}^{N-a'}C_1\;\underset{N\to\infty}{\longrightarrow}\;0.\qquad(41)$$

Using the Fejer kernel type F_N defined previously, we have

$$\begin{aligned}&\frac{N}{\sqrt{aa'}}\,\frac{(\mathbb{E}(e^2(a,0))\mathbb{E}(e^2(a',0)))^{-1}}{(N-a)(N-a')}\sum_{b=1}^{N-a}\sum_{b'=1}^{N-a'}C_2\\&\quad=\frac{N}{\sqrt{aa'}}\,\frac{(\mathbb{E}(e^2(a,0))\mathbb{E}(e^2(a',0)))^{-1}}{(N-a)(N-a')}\Big(-\frac{2}{\mu_4-1}\sum_{b=1}^{N-a}\sum_{b'=1}^{N-a'}C_1\\&\qquad+8\pi^2\,aa'\int_{[-\pi,\pi]^2}F_N(a,a',\lambda)S_a(a(\lambda+\mu))S_{a'}(a'(\lambda+\mu))\\&\qquad\qquad\times\overline{S_a(a\mu)S_{a'}(a'\mu)}\big|\widehat{\alpha}(\lambda+\mu)\big|^2\,\big|\widehat{\alpha}(\mu)\big|^2d\lambda d\mu\Big)\qquad(42)\\&\quad=I_1+I_2,\end{aligned}$$

where I_1 is the first sum with C_1, I_2 the second sum. Using the asymptotic expansion (41), it is clear that the term I_1 written with C_1 in (42) tends to 0 when $\min(a,a',N)\to\infty$ and $\max(a,a')/N\underset{N\to\infty}{\longrightarrow}0$. It remains to consider the limit of second term I_2. For this, let $a=ra_N$ and $a'=r'a_N$. Then, using changes of variables,

$$\begin{aligned}I_2&=8\pi^2\sqrt{rr'}\,\frac{(\mathbb{E}(e^2(ra_N,0))\mathbb{E}(e^2(r'a_N,0)))^{-1}}{(N-ra_N)(N-r'a_N)}\\&\quad\times\int_{-\pi a_N}^{\pi a_N}\int_{-\pi N}^{\pi N}F_N(ra_N,r'a_N,\frac{x}{N})S_{ra_N}(r(y+\frac{a_N}{N}x))S_{r'a_N}(r'(y+\frac{a_N}{N}x))\\&\qquad\times\overline{S_{ra_N}(ry)S_{r'a_N}(r'y)}\big|\widehat{\alpha}(\frac{y}{a_N}+\frac{x}{N})\big|^2\big|\widehat{\alpha}(\frac{y}{a_N})\big|^2dxdy\\&\underset{N\to\infty}{\longrightarrow}\;4\pi\,\frac{(rr')^{1/2-2d}}{K^2_{(\psi,2d)}}\int_{-\infty}^{\infty}\frac{\big|\widehat{\psi}(ry)\big|^2\big|\widehat{\psi}(r'y)\big|^2}{|y|^{4d}}\,dy.\end{aligned}$$

Therefore, with (41) and from (35), one deduces that:

$$\frac{N}{\sqrt{r a_N r' a_N}} \operatorname{Cov}\big(\widetilde{T}_N(r\, a_N), \widetilde{T}_N(r'\, a_N)\big) \underset{N\to\infty}{\longrightarrow} 4\pi \frac{(rr')^{1/2-2d}}{K^2_{(\psi,2d)}} \int_{-\infty}^{\infty} \frac{\big|\widehat{\psi}(r\lambda)\big|^2 \big|\widehat{\psi}(r'\lambda)\big|^2}{|\lambda|^{4d}} d\lambda. \quad (43)$$

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

Step 2 We prove here that if the distribution of the innovations $(\xi_t)_t$ is such that there exists $r_0 > 0$ satisfying

$$\mathbb{E}\big(e^{r_0 \xi_0}\big) \le \infty \qquad \text{(the so-called Cramèr condition)}, \quad (44)$$

then $(\widetilde{T}_N(r_i\, a_N))_{1\le i\le \ell} = \big(\frac{1}{N-r_i a_N} \sum_{k=1}^{N-r_i a_N} \widetilde{e}^2(r_i a_N, k)\big)_{1\le i\le \ell}$ satisfies a CLT for any sequence of positive integer numbers $(a_n)_{n\in\mathbb{N}}$ such as $\min(a_N, N/a_N) \to \infty$ when $N\to\infty$. Such theorem is implied by proving 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 $(u_i)_{1\le i\le\ell} \in \mathbb{R}^\ell$.

For establishing this result we are going to adapt a proof of [12] where central limit theorems for functionals of linear processes are proved using a decomposition with Appell polynomials. Indeed since X satisfies Assumption $A(d, d')$ and can be a two-sided linear process, martingale type results as in [29] or [11] can not be applied. Moreover, since $(a_N)_N$ is a sequence depending on N it is required to prove a central limit theorem for triangular arrays. Unfortunately the recent paper of [26] dealing with central limit theorems for arrays of decimated linear processes, and which can be applied to establish a multidimensional central limit for the variogram of wavelet coefficients associated to a multi-resolution analysis can not be applied here because in this paper this variogram is defined as in (7) with coefficients taken every n/n_j ($\simeq a_N$ with our notation) and the mean of n_j (N/a_N with our notation) coefficients is considered (with a convergence rate $\sqrt{n_j}$). Our definition of the wavelet coefficient variogram (6) is an average of $N - a_N$ terms. Then we chose to adapt the method and results of [12].

More precisely, consider first the case $\ell = 1$. We consider $H_2(x) = x^2 - 1$ the second-order Hermite polynomial and we would like to prove that

$$\Big(\frac{N}{a_N}\Big)^{1/2} \frac{1}{N - a_N} \sum_{b=1}^{N-a_N} \big(\widetilde{e}^2(a_N, b) - 1\big) \simeq \Big(\frac{1}{N a_N}\Big)^{-1/2} \sum_{b=1}^{N-a_N} H_2(\widetilde{e}(a_N, b)) \overset{\mathcal{L}}{\underset{N\to\infty}{\longrightarrow}} \mathcal{N}\big(0, \sigma^2_\psi(d)\big). \quad (45)$$

We are going to follow the proof of Proposition 6 of [12] which is devoted to CLT for polynomials in a linear process (note the Appell rank of H_2 is 2) but we have to adapt this proof to the case of triangular arrays (since $\widetilde{e}(a_N, b)$ depends on N).

For $a > 0$ and ξ_0 satisfying the Cramèr condition (44), the process $(\widetilde{e}(a,b))_{1\leq b\leq N-a}$ is a stationary linear process satisfying assumptions of the paper of Giraitis (called X_t in this article) since, from Step 1,

$$\widetilde{e}(a,b) = \sum_{s\in\mathbb{Z}} \beta_a(b-s)\xi_s \quad \text{with} \quad \beta_a(s) = \frac{\sqrt{a}}{\sqrt{\mathbb{E}e^2(a,b)}} \int_{-\pi}^{\pi} S_a(a\lambda)\widehat{\alpha}(\lambda)e^{i\lambda s}d\lambda, \quad (46)$$

where $\sum_{s\in\mathbb{Z}} \beta_a^2(s) = 1$ for any $a \in \mathbb{N}^*$. 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. The behavior of $\widehat{\beta}_a(u)$ will be considered in the sequel.

For random variables $(\eta_1, \ldots, \eta_k)$ we denote $\chi(\eta_1, \ldots, \eta_k)$ the cumulant of these random variables, defined by:

$$\chi(\eta_1, \ldots, \eta_k) = \frac{\partial^k \log\left[\mathbb{E}\exp\left(\sum_{j=1}^k a_j\eta_j\right)\right]}{\partial a_1 \partial a_2 \cdots \partial a_k}(0, 0, \ldots, 0).$$

We also recall that for any centered random variable Z satisfying the Cramèr condition (44) a sequence of Appel polynomials $(A_n^{(Z)}(x))_{n\in\mathbb{N}}$ can be defined following the relation:

$$A_n^{(Z)}(x) = \sum_{k=0}^{n} x^k \sum_{(v)(n-k)} (-1)^r \prod_{i=1}^{r} \chi^{(Z)}(|v_i|)$$

where $\chi^{(Z)}(k) = \chi(Z, Z, \ldots, Z)$ is the $k-th$ cumulant of the random variable Z, using also the convention $\sum_{(v)(0)} \cdots = 1$, $\sum_{(v)(1)} \cdots = 0$ and for $j \geq 2$, the sum $\sum_{(v)(j)}$ is taken over all partitions $(v_1, \ldots, v_r)$, $r = 1, 2, \ldots, j$ of the set $\{1, 2, \ldots, j\}$ such that $|v_i| \geq 2$ (with $|v_i| = \#(v_i)$). Note that if Z is a standard Gaussian random variable, $A_n^{(Z)}$ is the usual Hermite polynomial of degree n.

Following the proof of Proposition 6 [12], define

$$S_N^{(n)} = \sum_{b=1}^{N-a_N} A_n^{(\widetilde{e}(a_N,\cdot))}(\widetilde{e}(a_N, b))$$

where $A_n^{(\widetilde{e}(a_N,\cdot))}$ is the Appell polynomial of degree n of $\widetilde{e}(a_N,\cdot)$. Since there exists $(c_n^{(N)})_{n\geq 2}\in \ell^2$ such as $H_2=\sum_{n=2}^\infty c_n^{(N)} A_n^{(\widetilde{e}(a_N,\cdot))}$, for establishing (45), it is sufficient to prove that the cumulants of order $k\geq 3$ of $S_N(n)$ are such as

$$\chi\big(S_N^{(n(1))},\dots,S_N^{(n(k))}\big)=o\big((Na_N)^{k/2}\big) \tag{47}$$

for any $n(1),\dots,n(k)\geq 2$ (the computations of the cumulants of order 2 are induced by Step 1 of this proof).

From Proposition 5 and Corollary 1 of [12], we have

$$\chi\big(S_N^{(n(1))},\dots,S_N^{(n(k))}\big)=\sum_{\gamma\in\Gamma_0(T)} d_\gamma I_\gamma(N) \tag{48}$$

with:
- $I_{(V_1,\dots,V_r)}(N)=C\displaystyle\int_{[-\pi,\pi]^n} dx_{11}\dots dx_{kn(k)}\delta(x_{V_1})\dots\delta(x_{V_r})$
$$\times\prod_{j=1}^k D_N(x_{j1}+\cdots+x_{jn(j)})g_j(x_{j1},\dots,x_{jn(j)}) \tag{49}$$
- $g_j(x_{j1},\dots,x_{jn(j)})=\big(\widehat{\beta}_a\otimes\widehat{\beta}_a\otimes\cdots\otimes\widehat{\beta}_a\big)(x_{j1},\dots,x_{jn(j)})$ ($n(j)$ times),
- $D_N(x)=\sin(Nx/2)/\sin(x/2)$,
- $\delta(x_V)=1$ if $\displaystyle\sum_{(i,j)\in V} x_{ij}=0$ (mod. 2π), $=0$ else,

where $\Gamma_0(T)$ is the set of possible connected diagrams γ (a diagram $\gamma=(V_1,\dots,V_r)$ is a partition of the array $T=((i,j))_{1\leq i\leq k,1\leq j\leq n(i)}$ and such as $|V_i|>1$).

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

If γ is a non-Gaussian diagram, *mutatis mutandis*, we are going to follow the notation and proof of Lemma 2 of [12].

a/ Consider the case a/ of Lemma 2 of [12], corresponding to the case where there exist at least 3 different rows L_j such as $\exists V_1$, $|V_1\cap L_j|\geq 1$. Without loss of generality, we chose the diagram $V_1^*=\{(1,1),(2,1),(3,1)\}$ which is such that the rows L_1, L_2 and L_3 of the array T satisfy $|V_1^*\cap L_j|\geq 1$.

We are going to bound $I_{(V_1^*,V_2,\dots,V_r)}$ (see its formula (49)). First, $\delta(x_{V_1^*})=\delta(x_{11}+x_{21}+x_{31})$ which yields that the integration domain of I_γ in (49) is reduced to $[-\pi,\pi]^{n-1}$ since $x_{11}+x_{21}+x_{31}=0$ (mod. 2π). Then we have:

$$\begin{aligned}\big|I_{(V_1^*,V_2,\dots,V_r)}(N)\big| &\leq C\int_{[-\pi,\pi]^{n-3}} dx_{41}\dots dx_{kn(k)}\delta(x_{V_2})\dots\delta(x_{V_r})\\ &\quad\times\prod_{j=4}^k D_N(x_{j1}+\cdots+x_{jn(j)})g_j(x_{j1},\dots,x_{jn(j)})\\ &\quad\times\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} dx_{11}dx_{21}\big|D_N\big(x_{11}+\sum_{j=2}^{n(1)}x_{1j}\big)\widehat{\beta}_a(x_{11})D_N\big(x_{21}+\sum_{j=2}^{n(2)}x_{2j}\big)\end{aligned}$$

$$\times \widehat{\beta}_a(x_{21}) D_N\big(- x_{11} - x_{21} + \sum_{j=2}^{n(3)} x_{3j}\big) \widehat{\beta}_a(-x_{11} - x_{21})\big|.$$

Consider the second part (double integral) of the previous bound. As in (39) of [12], we can write:

$$\begin{aligned}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} dx_{11}dx_{21}\big|D_N\big(x_{11} + \sum_{j=2}^{n(1)} x_{1j}\big)\widehat{\beta}_a(x_{11}) D_N\big(x_{21} + \sum_{j=2}^{n(2)} x_{2j}\big)\widehat{\beta}_a(x_{21}) \\ \times D_N\big(- x_{11} - x_{21} + \sum_{j=2}^{n(3)} x_{3j}\big)\widehat{\beta}_a(-x_{11} - x_{21})\big| \\ \leq C\, \alpha_1(u_1)\alpha_2(u_2)\alpha_3(u_3).\end{aligned}$$

with $u_i = x_{i2} + \cdots + x_{in(i)}$, $\alpha_1(u) = \|D_N(\cdot + u)\widehat{\beta}_a(\cdot)\|_1$ and $\alpha_i(u) = \|D_N(\cdot + u)\widehat{\beta}_a(\cdot)\|_2$ for $i = 2, 3$. It remains to bound $\alpha_i(u)$. But, with the same approximations as in the proof of Property 1, for a_N and N large enough

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

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

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

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

For the $k - 3$ other terms (the first part of the bound of $I_{(V_1^*, V_2,\ldots,V_r)}(N)$), a result corresponding to Lemma 1 of [12] can also be obtained. Indeed, for a_N and N large enough, and with $g_{N,j} = D_N(x_{j1} + \cdots + x_{jn(j)})g_j(x_{j1}, \ldots, x_{jn(j)})$, we have

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

with $C' \ge 0$ not depending on N and a_N. Thus $\|g_{N,j}\|_2 \le C\,(Na_N)^{1/2}$ with $C \ge 0$. Using the same reasoning, there also exists $C' \ge 0$ such as $\|g'_{N,j}\|_2 \le C\,(Na_N)^{1/2}$ for $j \ge 2$ while $\|g'_{N,1}\|_2 = O(\sqrt{a_N}\,\log N) = o((Na_N)^{1/2})$. As a consequence, for γ such as $|V_1 \cap L_j| \ge 1$ for at least 3 different rows L_j,

$$I_\gamma(N) = o\big((Na_N)^{k/2}\big). \tag{50}$$

b/ For other diagrams γ, and following the part b/ of the proof of Lemma 2 in [12] (p. 32), it remains to bound the function $h(u_1,u_2) = \|D_N(u_1+\cdot)D_N(-u_2)\widehat{\beta}_a(\cdot)\|_1\|\widehat{\beta}_a(\cdot)\|_2^2$ as follows (with $x = x_{11}+x_{12}$) and with $u_1+u_2 \neq 0$:

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

But

$$\begin{aligned}
\int_{-\pi}^{\pi}\big|D_N\big(u_1+x\big)\,D_N\big(u_2-x\big)\big|dx &\le 2\int_{-2\pi N}^{2\pi N}\Big|\frac{\sin(x)}{x}\,\frac{\sin(\frac{N}{2}(u_1+u_2)-x)}{\sin(\frac{1}{2}(u_1+u_2)-\frac{x}{N})}\Big|dx \\
&\le \begin{cases} C\,\log N\big|\sin(\frac12(u_1+u_2))\big|^{-1} & \text{if } |u_1+u_2| \ge (N\log N)^{-1} \\ C\,N & \text{if } |u_1+u_2| < (N\log N)^{-1}\end{cases}.
\end{aligned}$$

Therefore,

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

and hence $\|h(u_1,u_2)\|_2 = o(Na_N)$. Then using Corollary 2 of [12], we also obtain (50).

Finally, (50) holds for all diagrams γ and it implies (47) and therefore (45).

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

$$\sqrt{\frac{N}{a_N}}\,\big(\widetilde{T}_N(r_i\, a_N)-1\big)_{1\le i\le \ell} \xrightarrow[N\to\infty]{d} \mathcal{N}\big(0\,,\,\Gamma(r_1,\ldots,r_\ell,\psi,d)\big), \tag{51}$$

with $\Gamma(r_1,\ldots,r_\ell,\psi,d)=(\gamma(r_i,r_j))_{1\le i,j\le \ell}$ given in (9).

Step 3 Now we extend the central limit obtained in Step 2 for linear processes with an innovation distribution satisfying a Cramèr condition ($\mathbb{E}\big(e^{r_0\xi_0}\big)<\infty$) to the weaker condition $\mathbb{E}\xi_0^4<\infty$ using a truncation procedure. Thus assume now that $\mathbb{E}\xi_0^4<\infty$. Let $M>0$ and define:

- $\xi_t^- = \xi_t\, \mathbb{I}_{|\xi_t|\le M}$ and $\xi_t^+ = \xi_t\, \mathbb{I}_{|\xi_t|> M}$ for $t\in\mathbb{Z}$;
- $X_t^- = \sum_{s\in\mathbb{Z}}\alpha(t-s)\,\xi_s^-$, $X_t^+ = \sum_{s\in\mathbb{Z}}\alpha(t-s)\,\xi_s^+$ for $t\in\mathbb{Z}$;
- For $(a,b)\in\mathbb{N}^*\times\mathbb{Z}$ and β_a defined in (46),

$$\widetilde{e}^-(a,b)=\frac{1}{\sqrt{\mathbb{E}e^2(a,b)}}\sum_{j=1}^{a}\Big(\frac{1}{\sqrt{a}}\,\psi(\frac{j}{a})\Big)\,X^-_{b+j}=\sum_{s\in\mathbb{Z}}\beta_a(b-s)\,\xi_s^- \tag{52}$$

$$\widetilde{e}^+(a,b)=\frac{1}{\sqrt{\mathbb{E}e^2(a,b)}}\sum_{j=1}^{a}\Big(\frac{1}{\sqrt{a}}\,\psi(\frac{j}{a})\Big)\,X^+_{b+j}=\sum_{s\in\mathbb{Z}}\beta_a(b-s)\,\xi_s^+. \tag{53}$$

Clearly $\widetilde{e}(a,b)=\widetilde{e}^+(a,b)+\widetilde{e}^-(a,b)$. We are going to prove that (51) holds. For this, we begin by writing

$$\widetilde{T}_N(r_i\,a_N)-1=\frac{1}{N-r_ia_N}\Big(\sum_{b=1}^{N-r_ia_N}\big(\widetilde{e}^-(r_ia_N,b)\big)^2-1\big) + 2\widetilde{e}^+(r_ia_N,b)\widetilde{e}^-(r_ia_N,b)+\big(\widetilde{e}^+(r_ia_N,b)\big)^2\Big). \tag{54}$$

We first prove that

$$\Big(\sqrt{\frac{N}{a_N}}\,\frac{1}{N-r_ia_N}\sum_{b=1}^{N-r_ia_N}\big(\widetilde{e}^-(r_ia_N,b)\big)^2-1\big)\Big)_{1\le i\le\ell} \xrightarrow[N\to\infty]{\mathcal{L}} \big(\gamma(r_i,r_j)\big)_{1\le i,j\le\ell}. \tag{55}$$

Indeed X^- is a linear process with innovations (ξ_t^-) satisfying the Cramèr condition, and X^- has the same (up to a multiplicative constant) spectral density as X. Therefore, from previous CLT (51),

$$\Big(\sqrt{\frac{N}{a_N}}\,\frac{1}{N-r_ia_N}\sum_{b=1}^{N-r_ia_N}\Big(\frac{(\widetilde{e}^-(r_ia_N,b))^2}{\mathbb{E}\big(\widetilde{e}^-(r_ia_N,b)\big)^2}-1\Big)\Big)_{1\le i\le\ell} \xrightarrow[N\to\infty]{\mathcal{L}} \big(\gamma(r_i,r_j)\big)_{1\le i,j\le\ell}. \tag{56}$$

Therefore it remains to prove that $\sqrt{\frac{N}{a_N}}\left(\mathbb{E}(\widetilde{e}^-(r_i a_N,b))^2-1\right)$ converges to 0 since then Slutsky Theorem will imply that CLT (55) yields. But, we have $\mathbb{E}\big(\widetilde{e}(r_i a_N,b)\big)^2=\big(\sum_{s\in\mathbb{Z}}\beta_a^2(s)\big)\,\mathbb{E}(\xi_0)^2=1$ and $\mathbb{E}\xi_0^2=1$. Then, from Cauchy-Schwarz Inequality and since $\widetilde{e}^-(r_i a_N,b)=\widetilde{e}(r_i a_N,b)-\widetilde{e}^+(r_i a_N,b)$,

$$\left|\mathbb{E}\big(\widetilde{e}^-(r_i a_N,b)\big)^2 \;-\; 1\right| \quad\le\quad 2\left(\mathbb{E}\big(\widetilde{e}^+(r_i a_N,b)\big)^2\right)^{1/2} \;+\; \mathbb{E}\big(\widetilde{e}^+(r_i a_N,b)\big)^2.$$

We have $\mathbb{E}\big(\widetilde{e}^+(r_i a_N,b)\big)^2=\big(\sum_{s\in\mathbb{Z}}\beta_a^2(s)\big)\,\mathbb{E}(\xi_0^+)^2=\mathbb{E}(\xi_0^+)^2$ from previous arguments and since we assume that the distribution of ξ_0 is symmetric. But using Hölder's and Markov's inequalities:

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

Hence, there exists $C>0$ not depending on M and N,

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

when $M=N$ (for instance). Therefore the CLT (55) holds.

From (54), it remains to prove that

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

Using the previous results, from Markov's and Hölder inequalities, this is implied when $\sqrt{\frac{N}{a_N}}\left(\mathbb{E}(\widetilde{e}^+(r_i a_N,b))^2+2\sqrt{\mathbb{E}(\widetilde{e}^+(r_i a_N,b))^2}\right)\;\underset{N\to\infty}{\longrightarrow}\;0$. Using also the inequality $\mathbb{E}(\widetilde{e}^+(r_i a_N,b))^2\le(\mathbb{E}\xi_0^4)\,M^{-2}$ obtained above, we deduce that this statement holds when $M=N$ (for instance). As a consequence, from (54), the CLT (51) holds even if the distribution of ξ_0 is symmetric and such that $\mathbb{E}\xi_0^4<\infty$.

Step 4 It remains to apply the Delta-method to (51) by considering the function $(x_1,\ldots,x_\ell)\mapsto(\log x_1,\ldots,\log x_\ell)$:

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

With $\mathbb{E}\big(e^2(a_N,0)\big)$ provided in Property 1, we obtain

$$\log\mathbb{E}\big(e^2(a_N,0)\big)=2d\,\log(a_N)+\log\big(2\pi\,c_d K_{(\psi,2d)}\big)+\frac{c_{d'}K_{(\psi,2d-d')}}{c_d K_{(\psi,2d)}}\,\frac{1}{a_N^{d'}}\big(1+o(1)\big)\qquad(57)$$

Therefore, when $\sqrt{\frac{N}{a_N}}\frac{1}{a_N^{d'}}\;\underset{N\to\infty}{\longrightarrow}\;0$, *i.e.* $N^{\frac{1}{1+2d'}}=o(a_N)$, the CLT (8) holds. □

Proof of Proposition 2. This proposition can be deduced from Proposition 1 by proving that Step 1 of its proof holds for $r_i = r_i(N)$ and $r_j = r_j(N)$ increasing to ∞ with $a_N \max(r_i(N), r_j(N)) = o(N)$ (Steps 2-3-4 are immediate because we consider a finite dimensional distribution convergence).

For Step 1, consider two sequences of integer numbers $(a_n)_{n\in\mathbb{N}}$ and $(a'_n)_{n\in\mathbb{N}}$ such as $\min(a_N, a'_N)\, N^{-1/(1+2d)} \to \infty$ and $\max(a_N, a'_N)/N \to 0$ when $N \to \infty$. Then, we would like to study the asymptotic behavior when $N \to \infty$ of $\frac{N}{\sqrt{a_N\, a'_N}} \operatorname{Cov}\big(\widetilde{T}_N(a_N), \widetilde{T}_N(a'_N)\big)$. The computations done in Proposition 1 can be used. Concerning the term C_1, we also have the asymptotic result (41). Concerning C_2 the case $a_N/a'_N \underset{N\to\infty}{\longrightarrow} 0$ implies another expansion of $I_1 + I_2$. On the one hand, we still have $I_1 \underset{N\to\infty}{\longrightarrow} 0$ from the asymptotic behavior of C_1. On the other hand, the behavior asymptotic of I_2 can be deduced from (42). Indeed, for N large enough we have:

$$\begin{aligned}
I_2 &= 8\pi^2 \frac{N}{\sqrt{a_N\, a'_N}} \frac{(\mathbb{E}(e^2(a_N,0))\mathbb{E}(e^2(a'_N,0)))^{-1}}{(N-a_N)(N-a'_N)}\, a_N\, a'_N \int_{[-\pi,\pi]^2} F_N(a_N, a'_N, \lambda) \\
&\quad \times S_{a_N}(a_N(\lambda+\mu)) S_{a'_N}(a'_N(\lambda+\mu)) \overline{S_{a_N}(a_N\mu) S_{a'_N}(a'_N\mu)} \big|\widehat{\alpha}(\lambda+\mu)\big|^2 \big|\widehat{\alpha}(\mu)\big|^2 d\lambda d\mu \\
&\le \frac{C\, a_N^{-1}}{(a_N a'_N)^{2d-1/2}} \int_{-\pi a_N}^{\pi a_N} \int_{-\pi N}^{\pi N} \frac{F_N(a_N, a'_N, \frac{x}{N})}{N^2} S_{a_N}\big(\frac{a_N}{N}x + y\big) S_{a'_N}\big(\frac{a'_N}{N}x + \frac{a'_N}{a_N}y\big) \\
&\qquad\qquad \times \overline{S_{a_N}(y) S_{a'_N}\big(\frac{a'_N}{a_N}y\big)} \big|\widehat{\alpha}\big(\frac{x}{N} + \frac{y}{a_N}\big)\big|^2 \big|\widehat{\alpha}\big(\frac{y}{a_N}\big)\big|^2 dx dy \\
&\le C' \big(\frac{a_N}{a'_N}\big)^{2d-1/2} \int_{-\pi a_N}^{\pi a_N} \frac{|\widehat{\psi}(y)|^2}{|y|^{4d}} S_{a'_N}\big(\frac{a'_N}{a_N}y\big) \overline{S_{a'_N}\big(\frac{a'_N}{a_N}y\big)} dy,
\end{aligned}$$

with C, $C' > 0$. From inequalities (26) and (27), we have for N large enough,

$$|S_{a'_N}(u)|^2 \le 2\, |\widehat{\psi}(u)|^2 \quad \text{for } 0 \le |u| \le \sqrt{a'_N} \qquad \text{and} \qquad |S_{a'_N}(u)|^2 \le \frac{D_\psi^2(k)}{u^{2k}} \quad \text{for } \sqrt{a'_N} \le |u| \le \pi a'_N.$$

Thus using also the assumptions on ψ (notably (3)), we can write:

$$\begin{aligned}
I_2 &\le C' \big(\frac{a_N}{a'_N}\big)^{2d-1/2} \Big(\int_0^{\frac{a_N}{\sqrt{a'_N}}} \frac{|\widehat{\psi}(y)|^2}{|y|^{4d}} |\widehat{\psi}\big(\frac{a'_N}{a_N}y\big)|^2 dy + \int_{\frac{a_N}{\sqrt{a'_N}}}^{\pi a_N} \frac{|\widehat{\psi}(y)|^2}{|y|^{4d}} \frac{1}{\big(\frac{a'_N}{a_N}y\big)^{2k}} dy \Big), \\
&\le C' \big(\frac{a_N}{a'_N}\big)^{2d-1/2} \Big(\int_0^{\sqrt{\frac{a_N}{a'_N}}} \frac{|y|^2}{|y|^{4d}} dy + \int_{\sqrt{\frac{a_N}{a'_N}}}^{\pi a_N} \frac{1}{|y|^{4d}} \frac{1}{\big(\frac{a'_N}{a_N}y\big)^{2k}} dy \Big) \\
&\le C'' \big(\frac{a_N}{a'_N}\big),
\end{aligned} \tag{58}$$

with $C'' > 0$. Therefore $I_2 \to 0$ when $N \to \infty$ and we can deduce that $\frac{N}{\sqrt{a_N\, a'_N}} \operatorname{Cov}\big(\widetilde{T}_N(a_N), \widetilde{T}_N(a'_N)\big) \to 0$ when $a_N/a'_N \to 0$ and $N \to \infty$. Note that same computations also imply that $\gamma(r_i, r_j) \to 0$ when $r_i/r_j \to 0$. Thus the proof of Proposition 2 is completed. □

Proof of Theorem 1. Here we use the results obtained in [4] concerning $\widehat{\alpha}_N$, $\widetilde{\alpha}_N$ and $\widetilde{d}_N$. The CLT (8) satisfied by $\big(\log\big(T_N(r_i a_N)\big)\big)_i$ in Proposition 1 is the same as the CLT (10) of [4], except that this CLT is devoted to Gaussian processes (while we consider linear processes here), the definition of T_N is not exactly the same (see Remark 2) and the asymptotic covariance Γ is not exactly the same. But the asymptotic properties of both these CLTs are the same: if $a_N = o(N^{\alpha^*})$ a bias appears in the asymptotic Gaussian distribution, and when $N^{\alpha^*} = o(a_N)$, the larger a_N the larger the asymptotic variance of $\log\big(T_N(r_i a_N)\big)$. Therefore the construction of $\widehat{\alpha}_N$ and $\widetilde{\alpha}_N$ proposed in [4] for Gaussian processes can be exactly reproduced here for linear processes. Hence Proposition 3 of [4] for Gaussian processes also holds here for linear processes and therefore the CLTs (8) and (10) are still valid when a_N is replaced by $N^{\widetilde{\alpha}_N}$. Then, 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 the limit in probability of $\widetilde{M}_N\,\Gamma(1,\ldots,\ell,d,\psi)\,\widetilde{M}_N'$, that is $\sigma_d^2(\ell)$.

The relation (15) is also an obvious consequence of Theorem 1 of [4]. □

Proof of Theorem 2. The theory of linear models can be applied: $Z_{N^{\widetilde{\alpha}_N}}\begin{pmatrix} \widetilde{c}_N \\ 2\widetilde{d}_N \end{pmatrix}$ 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 the CLT (8) 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_ℓ (identity matrix). Hence from the usual Cochran Theorem we deduce (17). □

Acknowledgments

The authors are very grateful to anonymous referees and associated editor for many relevant suggestions and corrections that strongly improve the content of the paper.

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. MR1957507

[2] Andrews, D.W.K. and Sun, Y. (2004). Adaptive local polynomial Whittle estimation of long-range dependence, *Econometrica*, 72, 569–614. MR2036733

[3] Bardet, J.M. and Bertrand, P. (2007) Definition, properties and wavelet analysis of multiscale fractional Brownian motion, *Fractals* 15, 73–87. MR2281946

[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. MR2537808

[5] Bardet, J.M., Billat V. and Kammoun, I. (2012). A new process for modeling heartbeat signals during exhaustive run with an adaptive estimator of its fractal parameters, *J. of Applied Statistics*, 39, 1331–1351. MR2914219

[6] 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. MR1819288

[7] 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. MR1956041

[8] Bhansali, R. Giraitis L. and Kokoszka, P.S. (2006). Estimation of the memory parameter by fitting fractionally-differenced autoregressive models, *J. Multivariate Analysis*, 97, 2101–2130. MR2301628

[9] Doukhan, P., Oppenheim, G. and Taqqu M.S. (Editors) (2003). *Theory and applications of long-range dependence*, Birkhäuser. MR1956041

[10] Fefferman, C. (1973). Pointwise convergence of Fourier series. *Ann. of Math.*, 98, 551–571. MR0340926

[11] 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. MR2445995

[12] Giraitis, L. (1985). The central limit theorem for functionals of a linear process. *Lith. Mat. J.*, 25, 43–57. MR0795854

[13] 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. MR1437741

[14] Giraitis, L., Robinson P.M., and Samarov, A. (2000). Adaptive semiparametric estimation of the memory parameter. *J. Multivariate Anal.*, 72, 183–207. MR1740636

[15] Henry, M. and Robinson, P.M. (1996). Bandwidth choice in Gaussian semiparametric estimation of long range dependence. In: Athens Conference on Applied Probability and Time Series Analysis, Vol. II, 220–232, Springer, New York. MR1466748

[16] Hurvich, C.M., Moulines, E. and Soulier, P. (2002). The FEXP estimator for potentially non-stationary linear time series. *Stochastic Process. Appl.*, 97, 307–340. MR1875337

[17] Iouditsky, A., Moulines, E. and Soulier, P. (2001). Adaptive estimation of the fractional differencing coefficient. *Bernoulli*, 7, 699–731. MR1867079

[18] Künsch, H. (1987). Statistical aspects of self-similar processes. Proceedings of the 1st World Congress of the Bernoulli Society, 67–74, VNU Sci. Press, Utrecht. MR1092336

[19] 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. MR2345656

[20] Moulines, E. and Soulier, P. (1999). Broadband log-periodogram regression of time series with long-range dependence, *Ann. Statist.*, 27, 1415–1439. MR1740105

[21] 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. MR1956053

[22] 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`.

[23] Robinson, P.M. (1995a). Log-periodogram regression of time series with long range dependence, *Ann. Statist.*, 23, 1048–1072. MR1345214

[24] Robinson, P.M. (1995b). Gaussian semiparametric estimation of long range dependence, *Ann. Statist.*, 23, 1630–1661. MR1370301

[25] 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. MR2560417

[26] Roueff, F. and Taqqu, M.S. (2009b). Central limit theorems for arrays of decimated linear processes. *Stochastic Process. Appl.*, 119, 3006–3041 MR2554037

[27] Samorodnitsky G. and Taqqu, M.S. (1994). Stable non-Gaussian random processes, London, U.K., Chapman & Hall. MR1280932

[28] Veitch, D., Abry, P. and Taqqu, M.S. (2003). On the Automatic Selection of the Onset of Scaling, *Fractals*, 11, 377–390. MR2028959

[29] Wu, W.B. (2002). Central limit theorems for functionals of linear processes and their applications. *Statist. Sinica*, 12, 635–649. MR1902729