Bernoulli **19**(1), 2013, 172–204
DOI: 10.3150/11-BEJ398

Central limit theorem for the robust log-regression wavelet estimation of the memory parameter in the Gaussian semi-parametric context

O. KOUAMO[1], C. LÉVY-LEDUC[2] and E. MOULINES[3]

[1]*ENSP, LIMSS, BP: 8390 Yaoundé Cameroun, Institut Telecom/Telecom ParisTech - 46, rue Barrault, 75634 Paris Cédex 13, France. E-mail: olaf.kouamo@telecom-paristech.fr*
[2]*CNRS/LTCI/Telecom ParisTech - 46, rue Barrault, 75634 Paris Cédex 13, France. E-mail: celine.levy-leduc@telecom-paristech.fr*
[3]*Institut Telecom/Telecom ParisTech - 46, rue Barrault, 75634 Paris Cédex 13, France. E-mail: eric.moulines@telecom-paristech.fr*

In this paper, we study robust estimators of the memory parameter d of a (possibly) non-stationary Gaussian time series with generalized spectral density f. This generalized spectral density is characterized by the memory parameter d and by a function f^* which specifies the short-range dependence structure of the process. Our setting is semi-parametric since both f^* and d are unknown, and d is the only parameter of interest. The memory parameter d is estimated by regressing the logarithm of the estimated variance of the wavelet coefficients at different scales. The two estimators of d that we consider are based on robust estimators of the variance of the wavelet coefficients, namely the square of the scale estimator proposed by Rousseeuw and Croux [*J. Amer. Statist. Assoc.* **88** (1993) 1273–1283] and the median of the square of the wavelet coefficients. We establish a central limit theorem, for these robust estimators as well as for the estimator of d, based on the classical estimator of the variance proposed by Moulines, Roueff and Taqqu [*Fractals* **15** (2007) 301–313]. Some Monte-Carlo experiments are presented to illustrate our claims and compare the performance of the different estimators. The properties of the three estimators are also compared to the Nile river data and the Internet traffic packet counts data. The theoretical results and the empirical evidence strongly suggest using the robust estimators as an alternative to estimate the memory parameter d of Gaussian time series.

Keywords: long-range dependence; memory parameter estimator; robustness; scale estimator; semiparametric estimation; wavelet analysis

1. Introduction

Long-range dependent processes are characterized by hyperbolically slowly decaying correlations or by a spectral density exhibiting a fractional pole at zero frequency. During the last decades, long-range dependence (and the closely related self-similarity phenomena) has been observed in many different fields, including financial econometrics, hydrology and analysis of Internet traffic. In most of these applications, however, the presence of atypical observations is

1350-7265

quite common. These outliers might be due to gross errors in the observations but also to unmodeled disturbances; see, for example, [31] and [30] for possible explanations of the presence of outliers in Internet traffic analysis. It is well known that even a few atypical observations can severely affect estimators, leading to incorrect conclusions. Hence, defining robust estimators of the memory parameter, which are less sensitive to the presence of additive outliers, is a challenging practical problem.

In this paper, we consider the class of fractional processes, denoted $M(d)$, defined as follows. Let $X = \{X_k\}_{k\in\mathbb{Z}}$ be a real-valued Gaussian process, not necessarily stationary, and denote by $\mathbf{\Delta} X$ the first order difference of X, defined by $[\mathbf{\Delta} X]_n = X_n - X_{n-1}$, $n \in \mathbb{Z}$. Define, for an integer $K \geq 1$, the Kth order difference recursively as follows: $\mathbf{\Delta}^K = \mathbf{\Delta} \circ \mathbf{\Delta}^{K-1}$. Let f^* be a bounded non-negative symmetric function, which is bounded away from zero in a neighborhood of the origin. Following [20], we say that X is an $M(d)$ process if, for any integer $K > d - 1/2$, $\mathbf{\Delta}^K X$ is stationary with spectral density function

$$f_{\mathbf{\Delta}^K X}(\lambda) = |1 - \mathrm{e}^{-\mathrm{i}\lambda}|^{2(K-d)} f^*(\lambda), \qquad \lambda \in (-\pi, \pi). \tag{1}$$

Observe that $f_{\mathbf{\Delta}^K X}(\lambda)$ in (1) is integrable since $-(K-d) < 1/2$. When $d \geq 1/2$, the process is not stationary. One can, nevertheless, associate to X the function

$$f(\lambda) = |1 - \mathrm{e}^{-\mathrm{i}\lambda}|^{-2d} f^*(\lambda), \tag{2}$$

which is called a *generalized spectral density function*. In the sequel, we assume that $f^* \in \mathcal{H}(\beta, L)$ with $0 < \beta \leq 2$ and $L > 0$, where $\mathcal{H}(\beta, L)$ denotes the set of non-negative and symmetric functions g satisfying, for all $\lambda \in (-\pi, \pi)$,

$$|g(\lambda) - g(0)| \leq L g(0) |\lambda|^{\beta}. \tag{3}$$

Our setting is semi-parametric in that both d and f^* in (2) are unknown. Here, f^* can be seen as a nuisance parameter, whereas d is the parameter of interest. This assumption on f^* is typical in the semi-parametric estimation setting; see for instance [25] and [21] and the references therein.

Different approaches have been proposed for building robust estimators of the memory parameter for $M(d)$ processes in the semi-parametric setting outlined above. Stoev *et al.* [31] have proposed a robustified wavelet-based regression estimator developed by [1]; the robustification is achieved by replacing the estimation of the wavelet coefficients variance at different scales by the median of the square of the wavelet coefficients. Another technique to robustify the wavelet regression technique has been outlined in [23], which consists of regressing the logarithm of the square of the wavelet coefficients at different scales. [18] proposed a robustified version of the log-periodogram regression estimator introduced in [14]. The method replaces the log-periodogram of the observation by a robust estimator of the spectral density in the neighborhood of the zero frequency, obtained as the discrete Fourier transform of a robust autocovariance estimator defined in [17]; the procedure is appealing and has been found to work well, but also lacks theoretical support in the semi-parametric context (note, however, that the consistency and the asymptotic normality of the robust estimator of the covariance have been discussed in [16]).

In the related context of the estimation of the fractal dimension of locally self-similar Gaussian processes, Coeurjolly [10] has proposed a robust estimator of the Hurst coefficient; instead of using the variance of the generalized discrete variations of the process (which are closely related to the wavelet coefficients, despite the facts that the motivations are quite different), this author proposes to use the empirical quantiles and the trimmed means. The consistency and asymptotic normality of this estimator is established for a class of locally self-similar processes, using a Bahadur-type representation of the sample quantile; see also [9]. Shen, Zhu and Lee [28] propose to replace the classical regression of the wavelet coefficients by a robust regression approach, based on Huberized M-estimators.

The two robust estimators of d that we propose consist of regressing the logarithm of robust variance estimators of the wavelet coefficients of the process X on a range of scales. We use, as robust variance estimators, the square of the scale estimator proposed by [27] and the square of the *mean absolute deviation* (MAD). These estimators are a robust alternative to the estimator of d, proposed by [19], which uses the same method, but with the classical variance estimator. Here, we derive a central limit theorem (CLT) for the two robust estimators of d and, by the way, we give another methodology for obtaining a central limit theorem for the estimator of d proposed by [19]. In this paper, we have also extended Theorem 4 of [2] and the Theorem of [11] to arrays of stationary Gaussian processes. These new results were very helpful in establishing the CLT for the three estimators of d that we propose.

The paper is organized as follows. In Section 2, we introduce the wavelet setting and define the wavelet-based regression estimators of d. Section 3 is dedicated to the asymptotic properties of the robust estimators of d. In this section, we derive asymptotic expansions of the wavelet spectrum estimators and provide a CLT for the estimators of d. In Section 4, some Monte-Carlo experiments are presented in order to support our theoretical claims. The Nile river data and two Internet traffic packet counts data sets, collected from the University of North Carolina, Chapel, are studied as an application in Section 5. Sections 6 and 7 detail the proofs of the theoretical results stated in Section 3.

2. Definition of the wavelet-based regression estimators of the memory parameter *d*

2.1. The wavelet setting

The wavelet setting involves two functions ϕ and ψ in $\mathrm{L}^2(\mathbb{R})$ and their Fourier transforms

$$\widehat{\phi}(\xi) \stackrel{\text{def}}{=} \int_{-\infty}^{\infty} \phi(t) \mathrm{e}^{-\mathrm{i}\xi t} \,\mathrm{d}t \quad \text{and} \quad \widehat{\psi}(\xi) \stackrel{\text{def}}{=} \int_{-\infty}^{\infty} \psi(t) \mathrm{e}^{-\mathrm{i}\xi t} \,\mathrm{d}t. \tag{4}$$

Assume the following:

(W-1) ϕ and ψ are compactly-supported, integrable and $\widehat{\phi}(0) = \int_{-\infty}^{\infty} \phi(t)\,\mathrm{d}t = 1$ and $\int_{-\infty}^{\infty} \psi^2(t)\,\mathrm{d}t = 1$.

(W-2) There exists $\alpha > 1$ such that $\sup_{\xi\in\mathbb{R}} |\widehat{\psi}(\xi)|(1+|\xi|)^{\alpha} < \infty$.

(W-3) The function ψ has M vanishing moments, that is, $\int_{-\infty}^{\infty} t^m \psi(t)\,\mathrm{d}t = 0$ for all $m = 0, \dots, M-1$.

(W-4) The function $\sum_{k\in\mathbb{Z}} k^m \phi(\cdot - k)$ is a polynomial of degree m for all $m = 0, \ldots, M-1$.

Condition (W-2) ensures that the Fourier transform $\widehat{\psi}$ decreases quickly to zero. Condition (W-3) ensures that ψ oscillates and that its scalar product with continuous-time polynomials up to degree $M-1$ vanishes. It is equivalent to asserting that the first $M-1$ derivatives of $\widehat{\psi}$ vanish at the origin and hence

$$|\widehat{\psi}(\lambda)| = \mathrm{O}(|\lambda|^M), \qquad \text{as } \lambda \to 0. \tag{5}$$

Daubechies wavelets (with $M \geq 2$) and the Coiflets satisfy these conditions; see [19]. Viewing the wavelet $\psi(t)$ as a basic template, define the family $\{\psi_{j,k}, j \in \mathbb{Z}, k \in \mathbb{Z}\}$ of translated and dilated functions

$$\psi_{j,k}(t) = 2^{-j/2}\psi(2^{-j}t - k), \qquad j \in \mathbb{Z}, k \in \mathbb{Z}. \tag{6}$$

Positive values of k translate ψ to the right, negative values to the left. The *scale index* j dilates ψ so that large values of j correspond to coarse scales and hence to low frequencies. We suppose throughout the paper that

$$(1+\beta)/2 - \alpha < d \leq M. \tag{7}$$

We now describe how the wavelet coefficients are defined in discrete time, that is, for a real-valued sequence $\{x_k, k \in \mathbb{Z}\}$ and for a finite sample $\{x_k, k = 1, \ldots, n\}$. Using the scaling function ϕ, we first interpolate these discrete values to construct the following continuous-time functions:

$$\mathbf{x}_n(t) \stackrel{\text{def}}{=} \sum_{k=1}^{n} x_k \phi(t-k) \quad \text{and} \quad \mathbf{x}(t) \stackrel{\text{def}}{=} \sum_{k\in\mathbb{Z}} x_k \phi(t-k), \qquad t \in \mathbb{R}. \tag{8}$$

Without loss of generality we may suppose that the support of the scaling function ϕ is included in $[-\mathrm{T}, 0]$ for some integer $\mathrm{T} \geq 1$. Then

$$\mathbf{x}_n(t) = \mathbf{x}(t) \qquad \text{for all } t \in [0, n - \mathrm{T} + 1].$$

We may also suppose that the support of the wavelet function ψ is included in $[0, \mathrm{T}]$. With these conventions, the support of $\psi_{j,k}$ is included in the interval $[2^j k, 2^j(k+\mathrm{T})]$. The wavelet coefficient $W_{j,k}$ at scale $j \geq 0$ and location $k \in \mathbb{Z}$ is formally defined as the scalar product in $\mathrm{L}^2(\mathbb{R})$ of the function $t \mapsto \mathbf{x}(t)$ and the wavelet $t \mapsto \psi_{j,k}(t)$.

$$W_{j,k} \stackrel{\text{def}}{=} \int_{-\infty}^{\infty} \mathbf{x}(t)\psi_{j,k}(t)\,\mathrm{d}t = \int_{-\infty}^{\infty} \mathbf{x}_n(t)\psi_{j,k}(t)\,\mathrm{d}t, \qquad j \geq 0, k \in \mathbb{Z}, \tag{9}$$

when $[2^j k, 2^j k + \mathrm{T}] \subseteq [0, n - \mathrm{T} + 1]$, that is, for all $(j,k) \in \mathcal{I}_n$, where

$$\mathcal{I}_n \stackrel{\text{def}}{=} \{(j,k)\colon\ j \geq 0, 0 \leq k \leq n_j - 1\} \qquad \text{with } n_j = [2^{-j}(n - \mathrm{T} + 1) - \mathrm{T} + 1]. \tag{10}$$

If $\mathbf{\Delta}^M X$ is stationary, then from [20], equation (17), the process $\{W_{j,k}\}_{k\in\mathbb{Z}}$ of wavelet coefficients at scale $j \geq 0$ is stationary, but the two-dimensional process $\{[W_{j,k}, W_{j',k}]^T\}_{k\in\mathbb{Z}}$ of wavelet coefficients at scales j and j', with $j \geq j'$, is not stationary. Here ^{T} denotes the transposition. This is why we consider instead the stationary, *between-scale* process

$$\{[W_{j,k}, \mathbf{W}_{j,k}(j-j')^T]^T\}_{k\in\mathbb{Z}}, \tag{11}$$

where $\mathbf{W}_{j,k}(j-j')$ is defined as follows:

$$\mathbf{W}_{j,k}(j-j') \stackrel{\text{def}}{=} [W_{j',2^{j-j'}k}, W_{j',2^{j-j'}k+1}, \ldots, W_{j',2^{j-j'}k+2^{j-j'}-1}]^T.$$

For all $j, j' \geq 1$, the covariance function of the between-scale process is given by

$$\operatorname{Cov}\big(\mathbf{W}_{j,k'}(j-j'), W_{j,k}\big) = \int_{-\pi}^{\pi} \mathrm{e}^{\mathrm{i}\lambda(k-k')}\mathbf{D}_{j,j-j'}(\lambda; f)\,\mathrm{d}\lambda, \tag{12}$$

where $\mathbf{D}_{j,j-j'}(\lambda; f)$ stands for the cross-spectral density function of this process. For further details, we refer the reader to [20], Corollary 1. The case $j = j'$ corresponds to the spectral density function of the *within-scale* process $\{W_{j,k}\}_{k\in\mathbb{Z}}$.

In the sequel, we shall use that the within- and between-scale spectral densities $\mathbf{D}_{j,j-j'}(\lambda; d)$ of the wavelet coefficients of the process X with memory parameter $d \in \mathbb{R}$ can be approximated by the corresponding spectral density of the generalized fractional Brownian motion $B_{(d)}$, defined, for $d \in \mathbb{R}$ and $u \in \mathbb{N}$, by

$$\begin{aligned}\mathbf{D}_{\infty,u}(\lambda; d) &= \big[\mathbf{D}^{(0)}_{\infty,u}(\lambda; d), \ldots, \mathbf{D}^{(2^u-1)}_{\infty,u}(\lambda; d)\big] \\ &= \sum_{l\in\mathbb{Z}} |\lambda + 2l\pi|^{-2d}\mathbf{e}_u(\lambda + 2l\pi)\overline{\widehat{\psi}(\lambda + 2l\pi)}\widehat{\psi}\big(2^{-u}(\lambda + 2l\pi)\big),\end{aligned} \tag{13}$$

where,

$$\mathbf{e}_u(\xi) \stackrel{\text{def}}{=} 2^{-u/2}\big[1, \mathrm{e}^{-\mathrm{i}2^{-u}\xi}, \ldots, \mathrm{e}^{-\mathrm{i}(2^u-1)2^{-u}\xi}\big]^T, \qquad \xi \in \mathbb{R}.$$

For further details, see page 307 of [19] and Theorem 1 and Remark 5 of [20].

2.2. Definition of the robust estimators of *d*

Let us now define robust estimators of the memory parameter d of the $M(d)$ process X from the observations $X_1, \ldots, X_n$. These estimators are derived from the construction of [1], and consist of regressing estimators of the scale spectrum

$$\sigma_j^2 \stackrel{\text{def}}{=} \operatorname{Var}(W_{j,0}), \tag{14}$$

with respect to the scale index j. The idea behind such a choice is that, by [19], equation (28),

$$\sigma_j^2 \sim C 2^{2jd}, \qquad \text{as } j \to \infty, \tag{15}$$

where C is a positive constant. More precisely, if $\widehat{\sigma}_j^2$ is an estimator of σ_j^2, based on $W_{j,0:n_j-1} = (W_{j,0}, \ldots, W_{j,n_j-1})$, then an estimator of the memory parameter d is obtained by regressing

$\log(\widehat{\sigma}_j^2)$ for a finite number of scale indices j in $\{J_0, \dots, J_0 + \ell\}$, where $J_0 = J_0(n) \geq 0$ is the lower scale and $1 + \ell \geq 2$ is the number of scales in the regression. The regression estimator can be expressed formally as

$$\widehat{d}_n(J_0, \mathbf{w}) \stackrel{\text{def}}{=} \sum_{j=J_0}^{J_0+\ell} w_{j-J_0} \log(\widehat{\sigma}_j^2), \tag{16}$$

where the vector $\mathbf{w} \stackrel{\text{def}}{=} [w_0, \dots, w_\ell]^T$ of weights satisfies $\sum_{i=0}^{\ell} w_i = 0$ and $2\log(2) \sum_{i=0}^{\ell} i w_i = 1$; see [1] and [20]. For $J_0 \geq 1$ and $\ell > 1$, one may choose, for example, $\mathbf{w}$ corresponding to the least squares regression matrix, defined by $\mathbf{w} = DB(B^T DB)^{-1}\mathbf{b}$ where

$$\mathbf{b} \stackrel{\text{def}}{=} [\,0 \quad (2\log(2))^{-1}\,], \qquad B \stackrel{\text{def}}{=} \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 0 & 1 & \cdots & \ell \end{bmatrix}^T \tag{17}$$

is the design matrix, and D is an arbitrary positive definite matrix. The best choice of D depends on the memory parameter d. However, a good approximation of this optimal matrix D is the diagonal matrix with diagonal entries $D_{i,i} = 2^{-i}$, $i = 0, \dots, \ell$; see [13] and the references therein. We will use this choice of the design matrix in the numerical experiments.

In the sequel, we shall consider three different estimators of d based on three different estimators of the scale spectrum σ_j^2, with respect to the scale index j, which are defined below.

2.2.1. *Classical scale estimator*

This estimator has been considered in the original contribution of [1] and consists of estimating the scale spectrum σ_j^2 with respect to the scale index j by the empirical variance

$$\widehat{\sigma}_{\mathrm{CL},j}^2 = \frac{1}{n_j} \sum_{i=1}^{n_j} W_{j,i}^2, \tag{18}$$

where for any j, n_j denotes the number of available wavelet coefficients at scale index j defined in (10).

2.2.2. *Median absolute deviation*

This estimator is well known to be a robust estimator of the scale, and, as mentioned in [27], it has several appealing properties: it is easy to compute and has the best possible breakdown point (50%). Since the wavelet coefficients $W_{j,i}$ are centered Gaussian observations, the square of the median absolute deviation of $W_{j,0:n_j-1}$ is defined by

$$\widehat{\sigma}_{\mathrm{MAD},j}^2 = \Big(m(\Phi) \operatorname*{med}_{0 \leq i \leq n_j - 1} |W_{j,i}| \Big)^2, \tag{19}$$

where Φ denotes the c.d.f. of a standard Gaussian random variable and

$$m(\Phi) = 1/\Phi^{-1}(3/4) = 1.4826. \tag{20}$$

The use of the median estimator to estimate the scalogram has been suggested to estimate the memory parameter in [29]; see also [24], page 420. A closely related technique is considered in [9] and [10] to estimate the Hurst coefficient of locally self-similar Gaussian processes. Note that the use of the median of the squared wavelet coefficients has been advocated to estimate the variance at a given scale in wavelet denoising applications; this technique is mentioned in [12] to estimate the scalogram of the noise in the i.i.d. context; Johnstone and Silverman [15] proposed to use this method in the long-range dependent context; the use of these estimators has not, however, been rigorously justified.

2.2.3. *The Croux and Rousseeuw estimator*

This estimator is another robust scale estimator introduced in [27]. Its asymptotic properties in several dependence contexts have been further studied in [16] and the square of this estimator is defined by

$$\widehat{\sigma}^2_{\mathrm{CR},j} = \bigl(c(\Phi)\{|W_{j,i} - W_{j,k}|;\ 0 \le i,k \le n_j - 1\}_{(k_{n_j})}\bigr)^2, \tag{21}$$

where $c(\Phi) = 2.21914$ and $k_{n_j} = \lfloor n_j^2/4 \rfloor$. That is, up to the multiplicative constant $c(\Phi)$, $\widehat{\sigma}_{\mathrm{CR},j}$ is the k_{n_j}th order statistics of the n_j^2 distances $|W_{j,i} - W_{j,k}|$ between all the pairs of observations.

3. Asymptotic properties of the robust estimators of *d*

3.1. Properties of the scale spectrum estimators

The following proposition gives an asymptotic expansion for $\widehat{\sigma}^2_{\mathrm{CL},j}$, $\widehat{\sigma}^2_{\mathrm{MAD},j}$ and $\widehat{\sigma}^2_{\mathrm{CR},j}$ defined in (18), (19) and (21), respectively. These asymptotic expansions are used for deriving central limit theorems for the different estimators of d.

Proposition 1. *Assume that X is a Gaussian $M(d)$ process with generalized spectral density function, defined in* (2), *such that $f^* \in \mathcal{H}(\beta, L)$ for some $L > 0$ and $0 < \beta \le 2$. Assume that* (W-1)–(W-4) *hold with d, α and M satisfying* (7). *Let $W_{j,k}$ be the wavelet coefficients associated to X defined by* (9). *If $n \mapsto J_0(n)$ is an integer valued sequence satisfying $J_0(n) \to \infty$ and $n2^{-J_0(n)} \to \infty$, as $n \to \infty$, then $\widehat{\sigma}^2_{*,j}$ defined in* (18), (19) *and* (21), *satisfies the following asymptotic expansion, as $n \to \infty$, for any given $\ell \ge 1$:*

$$\max_{J_0(n) \le j \le J_0(n)+\ell} \left| \sqrt{n_j}(\widehat{\sigma}^2_{*,j} - \sigma_j^2) - \frac{2\sigma_j^2}{\sqrt{n_j}} \sum_{i=0}^{n_j-1} \mathrm{IF}\left(\frac{W_{j,i}}{\sigma_j}, *, \Phi\right) \right| = o_P(1), \tag{22}$$

where $*$ denotes CL, CR *and* MAD, *σ_j^2 is defined in* (14) *and* IF *is given by*

$$\mathrm{IF}(x, \mathrm{CL}, \Phi) = \frac{1}{2} H_2(x), \tag{23}$$

$$\mathrm{IF}(x, \mathrm{CR}, \Phi) = c(\Phi)\left(\frac{1/4 - \Phi(x + 1/c(\Phi)) + \Phi(x - 1/c(\Phi))}{\int_{\mathbb{R}} \varphi(y)\varphi(y + 1/c(\Phi))\,\mathrm{d}y}\right), \tag{24}$$

$$\mathrm{IF}(x, \mathrm{MAD}, \Phi) = -m(\Phi)\left(\frac{(\mathbb{1}_{\{x \le 1/m(\Phi)\}} - 3/4) - (\mathbb{1}_{\{x \le -1/m(\Phi)\}} - 1/4)}{2\varphi(1/m(\Phi))}\right), \tag{25}$$

where φ denotes the p.d.f. of the standard Gaussian random variable, $m(\Phi)$ and $c(\Phi)$ being defined in (20) *and* (21), *respectively, and $H_2(x) = x^2 - 1$ is the second Hermite polynomial.*

The proof is postponed to Section 6.

We deduce from Proposition 1 and Theorem 6, given and proved in Section 6, the following multivariate central limit theorem for the wavelet coefficient scales.

Theorem 2. *Under the assumptions of Proposition* 1, $(\widehat{\sigma}^2_{*,J_0}, \ldots, \widehat{\sigma}^2_{*,J_0+\ell})^T$, *where $\widehat{\sigma}^2_{*,j}$ is defined in* (18), (19) *and* (21), *satisfies the following multivariate central limit theorem*:

$$\sqrt{n2^{-J_0}}2^{-2J_0 d}\left(\begin{bmatrix} \widehat{\sigma}^2_{*,J_0} \\ \widehat{\sigma}^2_{*,J_0+1} \\ \vdots \\ \widehat{\sigma}^2_{*,J_0+\ell} \end{bmatrix} - \begin{bmatrix} \sigma^2_{*,J_0} \\ \sigma^2_{*,J_0+1} \\ \vdots \\ \sigma^2_{*,J_0+\ell} \end{bmatrix}\right) \xrightarrow{d} \mathcal{N}(0, \mathbf{U}_*(d)), \tag{26}$$

where

$$\begin{aligned} \mathbf{U}_{*,i,j}(d) &= 4(f^*(0))^2 \sum_{p\ge 2} \frac{c_p^2(\mathrm{IF}_*)}{p!\mathrm{K}(d)^{p-2}} 2^{d(2+p)i\vee j} 2^{d(2-p)i\wedge j + i\wedge j} \\ &\quad \times \sum_{\tau\in\mathbb{Z}} \sum_{r=0}^{2^{|i-j|}-1} \left(\int_{-\pi}^{\pi} \mathbf{D}^{(r)}_{\infty,|i-j|}(\lambda; d) \mathrm{e}^{\mathrm{i}\lambda\tau}\,\mathrm{d}\lambda\right)^p, \qquad 0 \le i, j \le \ell. \end{aligned} \tag{27}$$

In (27), $\mathrm{K}(d) \stackrel{\text{def}}{=} \int_{\mathbb{R}} |\xi|^{-2d} |\widehat{\psi}(\xi)|\,\mathrm{d}\xi$, $\mathbf{D}_{\infty,|i-j|}(\cdot; d)$ *is the cross-spectral density defined in* (13), $c_p(\mathrm{IF}_*) = \mathbb{E}[\mathrm{IF}(X, *, \Phi) H_p(X)]$, *where H_p is the pth Hermite polynomial, and* $\mathrm{IF}(\cdot, *, \Phi)$ *is defined in* (23), (24) *and* (25).

The proof of Theorem 2 is postponed to Section 6.

Remark 1. Since, for $* = \mathrm{CL}$, $\mathrm{IF}(\cdot) = H_2(\cdot)/2$, Theorem 2 gives an alternative proof to ([19], Theorem 2) of the limiting covariance matrix of $(\widehat{\sigma}^2_{\mathrm{CL},J_0}, \ldots, \widehat{\sigma}^2_{\mathrm{CL},J_0+\ell})^T$ which is given, for $0 \le i, j \le \ell$, by

$$\mathbf{U}_{\mathrm{CL},i,j}(d) = 4\pi(f^*(0))^2 2^{4d(i\vee j)+i\wedge j} \int_{-\pi}^{\pi} \left|\mathbf{D}_{\infty,|i-j|}(\lambda; d)\right|^2 \mathrm{d}\lambda.$$

Thus, for $* = \mathrm{CR}$ and $* = \mathrm{MAD}$, we deduce the following:

$$\frac{\mathbf{U}_{\mathrm{CL},i,i}(d)}{\mathbf{U}_{*,i,i}(d)} \ge \frac{1/2}{\mathbb{E}[\mathrm{IF}_*^2(Z)]}, \tag{28}$$

where Z is a standard Gaussian random variable. With Lemma 8, we deduce, from inequality (28), that the asymptotic relative efficiency of $\widehat{\sigma}^2_{*,j}$ is larger than 36.76% when $* = \text{MAD}$ and larger than 82.27% when $* = \text{CR}$.

3.2. CLT for the robust wavelet-based regression estimator

Based on the results obtained in the previous section, we derive a central limit theorem for the robust wavelet-based regression estimators of d defined by

$$\widehat{d}_{*,n}(J_0, \mathbf{w}) \stackrel{\text{def}}{=} \sum_{j=J_0}^{J_0+\ell} w_{j-J_0} \log(\widehat{\sigma}^2_{*,j}), \tag{29}$$

where $\widehat{\sigma}^2_{*,j}$ are given for $* = \text{CL}$, MAD and CR by (18), (19) and (21), respectively.

Theorem 3. *Under the same assumptions as in Proposition 1 and if*

$$n2^{-(1+2\beta)J_0(n)} \to 0, \qquad \text{as } n \to \infty, \tag{30}$$

then $\widehat{d}_{*,n}(J_0, \mathbf{w})$ satisfies the following central limit theorem:

$$\sqrt{n2^{-J_0(n)}}\big(\widehat{d}_{*,n}(J_0, \mathbf{w}) - d\big) \xrightarrow{d} \mathcal{N}(0, \mathbf{w}^T \mathbf{V}_*(d)\mathbf{w}), \tag{31}$$

where $\mathbf{V}_*(d)$ is the $(1+\ell) \times (1+\ell)$ matrix defined by

$$\begin{aligned}\mathbf{V}_{*,i,j}(d) = \sum_{p\geq 2} \frac{4c_p^2(\text{IF}_*)}{p!\text{K}(d)^p} 2^{pd|i-j|+i\wedge j} \\ \times \sum_{\tau\in\mathbb{Z}} \sum_{r=0}^{2^{|i-j|}-1} \left(\int_{-\pi}^{\pi} \mathbf{D}^{(r)}_{\infty,|i-j|}(\lambda; d)\mathrm{e}^{\mathrm{i}\lambda\tau}\,\mathrm{d}\lambda\right)^p, \qquad 0 \leq i, j \leq \ell.\end{aligned} \tag{32}$$

In (32), $\text{K}(d) = \int_{\mathbb{R}} |\xi|^{-2d} |\widehat{\psi}(\xi)|\,\mathrm{d}\xi$, $\mathbf{D}_{\infty,|i-j|}(\cdot; d)$ *is the cross-spectral density defined in* (13), $c_p(\text{IF}_*) = \mathbb{E}[\text{IF}(X, *, \Phi)H_p(X)]$, *where H_p is the pth Hermite polynomial and* $\text{IF}(\cdot, *, \Phi)$ *is defined in* (23), (24) *and* (25).

The proof of Theorem 3 follows from Theorem 2 and the Delta method as explained in the proof of [19], Proposition 3.

Remark 2. Since it is difficult to provide a theoretical lower bound for the asymptotic relative efficiency (ARE) of $\widehat{d}_{*,n}(J_0, \mathbf{w})$ defined by

$$\text{ARE}_*(d) = \mathbf{w}^T \mathbf{V}_{\text{CL}}(d)\mathbf{w} / \mathbf{w}^T \mathbf{V}_*(d)\mathbf{w}, \tag{33}$$

Table 1. Asymptotic relative efficiency of $\widehat{d}_{n,\mathrm{CR}}$ and $\widehat{d}_{n,\mathrm{MAD}}$ with respect to $\widehat{d}_{n,\mathrm{CL}}$

d	−0.8	−0.4	−0.2	0	0.2	0.6	0.8	1	1.2	1.6	2	2.2	2.6	3
$\mathrm{ARE}_{\mathrm{CR}}(d)$	0.72	0.67	0.63	0.65	0.70	0.63	0.70	0.75	0.76	0.75	0.79	0.74	0.77	0.74
$\mathrm{ARE}_{\mathrm{MAD}}(d)$	0.48	0.39	0.38	0.36	0.43	0.39	0.44	0.47	0.45	0.50	0.48	0.5	0.49	0.49

where $* = \mathrm{CR}$ or MAD, we propose to compute this quantity empirically. We know from Theorem 3 that the expression of the limiting variance $\mathbf{w}^T \mathbf{V}_*(d)\mathbf{w}$ is valid for all Gaussian $M(d)$ processes satisfying the assumptions given in Proposition 1; thus it is enough to compute $\mathrm{ARE}_*(d)$ in the particular case of a Gaussian ARFIMA$(0, d, 0)$ process (X_t). Such a process is defined by

$$X_t = (I - B)^{-d} Z_t = \sum_{j\geq 0} \frac{\Gamma(j + d)}{\Gamma(j + 1)\Gamma(d)} Z_{t-j}, \tag{34}$$

where $\{Z_t\}$ are i.i.d. $\mathcal{N}(0, 1)$. We propose to evaluate $\mathrm{ARE}_*(d)$ empirically by simulating several Gaussian ARFIMA$(0, d, 0)$ processes for each d belonging to $[-0.8; 3]$ and computing the associated empirical standard error. With such a choice of d, both stationary and non-stationary processes are considered. The empirical values of $\mathrm{ARE}_*(d)$ are given in Table 1. The results were obtained from the observations $X_1, \ldots, X_n$ where $n = 2^{12}$, 1000 independent replications and $\mathbf{w} = DB(B^T DB)^{-1}\mathbf{b}$, where B and $\mathbf{b}$ are defined in (17) and D is a diagonal matrix with diagonal coefficients $D_{i,i} = 2^{-i}$. We used Daubechies wavelets with $M = 2$ vanishing moments when $d \leq 2$ and $M = 4$ when $d > 2$, which ensures that condition (7) is satisfied. The smallest scale is chosen to be $J_0 = 3$ and $J_0 + \ell = 8$.

From Table 1, we can see that $\widehat{d}_{n,\mathrm{CR}}$ is more efficient than $\widehat{d}_{n,\mathrm{MAD}}$ and that its asymptotic relative efficiency $\mathrm{ARE}_{\mathrm{CR}}$ ranges from 0.63 to 0.79. These results indicate empirically that the the loss of efficiency of the robust estimator $\widehat{d}_{n,\mathrm{CR}}$ is moderate and makes it an attractive robust procedure to the non-robust estimator $\widehat{d}_{n,\mathrm{CL}}$.

4. Numerical experiments

In this section the robustness properties of the different estimators of d, namely $\widehat{d}_{\mathrm{CL},n}(J_0, \mathbf{w})$, $\widehat{d}_{\mathrm{CR},n}(J_0, \mathbf{w})$ and $\widehat{d}_{\mathrm{MAD},n}(J_0, \mathbf{w})$, that are defined in Section 2.2 are investigated using Monte Carlo experiments. In the sequel, the memory parameter d is estimated from $n = 2^{12}$ observations of a Gaussian ARFIMA$(0, d, 0)$ process defined in (34), when $d = 0.2$ and 1.2 are eventually corrupted by additive outliers. We use the Daubechies wavelets with $M = 2$ vanishing moments which ensures that condition (7) is satisfied.

Let us first explain how to choose the parameters J_0 and $J_0 + \ell$. With $n = 2^{12}$, the maximal available scale is equal to 10. Choosing J_0 too small may introduce a bias in the estimation of d by Theorem 3. However, at coarse scales (large values of J_0), the number of observations may be too small, and thus choosing J_0 too large may yield a large variance. Since at scales $j = 9$ and $j = 10$, we have, respectively, 5 and 1 observations, we chose $J_0 + \ell = 8$. For the choice of J_0,

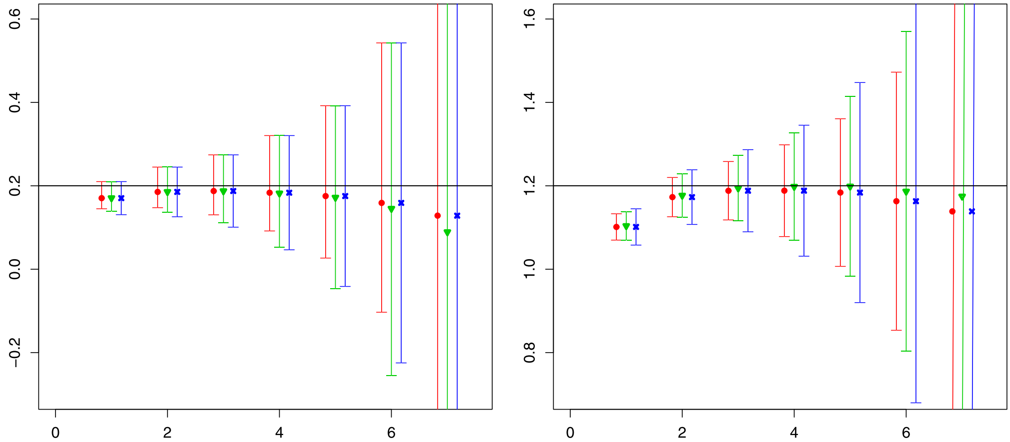

Figure 1. Confidence intervals of the estimates $\widehat{d}_{n,\mathrm{CL}}$, $\widehat{d}_{n,\mathrm{CR}}$ and $\widehat{d}_{n,\mathrm{MAD}}$ of an ARFIMA$(0,d,0)$ process with $d=0.2$ (left) and $d=1.2$ (right) for $J_0=1,\dots,8$ and $J_0+\ell=9$. For each J_0, are displayed confidence interval associated to $\widehat{d}_{n,\mathrm{CL}}$ (red), $\widehat{d}_{n,\mathrm{CR}}$ (green) and $\widehat{d}_{n,\mathrm{MAD}}$ (blue), respectively.

we use the empirical rule proposed by [19] and illustrated in Figure 1. In this figure, we display the estimates $\widehat{d}_{n,\mathrm{CL}}$, $\widehat{d}_{n,\mathrm{CR}}$ and $\widehat{d}_{n,\mathrm{MAD}}$ of the memory parameter d as well as their respective 95% confidence intervals from $J_0=1$ to $J_0=7$ with $J_0+\ell=8$. We propose to choose $J_0=3$ in both cases ($d=0.2$ and $d=1.2$) since the successive confidence intervals starting from $J_0=3$ to $J_0=7$ are such that the smallest one is included in the largest one. This choice is a way to achieve a bias/variance trade-off. A further justification for this choice of J_0 is given in Table 2, in which we provide the empirical coverage probabilities associated to $\widehat{d}_{*,n}(J_0,\mathbf{w})$, which correspond to the probability that d belongs to the 95% confidence intervals. Note that the confidence intervals in Figure 1 were computed by using Theorem 3. More precisely, an approximation of the limiting variance is obtained by computing the empirical standard error of $\sqrt{n2^{-J_0}}(\widehat{d}_{*,n}(J_0,\mathbf{w})-d)$ for the different values of J_0 by simulating and estimating the memory parameter of 5000 Gaussian ARFIMA$(0,d,0)$ processes with $d=0.2$ (left part of Figure 1) and with $d=1.2$ (right part of Figure 1).

Table 2. Coverage probabilities p_{CL}, p_{CR} and p_{MAD} of $\widehat{d}_{n,\mathrm{CL}}$ $\widehat{d}_{n,\mathrm{CR}}$ $\widehat{d}_{n,\mathrm{MAD}}$, respectively, for $n=2^{12}$ observations of an ARFIMA$(0,d,0)$ process with $d=0.2$ and $d=1.2$

	$d=0.2$							$d=1.2$						
J_0	1	2	3	4	5	6	7	1	2	3	4	5	6	7
p_{CL}	0.36	0.90	0.95	0.94	0.94	0.94	0.94	0	0.79	0.94	0.94	0.94	0.95	0.95
p_{CR}	0.52	0.92	0.95	0.95	0.95	0.94	0.95	0	0.85	0.95	0.95	0.95	0.96	0.95
p_{MAD}	0.73	0.93	0.95	0.95	0.96	0.95	0.95	0.01	0.89	0.95	0.95	0.95	0.95	0.95

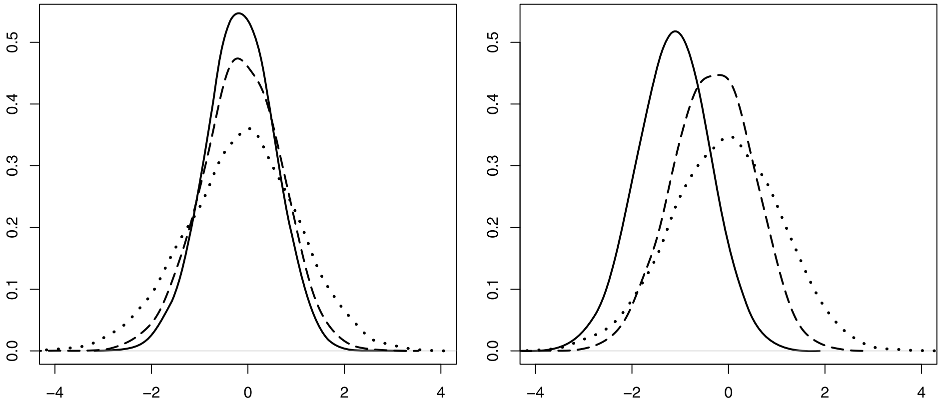

Figure 2. Empirical densities of the quantities $\sqrt{n2^{-J_0}}(\widehat{d}_{*,n} - d)$, with $* = \mathrm{CL}$ (solid line), $* = \mathrm{CR}$ (dashed line) and $* = \mathrm{MAD}$ (dotted line) of the ARFIMA$(0, 0.2, 0)$ model without outliers (left) and with 1% of outliers (right).

In the left panels of Figures 2 and 3, the empirical distribution of $\sqrt{n2^{-J_0}}(\widehat{d}_{*,n} - d)$ are displayed when $* = \mathrm{CL}$, MAD and CR for the ARFIMA$(0, d, 0)$ model with $d = 0.2$ (Figure 2) and $d = 1.2$ (Figure 3), respectively. They were computed using 5000 replications; their shapes are close to the Gaussian density (the standard deviations are of course different). In the right panels of Figures 2 and 3, the empirical distribution of $\sqrt{n2^{-J_0}}(\widehat{d}_{*,n} - d)$ are displayed when outliers are present. We introduce 1% of additive outliers in the observations; these outliers are obtained by choosing, uniformly at random, a time index and by adding to the selected observation 5 times

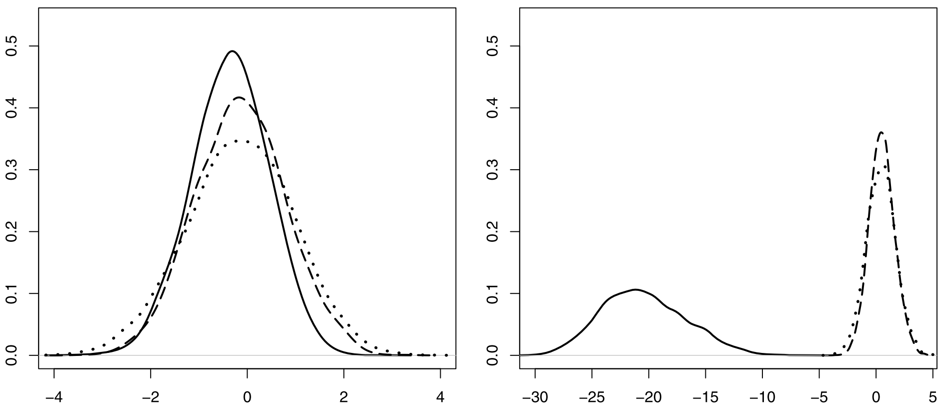

Figure 3. Empirical densities of the quantities $\sqrt{n2^{-J_0}}(\widehat{d}_{*,n} - d)$, with $* = \mathrm{CL}$ (solid line), $* = \mathrm{CR}$ (dashed line) and $* = \mathrm{MAD}$ (dotted line) of the ARFIMA$(0, 1.2, 0)$ model without outliers (left) and with 1% of outliers (right).

the standard error of the raw observations. The empirical distribution of $\sqrt{n2^{-J_0}}(\widehat{d}_{\mathrm{CL},n} - d)$ is clearly located far away from zero, especially in the non-stationary ARFIMA(0, 1.2, 0) model. One can also observe the considerable increase in the variance of the classical estimator. In sharp contrast, the distribution of the robust estimators $\sqrt{n2^{-J_0}}(\widehat{d}_{\mathrm{MAD},n} - d)$ and $\sqrt{n2^{-J_0}}(\widehat{d}_{\mathrm{CR},n} - d)$ stays symmetric and the variance stays constant.

5. Application to real data

In this section, we compare the performance of the different estimators of the long memory parameter d introduced in Section 2.2 on two different real data sets.

5.1. Nile river data

The Nile river data set is a well-known time series, which has been extensively analyzed; see [4], Section 1.4, page 20. The data consists of yearly minimal water levels of the Nile river measured at the Roda gauge, near Cairo, for the years 622–1284 AD and contains 663 observations; the units for the data, as presented by [4], are centimeters. The empirical mean and the standard deviation of the data are equal to 1148 and 89.05, respectively. The question has been raised as to whether the Nile time series contains outliers; see, for example, [3,8,25] and [18]. The test procedure developed by [8] suggests the presence of outliers at 646 AD (p-value 0.0308) and at 809 (p-value 0.0007). Another possible outliers is at 878 AD. Since the number of observations is small, in the estimation of d, we took $J_0 = 1$ and $J_0 + \ell = 6$. With this choice, we observe a significant difference between the classical estimators $\widehat{d}_{n,\mathrm{CL}} = 0.28$ (with 95% confidence interval [0.23, 0.32]) and the robust estimators $\widehat{d}_{n,\mathrm{CR}} = 0.408$ (with 95% confidence interval [0.34, 0.46]) and $\widehat{d}_{n,\mathrm{MAD}} = 0.414$ (with 95% confidence interval [0.34, 0.49]). Thus, to better understand the influence of outliers on the estimated memory parameter in practical situations, a new data set with artificial outliers was generated. Here, we replaced the presumed outliers of [8] by the value of the observation plus 10 times the standard deviation. The new memory parameter estimators are $\widehat{d}_{n,\mathrm{CL}} = 0.12$, $\widehat{d}_{n,\mathrm{CR}} = 0.4$ and $\widehat{d}_{n,\mathrm{MAD}} = 0.392$. As was expected, the values of the robust estimators remained stable. However, the classical estimator of d was significantly affected. A robust estimate of d for the Nile data is also given in [18]. The authors found 0.416, which is very close to $\widehat{d}_{n,\mathrm{CR}} = 0.408$ and $\widehat{d}_{n,\mathrm{MAD}} = 0.414$.

5.2. Internet traffic packet counts data

In this section, two Internet traffic packet counts data sets collected at the University of North Carolina, Chapel (UNC) are analyzed. These data sets are available from the website http://netlab.cs.unc.edu/public/old_research/net_lrd/. These data sets have been studied by [23].

Figure 4 (left) displays a packet count time series measured at the link of UNC on April 13, Saturday, from 7:30 p.m. to 9:30 p.m., 2002 (Sat1930). Figure 4 (right) displays the same type of time series but on April 11, a Thursday, from 1 p.m. to 3 p.m., 2002 (Thu1300). These packet counts were measured every 1 millisecond, but, for a better display, we aggregated them at 1 second.

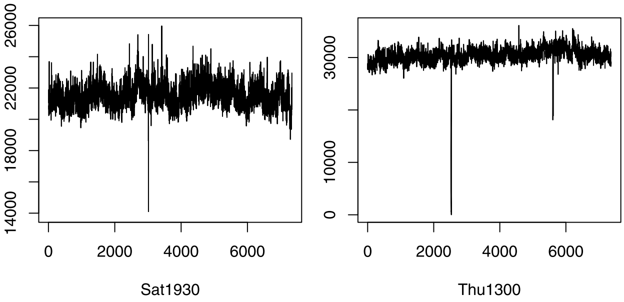

Figure 4. Packet counts of aggregated traffic every 1 second.

The maximal available scale for the two data sets is 20. Since we have less than 4 observations at this scale, we set the coarse scale $J_0 + \ell = 19$ and vary the finest scale J_0 from 1 to 17. The values of the three estimators of d are stored in Table 3 for $J_0 = 1$ to 14 as well as the standard errors of $\sqrt{n2^{-J_0}}(\widehat{d}_{n,*} - d)$ for the two data sets: Thu1300 and Sat1930. The standard errors in Table 3 were obtained as follows. For each estimated value of d, we simulated n observations of 1000 Gaussian ARFIMA$(0, d, 0)$ processes with this value of d, n being the number of observations of the data sets that we are studying (Thu1300 or Sat1930), and we computed the empirical standard errors of $\sqrt{n2^{-J_0}}(\widehat{d}_{n,*} - d)$ from these 1000 Gaussian ARFIMA$(0, d, 0)$ processes.

In Figure 5, we display the estimates $\widehat{d}_{n,\mathrm{CL}}$, $\widehat{d}_{n,\mathrm{CR}}$ and $\widehat{d}_{n,\mathrm{MAD}}$ of the memory parameter d as well as their respective 95% confidence intervals from $J_0 = 1$ to $J_0 = 14$. We propose to choose $J_0 = 9$ for Thu1300 and $J_0 = 10$ for Sat1930 since, from these values of J_0, the successive

Table 3. Estimators of d with $J_0 = 1$ to $J_0 = 14$ and $J_0 + \ell = 19$ obtained from Thu1300 and Sat1930. Here SE denotes the standard error of $\sqrt{n2^{-J_0}}(\widehat{d}_{n,*} - d)$

J_0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
							Thu1300							
$\widehat{d}_{n,\mathrm{CL}}$	0.08	0.09	0.11	0.15	0.19	0.25	0.31	0.39	0.43	0.47	0.51	0.49	0.44	0.41
$\mathrm{SE}_{\mathrm{CL}}$	(0.52)	(0.56)	(0.51)	(0.52)	(0.57)	(0.52)	(0.56)	(1.45)	(0.74)	(0.76)	(0.87)	(0.91)	(1.10)	(1.21)
$\widehat{d}_{n,\mathrm{CR}}$	0.08	0.07	0.07	0.09	0.13	0.19	0.28	0.34	0.37	0.40	0.42	0.43	0.48	0.45
$\mathrm{SE}_{\mathrm{CR}}$	(0.55)	(0.58)	(0.61)	(0.63)	(0.59)	(0.6)	(0.67)	(1.42)	(0.82)	(0.88)	(0.97)	(1.08)	(1.18)	(1.23)
$\widehat{d}_{n,\mathrm{MAD}}$	0.08	0.08	0.07	0.09	0.13	0.19	0.27	0.33	0.38	0.40	0.43	0.43	0.5	0.48
$\mathrm{SE}_{\mathrm{MAD}}$	(0.74)	(0.87)	(0.78)	(0.83)	(0.86)	(0.84)	(0.91)	(1.49)	(0.98)	(1.04)	(1.07)	(1.15)	(1.18)	(1.2)
							Sat1930							
$\widehat{d}_{n,\mathrm{CL}}$	0.05	0.06	0.08	0.11	0.14	0.17	0.23	0.28	0.33	0.36	0.37	0.39	0.42	0.42
$\mathrm{SE}_{\mathrm{CL}}$	(0.41)	(0.47)	(0.43)	(0.48)	(0.47)	(0.48)	(0.46)	(0.89)	(0.54)	(0.61)	(0.70)	(0.80)	(1.11)	(1.24)
$\widehat{d}_{n,\mathrm{CR}}$	0.06	0.06	0.06	0.09	0.12	0.16	0.23	0.3	0.34	0.38	0.4	0.42	0.44	0.42
$\mathrm{SE}_{\mathrm{CR}}$	(0.51)	(0.47)	(0.54)	(0.48)	(0.48)	(0.53)	(0.56)	(0.90)	(0.81)	(0.70)	(0.88)	(0.96)	(1.21)	(1.26)
$\widehat{d}_{n,\mathrm{MAD}}$	0.06	0.06	0.07	0.09	0.11	0.16	0.23	0.29	0.33	0.38	0.4	0.43	0.45	0.4
$\mathrm{SE}_{\mathrm{MAD}}$	(0.59)	(0.77)	(0.72)	(0.81)	(0.70)	(0.89)	(0.82)	(0.64)	(1.13)	(0.99)	(1.10)	(1.34)	(1.49)	(1.38)

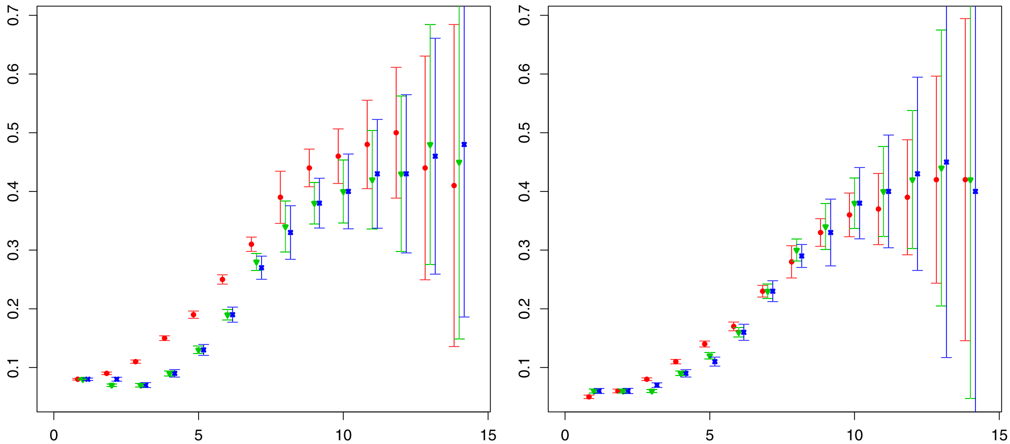

Figure 5. Confidence intervals of the estimates $\widehat{d}_{n,\mathrm{CL}}$ (red), $\widehat{d}_{n,\mathrm{CR}}$ (green) and $\widehat{d}_{n,\mathrm{MAD}}$ (blue) on the data Thu1300 (left) and Sat1930 (right) for $J_0 = 1, \dots, 14$ and $J_0 + \ell = 19$.

confidence intervals are such that the smallest one is included in the largest one (for the robust estimators). Note that Park and Park [23] chose the same values of J_0 using another methodology. For these values of J_0, we obtain $\widehat{d}_{n,\mathrm{CL}} = 0.43$ (with 95% confidence interval [0.412, 0.443]), $\widehat{d}_{n,\mathrm{CR}} = 0.37$ (with 95% confidence interval [0.358, 0.385]) and $\widehat{d}_{n,\mathrm{MAD}} = 0.38$ with (95% confidence interval [0.362, 0.397]) for Thu1300 and $\widehat{d}_{n,\mathrm{CL}} = 0.36$ (with 95% confidence interval [0.345, 0.374]), $\widehat{d}_{n,\mathrm{CR}} = \widehat{d}_{n,\mathrm{MAD}} = 0.38$ (with 95% confidence intervals [0.361, 0.398] for CR and [0.357, 0.402] for MAD) for Sat1930. These values are similar to the one found by [23].

With this choice of J_0 for Thu1300, we observe a significant difference between the classical estimator and the robust estimators. Thus to better understand the influence of outliers on the estimated memory parameter, a new data set with artificial outliers was generated. The Thu1300 time series shows two spikes shooting down. Especially, the first downward spike hits zero. Park *et al.* [22] have shown that this dropout lasted 8 seconds. Outliers are introduced by dividing by 6 the 8000 observations in this period. The new memory parameter estimators are $\widehat{d}_{n,\mathrm{CL}} = 0.445$, $\widehat{d}_{n,\mathrm{CR}} = 0.375$ and $\widehat{d}_{n,\mathrm{MAD}} = 0.377$. As for the Nile river data, the classical estimator was affected while the robust estimators remain stable.

6. Proofs

Theorem 4 is an extension of [2], Theorem 4, to arrays of stationary Gaussian processes in the unidimensional case, and Theorem 5 extends the result of [11] to arrays of stationary Gaussian processes. These two theorems are useful for the proof of Proposition 1.

Theorem 4. *Let $\{X_{j,i}, j \geq 1, i \geq 0\}$ be an array of standard stationary Gaussian processes such that for a fixed $j \geq 1$, $(X_{j,i})_{i\geq 0}$ has a spectral density f_j and an autocorrelation function ρ_j, defined by $\rho_j(k) = \mathbb{E}(X_{j,0}X_{j,k})$, for all $k \geq 0$. Assume also that there exists a sequence $\{u_j\}_{j\geq 1}$*

tending to zero as j tends to infinity, such that, for all $j \geq 1$,

$$\sup_{\lambda \in (-\pi, \pi)} |f_j(\lambda) - g_\infty(\lambda)| \leq u_j, \tag{35}$$

where g_∞ is a 2π-periodic function which is bounded on $(-\pi, \pi)$ and continuous at the origin. Let h be a function on $\mathbb{R}$ such that $\mathbb{E}[h(X)^2] < \infty$, where X is a standard Gaussian random variable, and of Hermite rank $\tau \geq 1$. Let $\{n_j\}_{j \geq 1}$ be a sequence of integers such that n_j tends to infinity as j tends to infinity. Then,

$$\frac{1}{\sqrt{n_j}} \sum_{i=1}^{n_j} h(X_{j,i}) \xrightarrow{d} \mathcal{N}(0, \widetilde{\sigma}^2), \qquad \text{as } j \to \infty, \tag{36}$$

where

$$\widetilde{\sigma}^2 = \lim_{j \to \infty} \operatorname{Var}\left(\frac{1}{\sqrt{n_j}} \sum_{i=1}^{n_j} h(X_{j,i}) \right) = 2\pi \sum_{\ell \geq \tau} \frac{c_\ell^2}{\ell!} g_\infty^{\star \ell}(0).$$

In the previous equality, $c_\ell = \mathbb{E}[h(X) H_\ell(X)]$, where H_ℓ is the ℓth Hermite polynomial, and X is a standard Gaussian random variable.

Proof. By observing that $h(x) = \sum_{\ell \geq \tau} c_\ell H_\ell(x)/\ell!$, we start by proving that, for any fixed t,

$$\frac{\sum_{i=1}^{n_j} \sum_{\tau \leq \ell \leq t} (c_\ell/\ell!) H_\ell(X_{j,i})}{\sqrt{\operatorname{Var}(\sum_{i=1}^{n_j} \sum_{\tau \leq \ell \leq t} (c_\ell/\ell!) H_\ell(X_{j,i}))}} \xrightarrow{d} \mathcal{N}(0, 1), \qquad \text{as } n_j \to \infty. \tag{37}$$

Using Mehler's formula given in [6, (2.1)], we have

$$\begin{aligned} \operatorname{Var}\left(\sum_{i=1}^{n_j} \sum_{\tau \leq \ell \leq t} \frac{c_\ell}{\ell!} H_\ell(X_{j,i}) \right) &= \sum_{i_1, i_2 = 1}^{n_j} \sum_{\tau \leq \ell_1, \ell_2 \leq t} \frac{c_{\ell_1} c_{\ell_2}}{\ell_1! \ell_2!} \mathbb{E}[H_{\ell_1}(X_{j,i_1}) H_{\ell_2}(X_{j,i_2})] \\ &= \sum_{\tau \leq \ell \leq t} \frac{c_\ell^2}{\ell!} \left[\sum_{i_1, i_2 = 1}^{n_j} \rho_j^\ell (i_2 - i_1) \right]. \end{aligned} \tag{38}$$

Since the Gaussian distribution is uniquely determined by its moments, it is enough to show the convergence of moments to prove (37), that is for $p \geq 1$,

$$\frac{\mathbb{E}[(\sum_{i=1}^{n_j} \sum_{\tau \leq \ell \leq t} (c_\ell/\ell!) H_\ell(X_{j,i}))^{2p+1}]}{(\sum_{\tau \leq \ell \leq t} (c_\ell^2/\ell!) [\sum_{i_1, i_2 = 1}^{n_j} \rho_j^\ell (i_2 - i_1)])^{(2p+1)/2}} \to 0, \qquad \text{as } n_j \to \infty \quad \text{and} \tag{39}$$

$$\frac{\mathbb{E}[(\sum_{i=1}^{n_j} \sum_{\tau \leq \ell \leq t} (c_\ell/\ell!) H_\ell(X_{j,i}))^{2p}]}{(\sum_{\tau \leq \ell \leq t} (c_\ell^2/\ell!) [\sum_{i_1, i_2 = 1}^{n_j} \rho_j^\ell (i_2 - i_1)])^{p}} \to \frac{(2p)!}{p! 2^p}, \qquad \text{as } n_j \to \infty. \tag{40}$$

Observe that, for all m in $\mathbb{N}^*$,

$$\begin{aligned}\mathbb{E}&\left[\left(\sum_{i=1}^{n_j}\sum_{\tau\le\ell\le t}\frac{c_\ell}{\ell!}H_\ell(X_{j,i})\right)^m\right]\\&=\sum_{1\le i_1,\dots,i_m\le n_j}\sum_{\tau\le\ell_1,\dots,\ell_m\le t}\frac{c_{\ell_1}\cdots c_{\ell_m}}{\ell_1!\cdots\ell_m!}\mathbb{E}[H_{\ell_1}(X_{j,i_1})\cdots H_{\ell_m}(X_{j,i_m})].\end{aligned}\tag{41}$$

By [26], equation (33), page 69,

$$\mathbb{E}[H_{\ell_1}(X_{j,i_1})\cdots H_{\ell_m}(X_{j,i_m})]=\ell_1!\cdots\ell_m!\sum_{\{\ell_1,\dots,\ell_m\}}\frac{\rho_j^\nu}{\nu!},\tag{42}$$

where it is understood that $\rho_j^\nu=\prod_{1\le q<k\le m}\rho_j^{\nu_{q,k}}(i_q-i_k)$, $\nu!=\prod_{1\le q<k\le m}\nu_{q,k}!$, and $\sum_{\{\ell_1,\dots,\ell_m\}}$ indicates that we are to sum over all symmetric matrices ν with nonnegative integer entries, $\nu_{ii}=0$ and the row sums equal to $\ell_1,\dots,\ell_m$.

(1) We start with the case where $m=2p+1$. By Lemma 9, we get that

$$\lim_{n_j\to\infty}\frac{1}{n_j}\sum_{\tau\le\ell\le t}\frac{c_\ell^2}{\ell!}\left[\sum_{i_1,i_2=1}^{n_j}\rho_j^\ell(i_2-i_1)\right]=2\pi\sum_{\tau\le\ell\le t}\frac{c_\ell^2}{\ell!}g_\infty^{\star\ell}(0).\tag{43}$$

Thus, in order to prove (39), it is enough to prove that

$$\lim_{n_j\to\infty}\frac{1}{n_j^{p+1/2}}\mathbb{E}\left[\left(\sum_{i=1}^{n_j}\sum_{\tau\le\ell\le t}\frac{c_\ell}{\ell!}H_\ell(X_{j,i})\right)^{2p+1}\right]=0.\tag{44}$$

Let us now prove that

$$\lim_{n_j\to\infty}\frac{1}{n_j^{p+1/2}}\sum_{1\le i_1,\dots,i_{2p+1}\le n_j}\sup_\nu\prod_{1\le q<k\le 2p+1}\rho_j^{\nu_{q,k}}(i_q-i_k)=0,\tag{45}$$

where $\sup_\nu$ indicates that we are taking the supremum over all symmetric matrices ν with nonnegative integer entries, $\nu_{ii}=0$ and the row sums equal to $\ell_1,\dots,\ell_{2p+1}$. By (41) and (42), (44) is a consequence of (45).

Let us first address the case where $|\{i_1,\dots,i_{2p+1}\}|=2p+1$; that is, the indices $i_1,\dots,i_{2p+1}$ are all different. Using that $\rho_j^{\nu_{q,k}}(i_q-i_k)=\int_{-\pi}^{\pi}\mathrm{e}^{\mathrm{i}\lambda_{q,k}(i_q-i_k)}f_j^{\star\nu_{q,k}}(\lambda_{q,k})\,\mathrm{d}\lambda_{q,k}$ and the notation $\mathcal{D}_{n_j}(\lambda)=\sum_{r=1}^{n_j}\mathrm{e}^{\mathrm{i}\lambda r}$, we obtain that

$$\begin{aligned}&\sum_{1\le i_1,\dots,i_{2p+1}\le n_j}\prod_{1\le q<k\le 2p+1}\rho_j^{\nu_{q,k}}(i_q-i_k)\\&\quad=\int_{[-\pi,\pi]^{p(2p+1)}}\mathcal{D}_{n_j}\left(\sum_{k=2}^{2p+1}\lambda_{1,k}\right)\mathcal{D}_{n_j}\left(-\lambda_{1,2}+\sum_{k=3}^{2p+1}\lambda_{2,k}\right)\end{aligned}\tag{46}$$

$$\times \mathcal{D}_{n_j}\left(-\lambda_{1,3}-\lambda_{2,3}+\sum_{k=4}^{2p+1}\lambda_{3,k}\right)\cdots$$
$$\times \mathcal{D}_{n_j}\left(-\sum_{q=1}^{2p}\lambda_{q,2p+1}\right)\prod_{1\le q<k\le 2p+1} f_j^{\star\nu_{q,k}}(\lambda_{q,k})\,\mathrm{d}\lambda_{q,k},$$

with the convention that $f_j^{\star 0}(\lambda)\,\mathrm{d}\lambda$ is a Dirac measure at 0. The number of $\mathcal{D}_{n_j}$ in the previous product is equal to $2p+1$. Since the ℓ_is are greater than 1, there is at least one $\lambda_{q,k}$ in each $\mathcal{D}_{n_j}$. Moreover, there exists at least one $\mathcal{D}_{n_j}$ having a sum of at least two $\lambda_{q,k}$s as argument; otherwise the matrix ν would have a null line which is impossible since the ℓ_is are greater than 1. The right-hand side of (46) can be bounded above by using that $ab \le (a^2+b^2)/2$ with

$$a = \mathcal{D}_{n_j}\left(\sum_{k=2}^{2p+1}\lambda_{1,k}\right)\left|\mathcal{D}_{n_j}\left(-\lambda_{1,2}+\sum_{k=3}^{2p+1}\lambda_{2,k}\right)\right|^{1/2},$$

and

$$b = \left|\mathcal{D}_{n_j}\left(-\lambda_{1,2}+\sum_{k=3}^{2p+1}\lambda_{2,k}\right)\right|^{1/2}\mathcal{D}_{n_j}\left(-\lambda_{1,3}-\lambda_{2,3}+\sum_{k=4}^{2p+1}\lambda_{3,k}\right)\cdots\mathcal{D}_{n_j}\left(-\sum_{q=1}^{2p}\lambda_{q,2p+1}\right).$$

Then, using Lemma 10 and (35), we get (45). Actually, the $\mathcal{D}_{n_j}$, which is common to a and b, can be any $\mathcal{D}_{n_j}$ having a sum of at least two $\lambda_{q,k}$s as argument. Such a $\mathcal{D}_{n_j}$ does exist according to the previous remark.

If $|\{i_1,\dots,i_{2p+1}\}| < 2p+1$, the number of $\mathcal{D}_{n_j}$ appearing in the right-hand side of (46) is equal to $|\{i_1,\dots,i_{2p+1}\}|$, and thus, using the same arguments as previously, we can also conclude, in this case, that (45) holds.

(2) Let us now study the case where m is even; that is, $m=2p$ with $p\ge 1$.

We shall prove that, among all the terms in the right-hand side of (42), the leading ones correspond to the case where we have p pairs of equal indices in the set $\{\ell_1,\dots,\ell_{2p}\}$, that is, for instance, $\ell_1=\ell_2$, $\ell_3=\ell_4,\dots,\ell_{2p-1}=\ell_{2p}$ and $\nu_{1,2}=\ell_1$, $\nu_{3,4}=\ell_3,\dots,\nu_{2p-1,2p}=\ell_{2p-1}$, the others $\nu_{i,j}$ being equal to zero. This gives

$$(\ell_2!)^2\cdots(\ell_{2p}!)^2\frac{\rho_j(i_2-i_1)^{\ell_2}\rho_j(i_4-i_3)^{\ell_4}\cdots\rho_j(i_{2p}-i_{2p-1})^{\ell_{2p}}}{\ell_2!\cdots\ell_{2p}!}.$$

The corresponding term in (41) with $m=2p$ is given by

$$\sum_{1\le i_1,\dots,i_{2p}\le n_j}\ \sum_{\tau\le \ell_2,\ell_4,\dots,\ell_{2p}\le t}\frac{c_{\ell_2}^2 c_{\ell_4}^2\cdots c_{\ell_{2p}}^2}{\ell_2!\ell_4!\cdots\ell_{2p}!}\rho_j(i_2-i_1)^{\ell_2}\rho_j(i_4-i_3)^{\ell_4}\cdots\rho_j(i_{2p}-i_{2p-1})^{\ell_{2p}}$$
$$=\left[\sum_{l\ge\tau}\frac{c_\ell^2}{\ell!}\left(\sum_{i_1,i_2=1}^{n_j}\rho_k^\ell(i_2-i_1)\right)\right]^p,$$

which corresponds to the denominator in the left-hand side of (40). Since there exists exactly $(2p)!/(2^p p!)$ possibilities to have pairs of equal indices among $2p$ indices, we obtain (40) if we prove that the other terms can be neglected.

We use the same line of reasoning as the one used in the case where m is odd and prove that

$$\lim_{n_j\to\infty} \frac{1}{n_j^p} \sum_{1\le i_1,\dots,i_{2p}\le n_j} \sup_{\nu} \prod_{1\le q<k\le 2p} \rho_j^{\nu_{q,k}}(i_q - i_k) = 0, \tag{47}$$

where $\sup_\nu$ indicates that we are taking the supremum over all symmetric matrices ν with non-negative integer entries such that $\nu_{ii}=0$, the row sums equal to $\ell_1,\dots,\ell_{2p}$ and such that there are at least two non-null values of ν_{qk} on a row of ν.

Let us first address the case where $|\{i_1,\dots,i_{2p}\}| = 2p$. Using the notation $\mathcal{D}_{n_j}(\lambda) = \sum_{r=1}^{n_j} \mathrm{e}^{\mathrm{i}\lambda r}$ and that $\rho_j^{\nu_{q,k}}(i_q - i_k) = \int_{-\pi}^{\pi} \mathrm{e}^{\mathrm{i}\lambda_{q,k}(i_q-i_k)} f_j^{\star\nu_{q,k}}(\lambda_{q,k})\,\mathrm{d}\lambda_{q,k}$, we obtain that

$$\begin{aligned}
&\sum_{1\le i_1,\dots,i_{2p}\le n_j} \prod_{1\le q<k\le 2p} \rho_j^{\nu_{q,k}}(i_q - i_k)\\
&\quad= \int_{[-\pi,\pi]^{p(2p-1)}} \mathcal{D}_{n_j}\left(\sum_{k=2}^{2p}\lambda_{1,k}\right)\mathcal{D}_{n_j}\left(-\lambda_{1,2}+\sum_{k=3}^{2p}\lambda_{2,k}\right)\\
&\qquad\times \mathcal{D}_{n_j}\left(-\lambda_{1,3}-\lambda_{2,3}+\sum_{k=4}^{2p}\lambda_{3,k}\right)\cdots\mathcal{D}_{n_j}\left(-\sum_{q=1}^{2p-1}\lambda_{q,2p}\right)\\
&\qquad\times \prod_{1\le q<k\le 2p} f_j^{\star\nu_{q,k}}(\lambda_{q,k})\,\mathrm{d}\lambda_{q,k},
\end{aligned} \tag{48}$$

with the convention that $f_j^{\star 0}(\lambda)\,\mathrm{d}\lambda$ is a Dirac measure at 0. The number of $\mathcal{D}_{n_j}$ in the previous product is equal to $2p$. Since the ℓ_is is greater than 1, there are at least one $\lambda_{q,k}$ in each $\mathcal{D}_{n_j}$. Moreover, there exists at least one $\mathcal{D}_{n_j}$, having a sum of at least two $\lambda_{q,k}$s, as argument; otherwise, there are p pairs of equal indices in the set $\{\ell_1,\dots,\ell_{2p}\}$, which corresponds to the case previously addressed. To conclude the proof of (47), we use the same arguments as those used in the odd case.

If $|\{i_1,\dots,i_{2p}\}| < 2p$, the number of $\mathcal{D}_{n_j}$ appearing in the right-hand side of (48) is equal to $|\{i_1,\dots,i_{2p}\}|$, and thus, using the same arguments as previously, we can also conclude in this case that (47) holds.

We conclude the proof by applying [7], Proposition 6.3.9. By (37), (38) and (43), Assumption (i) of [7], Proposition 6.3.9, holds true. Assumption (ii) comes from $\sum_{\tau\le\ell\le t} c_\ell^2 g_\infty^{\star\ell}(0)/\ell! \to \sum_{\ell\ge\tau} c_\ell^2 g_\infty^{\star\ell}(0)/\ell!$, as t tends to infinity. Let us now check Assumption (iii). For this, it is enough to prove that $\lim_{t\to\infty}\limsup_{n_j\to\infty} \operatorname{Var}(n_j^{-1/2}\sum_{i=1}^{n_j}\sum_{\ell>t} c_\ell H_\ell(X_{j,i})/\ell!) = 0$. Note that $\operatorname{Var}(n_j^{-1/2}\sum_{i=1}^{n_j}\sum_{\ell>t} c_\ell H_\ell(X_{j,i})/\ell!) = \sum_{\ell>t}\sum_{|s|<n_j} c_\ell^2(1-|s|/n_j)\rho_j^\ell(s)/\ell!$. We aim at ap-

plying Lemma 12 to f_{n_j}, g_{n_j}, f and g defined hereafter. Let

$$f_{n_j}(\ell) = \frac{c_\ell^2}{\ell!} \sum_{|s|<n_j} \left(1 - \frac{|s|}{n_j}\right) \rho_j^\ell(s).$$

Observe that for $\ell \geq 2$, $|f_{n_j}(\ell)| \leq g_{n_j}(\ell)$ where

$$g_{n_j}(\ell) = \frac{c_\ell^2}{\ell!} \sum_{|s|<n_j} \left(1 - \frac{|s|}{n_j}\right) \rho_j^2(s).$$

By Lemma 9 we get, as $n_j \to \infty$, that

$$f_{n_j}(\ell) \to f(\ell) = 2\pi \frac{c_\ell^2}{\ell!} g_\infty^{\star\ell}(0) \quad \text{and} \quad g_{n_j}(\ell) \to g(\ell) = 2\pi \frac{c_\ell^2}{\ell!} g_\infty^{\star 2}(0).$$

Moreover, $\sum_{\ell>t} g_{n_j}(\ell) \to 2\pi \sum_{\ell>t} c_\ell^2 g_\infty^{\star 2}(0)/\ell!$ and Lemma 12 yields

$$\lim_{n_j\to\infty} \frac{1}{n_j} \operatorname{Var}\left(\sum_{i=1}^{n_j} \sum_{\ell>t} \frac{c_\ell}{\ell!} H_\ell(X_{j,i})\right) = 2\pi \sum_{\ell>t} \frac{c_\ell^2}{\ell!} g_\infty^{\star\ell}(0),$$

which tends to zero as t tends to infinity since $\sum_\ell c_\ell^2 g_\infty^{\star\ell}(0)/\ell!$ is a convergent series. □

Theorem 5. *Let $\{X_{j,i}, j \geq 1, i \geq 0\}$ be an array of standard stationary Gaussian processes such that, for a fixed $j \geq 1$, $(X_{j,i})_{i\geq 0}$ has a spectral density f_j and an autocorrelation function ρ_j defined by $\rho_j(k) = \mathbb{E}(X_{j,0}X_{j,k})$, for all $k \geq 0$. Let F_j be the c.d.f. of $X_{j,1}$ and F_{n_j} the empirical c.d.f., computed from $X_{j,1}, \ldots, X_{j,n_j}$. If $\{n_j\}_{j\geq 1}$ is a sequence of integers such that n_j tends to infinity as j tends to infinity and if condition (35) holds, then*

$$\sqrt{n_j}(F_{n_j} - F_j) \xrightarrow{d} W \qquad \text{in } D([-\infty, \infty]), \tag{49}$$

as j tends to infinity, where $D([-\infty, \infty])$ denotes the Skorokhod space on $[-\infty, \infty]$, and W is a Gaussian process with covariance function

$$\mathbb{E}[W(x)W(y)] = 2\pi \sum_{q\geq 1} \frac{J_q(x)J_q(y)}{q!} g_\infty^{\star q}(0), \qquad x, y \in \mathbb{R},$$

where $J_q(x) = \mathbb{E}[H_q(X)\{\mathbb{1}_{\{X\leq x\}} - \mathbb{E}(\mathbb{1}_{\{X\leq x\}})\}]$, H_q is the qth Hermite polynomial and X is a standard Gaussian random variable.

Proof. Let $S_j(x) = n_j^{-1/2} \sum_{i=1}^{n_j} (\mathbb{1}_{\{X_{j,i}\leq x\}} - F_j(x))$, for all x in $\mathbb{R}$. We shall first prove that for $x_1, \ldots, x_Q$ and $a_1, \ldots, a_Q$ in $\mathbb{R}$

$$\sum_{q=1}^{Q} a_q S_j(x_q) \xrightarrow{d} \mathcal{N}\left(0, 2\pi \sum_{\ell\geq 1} \frac{c_\ell^2}{\ell!} g_\infty^{\star\ell}(0)\right), \qquad \text{as } j \to \infty, \tag{50}$$

where c_ℓ is the ℓth Hermite coefficient of the function h defined by

$$h(\cdot) = \sum_{q=1}^{Q} a_q \big(\mathbb{1}_{\{\cdot \le x_q\}} - \mathbb{E}(\mathbb{1}_{\{\cdot \le x_q\}})\big).$$

Thus $\sum_{q=1}^{Q} a_q S_j(x_q) = n_j^{-1/2} \sum_{i=1}^{n_j} h(X_{j,i})$, where h is bounded and of Hermite rank $\tau \ge 1$ since, for all t in $\mathbb{R}$, $\mathbb{E}(X\mathbb{1}_{X\le t}) = \int_{\mathbb{R}} x\mathbb{1}_{x\le t}\varphi(x)\,\mathrm{d}x = \int_{-\infty}^{t}(-\varphi(x))'\,\mathrm{d}x = -\varphi(t) \ne 0$, and the CLT (50) follows from Theorem 4.

Let us now prove that there exists a positive constant C and $\beta > 1$ such that for all $r \le s \le t$,

$$\mathbb{E}\big(|S_j(s) - S_j(r)|^2 |S_j(t) - S_j(s)|^2\big) \le C|t-r|^{\beta}. \tag{51}$$

The convergence (49) then follows from (50), (51) and [5], Theorem 13.5. Note that

$$\begin{aligned}
&\mathbb{E}\big(|S_j(s) - S_j(r)|^2 |S_j(t) - S_j(s)|^2\big) \\
&\quad = \frac{1}{n_j^2} \sum_{i,i'=1}^{n_j} \sum_{l,l'=1}^{n_j} \mathbb{E}\big((k_s - k_r)(X_{j,i})(k_s - k_r)(X_{j,i'})(k_t - k_s)(X_{j,l})(k_t - k_s)(X_{j,l'})\big),
\end{aligned}$$

where $k_t(X) = \mathbb{1}_{\{X\le t\}} - \mathbb{E}(\mathbb{1}_{\{X\le t\}})$. By expanding each difference of functions in Hermite polynomials, we get

$$\begin{aligned}
&\mathbb{E}\big(|S_j(s) - S_j(r)|^2 |S_j(t) - S_j(s)|^2\big) \\
&\quad = \frac{1}{n_j^2} \sum_{i,i'=1}^{n_j} \sum_{l,l'=1}^{n_j} \sum_{p_1,\ldots,p_4 \ge 1} \frac{c_{p_1}(k_s - k_r) c_{p_2}(k_s - k_r) c_{p_3}(k_t - k_s) c_{p_4}(k_t - k_s)}{p_1! \cdots p_4!} \\
&\qquad \times \mathbb{E}(H_{p_1}(X_{j,i}) H_{p_2}(X_{j,i'}) H_{p_3}(X_{j,l}) H_{p_4}(X_{j,l'})).
\end{aligned}$$

Using the same arguments as in the case where m is even in the proof of Theorem 4, we obtain

$$\begin{aligned}
&\mathbb{E}\big(|S_j(s) - S_j(r)|^2 |S_j(t) - S_j(s)|^2\big) \\
&\quad = \frac{1}{n_j^2} \sum_{p_1,p_2 \ge 1} \sum_{i,i',l,l'=1}^{n_j} \bigg[\frac{c_{p_1}^2(k_t - k_s) c_{p_2}^2(k_s - k_r)}{p_1! p_2!} \rho_j^{p_1}(i'-i) \rho_j^{p_2}(l'-l) \\
&\qquad + \frac{c_{p_1}(k_t - k_s) c_{p_1}(k_s - k_r) c_{p_2}(k_t - k_s) c_{p_2}(k_s - k_r)}{p_1! p_2!} \\
&\qquad \times \rho_j^{p_1}(l-i) \rho_j^{p_2}(l'-i') \\
&\qquad + \frac{c_{p_1}(k_t - k_s) c_{p_1}(k_s - k_r) c_{p_2}(k_t - k_s) c_{p_2}(k_s - k_r)}{p_1! p_2!} \\
&\qquad \times \rho_j^{p_1}(l'-i) \rho_j^{p_2}(l-i')\bigg] + \mathrm{O}(n_j^{-1}).
\end{aligned}$$

Let $\|\cdot\|_2 = (\mathbb{E}(\cdot)^2)^{1/2}$ and $\langle f, g\rangle = \mathbb{E}[f(X)g(X)]$, where X is a standard Gaussian random variable. Since, by (64), $\sum_{i,i',l,l'=1}^{n_j} \rho_j^{p_1}(l-i)\rho_j^{p_2}(l'-i') = \mathrm{O}(n_j^2)$, we get with the Cauchy–Schwarz inequality that there exists a positive constant C such that

$$\begin{aligned}
&\mathbb{E}\big(|S_j(s)-S_j(r)|^2|S_j(t)-S_j(s)|^2\big)\\
&\quad\le C\sum_{p_1,p_2\ge 1}\left[\frac{c_{p_1}^2(k_t-k_s)c_{p_2}^2(k_s-k_r)}{p_1!p_2!}+\frac{c_{p_1}(k_t-k_s)c_{p_1}(k_s-k_r)c_{p_2}(k_t-k_s)c_{p_2}(k_s-k_r)}{p_1!p_2!}\right]\\
&\quad\le C(\|k_t-k_s\|_2^2\|k_s-k_r\|_2^2+|\langle k_t-k_s,k_s-k_r\rangle|^2)\\
&\quad\le C\|k_t-k_s\|_2^2\|k_s-k_r\|_2^2.
\end{aligned}$$

Note that $\|k_t-k_s\|_2^2 \le 2(\|\mathbb{1}_{\{X\le t\}}-\mathbb{1}_{\{X\le s\}}\|_2^2+\|\mathbb{E}(\mathbb{1}_{\{X\le s\}})-\mathbb{E}(\mathbb{1}_{\{X\le t\}})\|_2^2)$. Since $s\le t$, $\|\mathbb{1}_{\{X\le t\}}-\mathbb{1}_{\{X\le s\}}\|_2^2=\Phi(t)-\Phi(s)\le C|t-s|$, where Φ denotes the c.d.f. of a standard Gaussian random variable. Moreover, $\|\mathbb{E}(\mathbb{1}_{\{X\le s\}})-\mathbb{E}(\mathbb{1}_{\{X\le t\}})\|_2^2=|\Phi(t)-\Phi(s)|^2\le C|t-s|^2$, which concludes the proof of (51) with $\beta=2$. □

Proof of Proposition 1. For $*=\mathrm{CL}$, the proof of (22) is immediate, since

$$\sqrt{n_j}(\widehat{\sigma}_{\mathrm{CL},j}^2-\sigma_j^2)=\frac{1}{\sqrt{n_j}}\sum_{i=0}^{n_j-1}(W_{j,i}^2-\sigma_j^2)=\frac{2\sigma_j^2}{\sqrt{n_j}}\sum_{i=0}^{n_j-1}\mathrm{IF}\left(\frac{W_{j,i}}{\sigma_j},\mathrm{CL},\Phi\right).$$

Let us now prove (22) for $*=\mathrm{MAD}$. Let us denote by F_{n_j} the empirical c.d.f. of $W_{j,0:n_j-1}$ and by F_j the c.d.f. of $W_{j,0}$. Note that

$$\widehat{\sigma}_{\mathrm{MAD},j}=m(\Phi)T_0(F_{n_j}),$$

where $T_0=T_2\circ T_1$ with $T_1:F\mapsto\{r\mapsto\int_{\mathbb{R}}\mathbb{1}_{\{|x|\le r\}}\,\mathrm{d}F(x)\}$ and $T_2:U\mapsto U^{-1}(1/2)$. To prove (22), we start by proving that $\sqrt{n_j}(F_{n_j}-F_j)$ converges in distribution in the space of cadlag functions equipped with the topology of uniform convergence. This convergence follows by applying Theorem 5 to $X_{j,i}=W_{j,i}/\sigma_j$ which is an array of zero mean stationary Gaussian processes by [20], Corollary 1. The spectral density f_j of $(X_{j,i})_{i\ge 0}$ is given by $f_j(\lambda)=\mathbf{D}_{j,0}(\lambda;f)/\sigma_j^2$ where $\mathbf{D}_{j,0}(\cdot;f)$ is the within scale spectral density of the process $\{W_{j,k}\}_{k\ge 0}$, defined in (12), and σ_j^2 is the wavelet spectrum defined in (14). Here, $g_\infty(\lambda)=\mathbf{D}_{\infty,0}(\lambda;d)/\mathrm{K}(d)$, with $\mathbf{D}_{\infty,0}(\cdot;d)$ defined in (13) and $\mathrm{K}(d)=\int_{-\infty}^{+\infty}|\xi|^{-2d}|\widehat{\psi}(\xi)|^2\,\mathrm{d}\xi$ since, by [20], (26) and (29) in Theorem 1,

$$\left|\frac{\mathbf{D}_{j,0}(\lambda;f)}{f^*(0)\mathrm{K}(d)2^{2dj}}-\frac{\mathbf{D}_{\infty,0}(\lambda;d)}{\mathrm{K}(d)}\right|\le CL\mathrm{K}(d)^{-1}2^{-\beta j}\to 0,\qquad \text{as } j\to\infty,$$

$$\left|\frac{\sigma_j^2}{f^*(0)\mathrm{K}(d)2^{2dj}}-1\right|\le CL2^{-\beta j}\to 0,\qquad \text{as } j\to\infty.$$

Note also that, by [20], Theorem 1, $g_\infty(\lambda)$ is a continuous and 2π-periodic function on $(-\pi,\pi)$. Moreover, $g_\infty(\lambda)$ is bounded on $(-\pi,\pi)$ by Lemma 11, and (35) holds with

$$u_j = C_1 \frac{2^{-\beta j}}{\sigma_j^2/2^{2dj}}\left(2^{-\beta j} + C_2 \frac{\sigma_j^2}{2^{2dj}}\right) \to 0, \qquad \text{as } j \to \infty,$$

where C_1 and C_2 are positive constants. The asymptotic expansion (22) for $\widehat{\sigma}_{\mathrm{MAD},j}$ can be deduced from the functional Delta method (stated, e.g., in [32], Theorem 20.8) and the classical Delta method, stated, for example, in [32], Theorem 3.1. To show this, we have to prove that $T_0 = T_1 \circ T_2$ is Hadamard differentiable and that the corresponding Hadamard differential is defined and continuous on the whole space of cadlag functions. We prove first the Hadamard differentiability of the functional T_1. Let (g_t) be a sequence of cadlag functions with bounded variations such that $\|g_t - g\|_\infty \to 0$, as $t \to 0$, where g is a cadlag function. For any non-negative r, we consider

$$\begin{aligned}\frac{T_1(F_j + t g_t)[r] - T_1(F_j)[r]}{t} &= \frac{(F_j + t g_t)(r) - (F_j + t g_t)(-r) - F_j(r) + F_j(-r)}{t}\\ &= \frac{t g_t(r) - t g_t(-r)}{t} = g_t(r) - g_t(-r) \to g(r) - g(-r),\end{aligned}$$

since $\|g_t - g\|_\infty \to 0$, as $t \to 0$. The Hadamard differential of T_1 at g is given by

$$(DT_1(F_j).g)(r) = g(r) - g(-r).$$

By [32], Lemma 21.3, T_2 is Hadamard differentiable. Finally, using the Chain rule [32], Theorem 20.9, we obtain the Hadamard differentiability of T_0 with the following Hadamard differential:

$$DT_0(F_j).g = -\frac{(DT_1(F_j).g)(T_0(F_j))}{(T_1(F_j))'[T_0(F_j)]} = -\frac{g(T_0(F_j)) - g(-T_0(F_j))}{(T_1(F_j))'[T_0(F_j)]}.$$

In view of the last expression, $DT_0(F_j)$ is a continuous function of g and is defined on the whole space of cadlag functions. Thus by [32], Theorem 20.8, we obtain:

$$m(\Phi)\sqrt{n_j}\big(T_0(F_{n_j}) - T_0(F_j)\big) = m(\Phi) DT_0(F_j)\big\{\sqrt{n_j}(F_{n_j} - F_j)\big\} + \mathrm{o}_P(1),$$

where $m(\Phi)$ is the constant defined in (20). Since $T_0(F_j) = \sigma_j/m(\Phi)$ and $(T_1(F_j))'(r) = 2{\sigma_j}^{-1}\varphi(r/\sigma_j)$, where φ is the p.d.f. of a standard Gaussian random variable, we get

$$\sqrt{n_j}(\widehat{\sigma}_{\mathrm{MAD},j} - \sigma_j) = \frac{\sigma_j}{\sqrt{n_j}} \sum_{i=0}^{n_j - 1} \mathrm{IF}\left(\frac{W_{j,i}}{\sigma_j}, \mathrm{MAD}, \Phi\right) + \mathrm{o}_P(1)$$

and the expansion (22) for $* = \mathrm{MAD}$ follows from the classical Delta method, applied with $f(x) = x^2$. We end the proof of Proposition 1 by proving the asymptotic expansion (22) for $* = \mathrm{CR}$. We use the same arguments as those used previously. In this case the Hadamard differentiability comes from [16], Lemma 1. □

The following theorem is an extension of [2], Theorem 4, to arrays of stationary Gaussian processes in the multidimensional case.

Theorem 6. *Let $\underline{\mathbf{X}}_{J,i} = \{X_{J,i}^{(0)}, \dots, X_{J,i}^{(d)}\}$ be an array of standard stationary Gaussian processes such that for j, j' in $\{0, \dots, d\}$, the vector $\{X_{J,i}^{(j)}, X_{J,i}^{(j')}\}$ has a cross-spectral density $f_J^{(j,j')}$ and a cross-correlation function $\rho_J^{(j,j')}$ defined by $\rho_J^{(j,j')}(k) = \mathbb{E}(X_{J,i}^{(j)} X_{J,i+k}^{(j')})$, for all $k \geq 0$. Assume also that there exists a non-increasing sequence $\{u_J\}_{J\geq 1}$ such that u_J tends to zero as J tends to infinity, and, for all $J \geq 1$,*

$$\sup_{\lambda \in (-\pi,\pi)} \left| f_J^{(j,j')}(\lambda) - g_\infty^{(j,j')}(\lambda) \right| \leq u_J, \tag{52}$$

where $g_\infty^{(j,j')}$ is a 2π-periodic function which is bounded on $(-\pi, \pi)$ and continuous at the origin. Let h be a function on $\mathbb{R}$ such that $\mathbb{E}[h(X)^2] < \infty$, where X is a standard Gaussian random variable, and of Hermite rank $\tau \geq 1$. Let $\boldsymbol{\beta} = \{\beta_0, \dots, \beta_d\}$ in $\mathbb{R}^{d+1}$ and $\mathcal{H} : \mathbb{R}^{d+1} \to \mathbb{R}$, the real-valued function defined by $\mathcal{H}(\mathbf{x}) = \sum_{j=0}^{d} \beta_j h(x_j)$. Let $\{n_J\}_{J\geq 1}$ be a sequence of integers such that n_J tends to infinity as J tends to infinity. Then

$$\frac{1}{\sqrt{n_J}} \sum_{i=1}^{n_J} \mathcal{H}(\underline{\mathbf{X}}_{J,i}) \xrightarrow{d} \mathcal{N}(0, \widetilde{\sigma}^2), \qquad \text{as } J \to \infty, \tag{53}$$

where

$$\begin{aligned}
\widetilde{\sigma}^2 &= \lim_{n\to\infty} \operatorname{Var}\left(\frac{1}{\sqrt{n_J}} \sum_{i=1}^{n_J} \mathcal{H}(\underline{\mathbf{X}}_{J,i}) \right) \\
&= 2\pi \sum_{\ell \geq \tau} \frac{c_\ell^2}{\ell!} \sum_{0 \leq j,j' \leq d} \beta_j \beta_{j'} \bigl(g_\infty^{(j,j')}\bigr)^{\star \ell}(0).
\end{aligned}$$

In the previous equality, $c_\ell = \mathbb{E}[h(X) H_\ell(X)]$, where H_ℓ is the ℓth Hermite polynomial, and X is a standard Gaussian random variable.

The proof of Theorem 6 follows the same lines as the one of Theorem 4 and is thus omitted.

Proof of Theorem 2. Without loss of generality, we set $f^*(0) = 1$. In order to prove (26), let us first prove that, for $\boldsymbol{\alpha} = (\alpha_0, \dots, \alpha_\ell)$, where the α_is are in $\mathbb{R}$,

$$\begin{aligned}
&\sqrt{n 2^{-J_0}} 2^{-2J_0 d} \sum_{j=0}^{\ell} \alpha_j \bigl(\widehat{\sigma}_{*,J_0+j}^2 (W_{J_0+j,0:n_{J_0+j}-1}) - \sigma_{*,J_0+j}^2 \bigr) \\
&\quad \xrightarrow{d} \mathcal{N}(0, \boldsymbol{\alpha}^T \mathbf{U}_*(d) \boldsymbol{\alpha}).
\end{aligned} \tag{54}$$

By Proposition 1,

$$\begin{aligned}
&\sqrt{n2^{-J_0}}2^{-2J_0 d}\sum_{j=0}^{\ell}\alpha_j\big(\widehat{\sigma}^2_{*,J_0+j}(W_{J_0+j,0:n_{J_0+j}-1})-\sigma^2_{*,J_0+j}\big)\\
&\quad=\sum_{j=0}^{\ell}\frac{\sqrt{n2^{-J_0}}2^{-2J_0 d}}{n_{J_0+j}}2\alpha_j\sigma^2_{J_0+j}\sum_{i=0}^{n_{J_0+j}-1}\mathrm{IF}\bigg(\frac{W_{J_0+j,i}}{\sigma_{J_0+j}},*,\Phi\bigg)+\mathrm{o}_P(1).
\end{aligned}\tag{55}$$

Thus, proving (54) amounts to proving that

$$\frac{2^{-\ell/2}f^*(0)\mathrm{K}(d)}{\sqrt{n_{J_0+\ell}}}\sum_{j=0}^{\ell}2\alpha_j 2^{2dj+j}\sum_{i=0}^{n_{J_0+j}-1}\mathrm{IF}\bigg(\frac{W_{J_0+j,i}}{\sigma_{J_0+j}},*,\Phi\bigg)\xrightarrow{d}\mathcal{N}(0,\boldsymbol{\alpha}^T\mathbf{U}_*(d)\boldsymbol{\alpha}),\tag{56}$$

since $\sigma^2_{J_0+j}\sqrt{n2^{-J_0}}2^{-2J_0 d}/n_{J_0+j}\sim 2^{2dj-\ell/2+j}\mathrm{K}(d)f^*(0)/\sqrt{n_{J_0+\ell}}$, as n tends to infinity, by [20], (29) in Theorem 1. Note that

$$\begin{aligned}
\sum_{i=0}^{n_{J_0+j}-1}\mathrm{IF}\bigg(\frac{W_{J_0+j,i}}{\sigma_{J_0+j}},*,\Phi\bigg)&=\sum_{i=0}^{n_{J_0+\ell}-1}\sum_{v=0}^{2^{\ell-j}-1}\mathrm{IF}\bigg(\frac{W_{j+J_0,2^{\ell-j}i+v}}{\sigma_{J_0+j}},*,\Phi\bigg)\\
&\quad+\sum_{q=n_{J_0+j}-(\mathrm{T}-1)(2^{\ell-j}-1)}^{n_{J_0+j}-1}\mathrm{IF}\bigg(\frac{W_{j+J_0,q}}{\sigma_{J_0+j}},*,\Phi\bigg).
\end{aligned}$$

Using the notation $\beta_j=2\alpha_j 2^{2dj-\ell/2+j}\mathrm{K}(d)f^*(0)$ and that IF is either bounded or equal to $H_2/2$,

$$\begin{aligned}
&\frac{1}{\sqrt{n_{J_0+\ell}}}\sum_{j=0}^{\ell}\beta_j\sum_{i=0}^{n_{J_0+j}-1}\mathrm{IF}\bigg(\frac{W_{J_0+j,i}}{\sigma_{J_0+j}},*,\Phi\bigg)\\
&\quad=\frac{1}{\sqrt{n_{J_0+\ell}}}\sum_{j=0}^{\ell}\beta_j\sum_{i=0}^{n_{J_0+\ell}-1}\sum_{v=0}^{2^{\ell-j}-1}\mathrm{IF}\bigg(\frac{W_{j+J_0,2^{\ell-j}i+v}}{\sigma_{J_0+j}},*,\Phi\bigg)+\mathrm{o}_P(1)\\
&\quad=\frac{1}{\sqrt{n_{J_0+\ell}}}\sum_{i=0}^{n_{J_0+\ell}-1}\mathbf{F}(Y_{J_0,\ell,i},*)+\mathrm{o}_P(1),
\end{aligned}$$

where

$$\mathbf{F}(Y_{J_0,\ell,i},*)=\sum_{j=0}^{\ell}\beta_j\sum_{v=0}^{2^{\ell-j}-1}\mathrm{IF}\bigg(\frac{W_{j+J_0,2^{\ell-j}i+v}}{\sigma_{J_0+j}},*,\Phi\bigg)$$

and

$$Y_{J_0,\ell,i} = \left(\frac{W_{J_0+\ell,i}}{\sigma_{J_0+\ell}}, \frac{W_{J_0+\ell-1,2i}}{\sigma_{J_0+\ell-1}}, \frac{W_{J_0+\ell-1,2i+1}}{\sigma_{J_0+\ell-1}}, \dots, \frac{W_{J_0+j,2^{\ell-j}i}}{\sigma_{J_0+j}}, \dots, \frac{W_{J_0+j,2^{\ell-j}i+2^{\ell-j}-1}}{\sigma_{J_0+j}}, \dots, \frac{W_{J_0,2^{\ell}i}}{\sigma_{J_0}}, \dots, \frac{W_{J_0,2^{\ell}i+2^{\ell}-1}}{\sigma_{J_0}}\right)^T$$

is a $(2^{\ell+1}-1)$-dimensional stationary Gaussian vector. By Lemma 7, $\mathbf{F}$ is of Hermite rank larger than 2. Hence, from Theorem 6 applied to $\mathcal{H}(\cdot) = \mathbf{F}(\cdot)$, $\underline{\mathbf{X}}_{J,i} = Y_{J_0,\ell,i}$ and $h(\cdot) = \mathrm{IF}(\cdot)$, we get

$$\frac{1}{\sqrt{n_{J_0+\ell}}} \sum_{i=0}^{n_{J_0+\ell}-1} \mathbf{F}(Y_{J_0,\ell,i}, *) \xrightarrow{d} \mathcal{N}(0, \widetilde{\sigma}_*^2), \tag{57}$$

where $\widetilde{\sigma}_*^2 = \lim_{n\to\infty} n_{J_0+\ell}^{-1} \operatorname{Var}(\sum_{i=0}^{n_{J_0+\ell}-1} \mathbf{F}(Y_{J_0,\ell,i}, *))$. By [20], (26) and (29), and by using the same arguments as those used in the proof of Proposition 1, condition (52) of Theorem 6 holds with $f_J^{(j,j')}(\lambda) = \mathbf{D}^{(r)}_{J_0+j,j-j'}(\lambda; f)/\sigma_{J_0+j}\sigma_{J_0+j'}$ and $g_\infty^{(j,j')} = \mathbf{D}^{(r)}_{\infty,j-j'}(\lambda; d)/\mathrm{K}(d)$, where $0 \le r \le 2^{j-j'} - 1$ and $\mathbf{D}_{J_0+j,j-j'}(\cdot; f)$ is the cross-spectral density of the stationary between scale process defined in (12). Lemma 11 and [20], Theorem 1, ensure that $\mathbf{D}^{(r)}_{\infty,j-j'}(\cdot; d)$ is a bounded, continuous and 2π-periodic function.

By using Mehler's formula [6], equation (2.1), and the expansion of IF onto the Hermite polymials basis given by: $\mathrm{IF}(x, *, \Phi) = \sum_{p\ge 2} c_p(\mathrm{IF}_*) H_p(x)/p!$, where $c_p(\mathrm{IF}_*) = \mathbb{E}[\mathrm{IF}(X, *, \Phi) H_p(X)]$, H_p being the pth Hermite polynomial, we get

$$\begin{aligned}
&\frac{1}{n_{J_0+\ell}} \operatorname{Var}\left(\sum_{i=0}^{n_{J_0+\ell}-1} \mathbf{F}(Y_{J_0,\ell,i}, *)\right) \\
&= \frac{1}{n_{J_0+\ell}} \sum_{j,j'=1}^{\ell} \beta_j \beta_{j'} \sum_{i,i'=0}^{n_{J_0+\ell}-1} \sum_{v=0}^{2^{\ell-j}-1} \sum_{v'=0}^{2^{\ell-j'}-1} \mathbb{E}\left[\mathrm{IF}\left(\frac{W_{J_0+j,2^{\ell-j}i+v}}{\sigma_{J_0+j}}, *, \Phi\right) \times \mathrm{IF}\left(\frac{W_{J_0+j',2^{\ell-j'}i'+v'}}{\sigma_{J_0+j'}}, *, \Phi\right)\right] \\
&= \frac{1}{n_{J_0+\ell}} \sum_{j,j'=1}^{\ell} \beta_j \beta_{j'} \sum_{i=0}^{n_{J_0+j}-1} \sum_{i'=0}^{n_{J_0+j'}-1} \mathbb{E}\left[\mathrm{IF}\left(\frac{W_{J_0+j,i}}{\sigma_{J_0+j}}, *, \Phi\right) \times \mathrm{IF}\left(\frac{W_{J_0+j',i'}}{\sigma_{J_0+j'}}, *, \Phi\right)\right] + \mathrm{o}(1) \\
&= \frac{1}{n_{J_0+\ell}} \sum_{j,j'=1}^{\ell} \beta_j \beta_{j'} \sum_{i=0}^{n_{J_0+j}-1} \sum_{i'=0}^{n_{J_0+j'}-1} \sum_{p\ge 2} \frac{c_p^2(\mathrm{IF}_*)}{p!} \left(\mathbb{E}\left[\frac{W_{J_0+j,i}}{\sigma_{J_0+j}} \frac{W_{J_0+j',i'}}{\sigma_{J_0+j'}}\right]\right)^p + \mathrm{o}(1).
\end{aligned} \tag{58}$$

Without loss of generality, we shall assume in the sequel that $j \geq j'$. Equation (58) can be rewritten as follows by using that $i' = 2^{j-j'}q + r$, where $q \in \mathbb{N}$ and $r \in \{0, 1, \ldots, 2^{j-j'} - 1\}$ and equation (18) in [20]:

$$
\begin{aligned}
&\frac{1}{n_{J_0+\ell}} \sum_{j,j'=1}^{\ell} \beta_j \beta_{j'} \sum_{i=0}^{n_{J_0+j}-1} \sum_{q=0}^{n_{J_0+j}-1} \sum_{r=0}^{2^{j-j'}-1} \sum_{p \geq 2} \frac{c_p^2(\mathrm{IF}_*)}{p!} \\
&\qquad \times \left(\mathbb{E}\left[\frac{W_{J_0+j,0}}{\sigma_{J_0+j}} \frac{W_{J_0+j',2^{j-j'}(q-i)+r}}{\sigma_{J_0+j'}} \right] \right)^p + \mathrm{o}(1) \\
&= \frac{n_{J_0+j}}{n_{J_0+\ell}} \sum_{j,j'=1}^{\ell} \beta_j \beta_{j'} \sum_{|\tau| < n_{J_0+j}} \sum_{r=0}^{2^{j-j'}-1} \sum_{p \geq 2} \frac{c_p^2(\mathrm{IF}_*)}{p!} \left(1 - \frac{|\tau|}{n_{J_0+j}} \right) \\
&\qquad \times \left(\mathbb{E}\left[\frac{W_{J_0+j,0}}{\sigma_{J_0+j}} \frac{W_{J_0+j',2^{j-j'}\tau+r}}{\sigma_{J_0+j'}} \right] \right)^p + \mathrm{o}(1) \\
&= \frac{n_{J_0+j}}{n_{J_0+\ell}} \sum_{j,j'=1}^{\ell} \beta_j \beta_{j'} \sum_{|\tau| < n_{J_0+j}} \sum_{r=0}^{2^{j-j'}-1} \sum_{p \geq 2} \frac{c_p^2(\mathrm{IF}_*)}{p!} \left(1 - \frac{|\tau|}{n_{J_0+j}} \right) \\
&\qquad \times \left(\int_{-\pi}^{\pi} \frac{\mathbf{D}^{(r)}_{J_0+j,j-j'}(\lambda; f) \mathrm{e}^{\mathrm{i}\lambda\tau}}{\sigma_{J_0+j}\sigma_{J_0+j'}} \, \mathrm{d}\lambda \right)^p + \mathrm{o}(1),
\end{aligned}
$$

where $\mathbf{D}_{J_0+j,j-j'}(\cdot; f)$ is the cross-spectral density of the stationary between scale process defined in (12). We aim at applying Lemma 12 with f_n, g_n, f and g defined hereafter.

$$
f_{n_{J_0+j}}(\tau, p) = \frac{c_p^2(\mathrm{IF}_*)}{p!} \sum_{r=0}^{2^{j-j'}-1} \mathbb{1}\{|\tau| < n_{J_0+j}\} \left(1 - \frac{|\tau|}{n_{J_0+j}} \right) \left(\mathbb{E}\left[\frac{W_{J_0+j,0}}{\sigma_{J_0+j}} \frac{W_{J_0+j',2^{j-j'}\tau+r}}{\sigma_{J_0+j'}} \right] \right)^p.
$$

Observe that $|f_{n_{J_0+j}}| \leq g_{n_{J_0+j}}$, where

$$
g_{n_{J_0+j}}(\tau, p) = \frac{c_p^2(\mathrm{IF}_*)}{p!} \sum_{r=0}^{2^{j-j'}-1} \mathbb{1}\{|\tau| < n_{J_0+j}\} \left(1 - \frac{|\tau|}{n_{J_0+j}} \right) \left(\mathbb{E}\left[\frac{W_{J_0+j,0}}{\sigma_{J_0+j}} \frac{W_{J_0+j',2^{j-j'}\tau+r}}{\sigma_{J_0+j'}} \right] \right)^2.
$$

Using [20], (26) and (29) in Theorem 1, we get that

$$
\lim_{n \to \infty} \frac{\mathbf{D}_{J_0+j,j-j'}(\lambda; f)}{\sigma_{J_0+j}\sigma_{J_0+j'}} = \frac{2^{d(j-j')}}{\mathrm{K}(d)} \mathbf{D}_{\infty,j-j'}(\lambda; d).
$$

This implies that $\lim_{n\to\infty} f_{n_{J_0+j}}(\tau, p) = f(\tau, p)$ where

$$f(\tau, p) = \frac{c_p^2(\mathrm{IF}_*)}{p!} \sum_{r=0}^{2^{j-j'}-1} \left(\frac{2^{d(j-j')}}{\mathrm{K}(d)} \int_{-\pi}^{\pi} \mathbf{D}^{(r)}_{\infty, j-j'}(\lambda; d) \mathrm{e}^{\mathrm{i}\lambda\tau} \, \mathrm{d}\lambda \right)^p .$$

Furthermore, $\lim_{n\to\infty} g_{n_{J_0+j}}(\tau, p) = g(\tau, p)$ where

$$g(\tau, p) = \frac{c_p^2(\mathrm{IF}_*)}{p!} \frac{2^{2d(j-j')}}{\mathrm{K}(d)^2} \left| \int_{-\pi}^{\pi} \mathbf{D}_{\infty, j-j'}(\lambda; d) \mathrm{e}^{\mathrm{i}\lambda\tau} \, \mathrm{d}\lambda \right|_2^2,$$

and $|\mathbf{x}|_2^2 = \sum_{k=1}^r x_k^2$ for $\mathbf{x} = (x_1, \ldots, x_r) \in \mathbb{R}^r$. Using (63)–(65) in [20], we get

$$\sum_{p\geq 2} \sum_{\tau\in\mathbb{Z}} g_{n_{J_0+j}}(\tau, p) \longrightarrow \left(\sum_{p\geq 2} \frac{c_p^2(\mathrm{IF}_*)}{p!} \right) \frac{2^{2d(j-j')}}{\mathrm{K}(d)^2} 2\pi \int_{-\pi}^{\pi} |\mathbf{D}_{\infty, j-j'}(\lambda; d)|_2^2 \, \mathrm{d}\lambda, \qquad \text{as } n \to \infty.$$

Then, with Lemma 12, we obtain

$$\widetilde{\sigma}_*^2 = \sum_{p\geq 2} \frac{c_p^2(\mathrm{IF}_*)(f^*(0))^2}{p!\mathrm{K}(d)^{p-2}} \sum_{j,j'=0}^{\ell} 4\alpha_j \alpha_{j'} 2^{dj(2+p)} 2^{dj'(2-p)+j'} \times \sum_{\tau\in\mathbb{Z}} \sum_{r=0}^{2^{j-j'}-1} \left(\int_{-\pi}^{\pi} \mathbf{D}^{(r)}_{\infty, j-j'}(\lambda; d) \mathrm{e}^{\mathrm{i}\lambda\tau} \, \mathrm{d}\lambda \right)^p . \qquad \square$$

7. Technical lemmas

Lemma 7. *Let X be a standard Gaussian random variable. The influence functions* IF *defined in Proposition 1 have the following properties*:

$$\mathbb{E}[\mathrm{IF}(X, *, \Phi)] = 0; \tag{59}$$

$$\mathbb{E}[X\mathrm{IF}(X, *, \Phi)] = 0; \tag{60}$$

$$\mathbb{E}[X^2\mathrm{IF}(X, *, \Phi)] \neq 0. \tag{61}$$

Proof. We only have to prove the result for $* = \mathrm{MAD}$ since the result for $* = \mathrm{CR}$ follows from [16], Lemma 12. (59) comes from $\mathbb{E}(\mathbb{1}_{\{X\leq 1/m(\Phi)\}}) = \mathbb{E}(\mathbb{1}_{\{X\leq\Phi^{-1}(3/4)\}}) = 3/4$ and $\mathbb{E}(\mathbb{1}_{\{X\leq -1/m(\Phi)\}}) = 1/4$, where X is a standard Gaussian random variable. (60) follows from $\int_{\mathbb{R}} x\mathbb{1}_{\{x\leq\Phi^{-1}(3/4)\}}\varphi(x)\,\mathrm{d}x - \int_{\mathbb{R}} x\mathbb{1}_{\{x\leq-\Phi^{-1}(3/4)\}}\varphi(x)\,\mathrm{d}x = -\varphi(\Phi^{-1}(3/4)) + \varphi(-\Phi^{-1}(3/4)) = 0$, where φ is the p.d.f. of a standard Gaussian random variable and the fact that $\mathbb{E}(X) = 0$. Let us now compute $\mathbb{E}[X^2\mathrm{IF}(X, \mathrm{MAD}, \Phi)]$. Integrating by parts, we get $\int_{\mathbb{R}} x^2\mathbb{1}_{\{x\leq\Phi^{-1}(3/4)\}}\varphi(x)\,\mathrm{d}x - 3/4 - \int_{\mathbb{R}} x^2\mathbb{1}_{\{x\leq-\Phi^{-1}(3/4)\}}\varphi(x)\,\mathrm{d}x + 1/4 = -2\varphi(\Phi^{-1}(3/4))$. Thus, $\mathbb{E}[X^2\mathrm{IF}(X, \mathrm{MAD}, \Phi)] = 2 \neq 0$, which concludes the proof. $\square$

Lemma 8. *Let X be a standard Gaussian random variable. The influence functions* IF, *defined in Lemma 1, have the following properties*:

$$\mathbb{E}[\mathrm{IF}^2(X, \mathrm{MAD}, \Phi)] = \frac{m^2(\Phi)}{16\varphi(\Phi^{-1}(3/4)^2)} = 1.3601, \tag{62}$$

$$\mathbb{E}[\mathrm{IF}^2(X, \mathrm{CR}, \Phi)] \approx 0.6077. \tag{63}$$

Proof. Equation (63) comes from [27]. Since

$$\mathbb{E}[\mathrm{IF}^2(X, \mathrm{MAD}, \Phi)] = \frac{m^2(\Phi)}{4\varphi(\Phi^{-1}(3/4)^2)} \operatorname{Var}\bigl(\mathbb{1}_{\{|X|\le\Phi^{-1}(3/4)\}}\bigr),$$

where $\mathbb{1}_{\{|X|\le\Phi^{-1}(3/4)\}}$ is a Bernoulli random variable with parameter $1/2$, (62) follows. □

Lemma 9. *Under the assumptions of Theorem 4 and for any fixed $\ell \ge 1$*,

$$\frac{1}{n_j}\sum_{r,s=1}^{n_j} \rho_j^{\ell}(r-s) \to 2\pi g_{\infty}^{\star\ell}(0), \qquad \textit{as } n_j \to \infty. \tag{64}$$

Proof. Let us first prove that

$$\frac{1}{n_j}\sum_{1\le r,s\le n_j} \rho_j(r-s) \to 2\pi g_{\infty}(0), \qquad \text{as } n_j \to \infty. \tag{65}$$

Using that $\mathcal{F}_{n_j}$, defined by $\mathcal{F}_{n_j}(\lambda) = n_j^{-1}|\sum_{r=1}^{n_j} \mathrm{e}^{\mathrm{i}\lambda r}|^2$, for all λ in $[-\pi, \pi]$, satisfies $\int_{-\pi}^{\pi} \mathcal{F}_{n_j}(\lambda)\,\mathrm{d}\lambda = 2\pi$, we obtain

$$\begin{aligned}\frac{1}{n_j}\Biggl(\sum_{1\le r,s\le n_j} \rho_j(r-s)\Biggr) - 2\pi g_{\infty}(0) &= \int_{-\pi}^{\pi} \bigl(f_j(\lambda) - g_{\infty}(\lambda)\bigr)\mathcal{F}_{n_j}(\lambda)\,\mathrm{d}\lambda \\ &\quad + \int_{-\pi}^{\pi} \bigl(g_{\infty}(\lambda) - g_{\infty}(0)\bigr)\mathcal{F}_{n_j}(\lambda)\,\mathrm{d}\lambda.\end{aligned} \tag{66}$$

Using that $\mathcal{F}_{n_j}$ is non-negative, $\int_{-\pi}^{\pi} \mathcal{F}_{n_j}(\lambda)\,\mathrm{d}\lambda = 2\pi$ and (35), the first term on the right-hand side of (66) tends to zero as j tends to infinity. The second term on the right-hand side of (66) can be bounded above as follows. For $0 < \eta \le \pi$,

$$\begin{aligned}&\left|\int_{-\pi}^{\pi} \bigl(g_{\infty}(\lambda) - g_{\infty}(0)\bigr)\mathcal{F}_{n_j}(\lambda)\,\mathrm{d}\lambda\right| \\ &\quad\le \int_{-\pi}^{-\eta} |g_{\infty}(\lambda) - g_{\infty}(0)|\mathcal{F}_{n_j}(\lambda)\,\mathrm{d}\lambda \\ &\qquad + \int_{-\eta}^{\eta} |g_{\infty}(\lambda) - g_{\infty}(0)|\mathcal{F}_{n_j}(\lambda)\,\mathrm{d}\lambda + \int_{\eta}^{\pi} |g_{\infty}(\lambda) - g_{\infty}(0)|\mathcal{F}_{n_j}(\lambda)\,\mathrm{d}\lambda.\end{aligned} \tag{67}$$

Since there exists a positive constant C such that $\mathcal{F}_{n_j}(\lambda) \le C/(n_j|\lambda|^2)$, for all λ in $[-\pi, \pi]$, the first and last terms on the right-hand side of (67) are bounded by $C\pi/(n_j\eta^2)$. The continuity of g_∞ at 0 and the fact that $\int_{-\eta}^{\eta} \mathcal{F}_{n_j}(\lambda)\,\mathrm{d}\lambda \le \int_{-\pi}^{\pi} \mathcal{F}_{n_j}(\lambda)\,\mathrm{d}\lambda = 2\pi$ ensure that the right-hand side of (67) tends to zero as j tends to infinity. This completes the proof of (65).

Observe that, for all k in $\mathbb{Z}$, $\rho_j^{\ell}(k) = \int_{-\pi}^{\pi} \mathrm{e}^{\mathrm{i}\lambda k} f_j^{\star\ell}(\lambda)\,\mathrm{d}\lambda$, where $f_j^{\star\ell}$ is the ℓth self-convolution of f_j. Since (35) implies that $\sup_{\lambda\in(-\pi,\pi)} |f_j^{\star\ell}(\lambda) - g_\infty^{\star\ell}(\lambda)|$ tends to zero as j tends to infinity for any fixed $\ell \ge 2$, the same arguments as those used to prove (65) lead to (64). $\square$

Lemma 10. *Under the assumptions of Theorem 4, let $\mathcal{D}_{n_j}$ and $\mathcal{F}_{n_j}$ be defined by $\mathcal{D}_{n_j}(\lambda) = \sum_{r=1}^{n_j} \mathrm{e}^{\mathrm{i}\lambda r}$ and $\mathcal{F}_{n_j}(\lambda) = n_j^{-1}|\sum_{r=1}^{n_j} \mathrm{e}^{\mathrm{i}\lambda r}|^2$, respectively, for all λ in $[-\pi, \pi]$. Then the following two statements hold true*:

(i) *For any fixed $j, \ell \ge 1$,*

$$\|\mathcal{F}_{n_j} \star f_j^{\star\ell}\|_\infty \le 2\pi(\|f_j^{\star\ell} - g_\infty^{\star\ell}\|_\infty + \|g_\infty^{\star\ell}\|_\infty), \tag{68}$$

where $\|\mathcal{F}_{n_j} \star f_j^{\star\ell}\|_\infty = \sup_{t\in\mathbb{R}} |\int_{-\pi}^{\pi} \mathcal{F}_{n_j}(t-\lambda) f_j^{\star\ell}(\lambda)\,\mathrm{d}\lambda|$.

(ii) *For any fixed $\ell \ge 1$,*

$$\sup_{x\in\mathbb{R}} \int_{-\pi}^{\pi} \frac{|\mathcal{D}_{n_j}(\lambda+x)|}{\sqrt{n_j}} f_j^{\star\ell}(\lambda)\,\mathrm{d}\lambda \to 0, \qquad \text{as } n_j \to \infty. \tag{69}$$

Proof. (i) Using that $\mathcal{F}_{n_j}$ is non-negative and such that $\int_{-\pi}^{\pi} \mathcal{F}_{n_j}(\lambda)\,\mathrm{d}\lambda = 2\pi$, we get that $\|\mathcal{F}_{n_j} \star f_j^{\star\ell}\|_\infty \le 2\pi\|f_j^{\star\ell}\|_\infty$, and thus (68) follows from the triangle inequality.

(ii) Writing $f_j^{\star\ell}(\lambda) = (f_j^{\star\ell}(\lambda) - g_\infty^{\star\ell}(\lambda)) + g_\infty^{\star\ell}(\lambda)$ and using the Cauchy–Schwarz inequality,

$$\begin{aligned}\sup_{x\in\mathbb{R}} \int_{-\pi}^{\pi} \frac{|\mathcal{D}_{n_j}(\lambda+x)|}{\sqrt{n_j}} f_j^{\star\ell}(\lambda)\,\mathrm{d}\lambda &\le \sqrt{2\pi}\|f_j^{\star\ell} - g_\infty^{\star\ell}\|_\infty \sup_{x\in\mathbb{R}} \left(\int_{-\pi}^{\pi} \mathcal{F}_{n_j}(\lambda+x)\,\mathrm{d}\lambda\right)^{1/2} \\ &\quad + \sup_{x\in\mathbb{R}} \int_{-\pi}^{\pi} \frac{|\mathcal{D}_{n_j}(\lambda+x)|}{\sqrt{n_j}} g_\infty^{\star\ell}(\lambda)\,\mathrm{d}\lambda.\end{aligned} \tag{70}$$

By (35), $\|f_j^{\star\ell} - g_\infty^{\star\ell}\|_\infty$ tends to zero for any fixed $\ell \ge 1$, as n_j tends to infinity. Since $\mathcal{F}_{n_j}$ is a 2π-periodic function, $\int_{-\pi}^{\pi} \mathcal{F}_{n_j}(\lambda+x)\,\mathrm{d}\lambda = 2\pi$ for all x in $\mathbb{R}$, and thus the first term in the right-hand side of (70) tends to zero as n_j tends to infinity.

Let us now study the second term on the right-hand side of (70). Since $\mathcal{D}_{n_j}$ and g_∞ are 2π-periodic functions, $n_j^{-1/2} \int_{-\pi}^{\pi} |\mathcal{D}_{n_j}(\lambda+x)| g_\infty^{\star\ell}(\lambda)\,\mathrm{d}\lambda = n_j^{-1/2} \int_{-\pi}^{\pi} |\mathcal{D}_{n_j}(u)| g_\infty^{\star\ell}(u-x)\,\mathrm{d}u$, for all x in $\mathbb{R}$. Then, by splitting the interval $[-\pi, \pi]$ into $[-\delta, \delta]$ and $[-\pi, \pi] \setminus [-\delta, \delta]$ and by using the Cauchy–Schwarz inequality, we get that, for all x in $\mathbb{R}$,

$$
\begin{aligned}
&\int_{-\pi}^{\pi} \frac{|\mathcal{D}_{n_j}(\lambda+x)|}{\sqrt{n_j}} g_\infty^{\star\ell}(\lambda)\,\mathrm{d}\lambda \\
&\quad \le \left(\int_{-\delta}^{\delta} \mathcal{F}_{n_j}(u)\,\mathrm{d}u\right)^{1/2} \left(\int_{-\delta}^{\delta} g_\infty^{\star\ell}(u-x)^2\,\mathrm{d}u\right)^{1/2} \\
&\qquad + \left(\int_{[-\pi,\pi]\setminus[-\delta,\delta]} \mathcal{F}_{n_j}(u)\,\mathrm{d}u\right)^{1/2} \left(\int_{[-\pi,\pi]\setminus[-\delta,\delta]} g_\infty^{\star\ell}(u-x)^2\,\mathrm{d}u\right)^{1/2}.
\end{aligned}
$$

Using that $\int_{-\delta}^{\delta} \mathcal{F}_{n_j}(u)\,\mathrm{d}u$ and $\int_{[-\pi,\pi]\setminus[-\delta,\delta]} \mathcal{F}_{n_j}(u)\,\mathrm{d}u$ are bounded above by $\int_{[-\pi,\pi]} \mathcal{F}_{n_j}(u)\,\mathrm{d}u = 2\pi$ and that g_∞ is bounded, we get that there exists a positive constant $C(\ell)$ depending on ℓ such that

$$
\sup_{x\in\mathbb{R}} \int_{-\pi}^{\pi} \frac{|\mathcal{D}_{n_j}(\lambda+x)|}{\sqrt{n_j}} g_\infty^{\star\ell}(\lambda)\,\mathrm{d}\lambda \le C(\ell)\|g_\infty\|_\infty^{\ell}\left(\sqrt{\delta}+\frac{1}{\sqrt{n_j}\delta}\right).
$$

Setting $\delta = \delta_{n_j} = n_j^{\alpha-1/2}$, with α in $(0, 1/2)$, (69) follows. □

Lemma 11. *Let* $\mathbf{e}_u(\xi) = 2^{-u/2}[1, \mathrm{e}^{-\mathrm{i}2^{-u}\xi}, \ldots, \mathrm{e}^{-\mathrm{i}(2^u-1)2^{-u}\xi}]^T$, *where* $\xi \in \mathbb{R}$. *For all* $u \ge 0$, *each component of the vector*

$$
\mathbf{D}_{\infty,u}(\lambda; d) = \sum_{l\in\mathbb{Z}} |\lambda+2l\pi|^{-2d}\mathbf{e}_u(\lambda+2l\pi)\overline{\widehat{\psi}(\lambda+2l\pi)}\widehat{\psi}\big(2^{-u}(\lambda+2l\pi)\big),
$$

is bounded on $(-\pi, \pi)$, *where* $\widehat{\psi}$ *is defined in* (4).

Proof. We start with the case where $l = 0$. Using (5), we obtain that $2^{-u/2}|\lambda|^{-2d}|\widehat{\psi}(\lambda)| \times |\widehat{\psi}(2^{-u}\lambda)| = \mathrm{O}(|\lambda|^{2M-2d})$, as $\lambda \to 0$; hence, (7) ensures that $2^{-u/2}|\lambda|^{-2d}|\widehat{\psi}(\lambda)||\widehat{\psi}(2^{-u}\lambda)| = \mathrm{O}(1)$. Let $\mathbf{e}_u^{(k)}$ denotes the kth component of the vector $\mathbf{e}_u$. For $l \neq 0$, (W-2) ensures that, for all λ in $(-\pi, \pi)$, there exists a positive constant C such that $|\widehat{\psi}(\lambda)| \le C/(1+|\lambda|)^\alpha$. Then there exists a positive constant C' such that

$$
\sum_{l\in\mathbb{Z}^*} |\lambda+2\pi l|^{-2d}\overline{\widehat{\psi}(\lambda+2\pi l)}\widehat{\psi}\big(2^{-u}(\lambda+2\pi l)\big)\mathbf{e}_u^{(k)}(\lambda) \le C' \sum_{l\in\mathbb{Z}^*} |\lambda+2\pi l|^{-2d-2\alpha}.
$$

If $\lambda = 0$, $\sum_{l\in\mathbb{Z}^*} 1/|2\pi l|^{2d+2\alpha} < \infty$ by (7). If $\lambda \neq 0$, then, since $-\pi \le \lambda \le \pi$, $\sum_{l\in\mathbb{Z}^*} 1/|\lambda+2\pi l|^{2d+2\alpha} \le \sum_{l\in\mathbb{Z}^*} 1/|\pi(2l-1)|^{2d+2\alpha} < \infty$ by (7). □

Lemma 12. *Let* f_n *and* g_n *be two sequences of measurable functions on a measure space* $(\Omega, \mathcal{F}, \mu)$ *such that, for all* n, $|f_n| \le g_n$. *Assume that* $\lim_{n\to\infty} g_n = g$, $\lim_{n\to\infty}\int g_n\,\mathrm{d}\mu = \int g\,\mathrm{d}\mu$ *and that* $\lim_{n\to\infty} f_n = f$. *Then* $\int \lim_{n\to\infty} f_n\,\mathrm{d}\mu = \lim_{n\to\infty}\int f_n\,\mathrm{d}\mu$.

Proof. Since $|f_n| \le g_n$, we have $-g_n \le f_n \le g_n$. By applying Fatou's Lemma first to $g_n - f_n$ and after to $f_n + g_n$, we get $\liminf_{n\to\infty}\int(g_n - f_n) \ge \int g - \int f$ and $\liminf_{n\to\infty}\int(g_n + f_n) \ge$

$\int g + \int f$. Since $\lim_{n\to\infty} \int g_n = \int g$, $\liminf_{n\to\infty} \int (g_n - f_n) \le \int g + \liminf_{n\to\infty}(-\int f_n)$, and thus we deduce from the first inequality that $\limsup_{n\to\infty} \int f_n \le \int f$. In the same way, we deduce from the second inequality that $\int f \le \liminf_{n\to\infty} \int f_n$, which concludes the proof. □

References

[1] Abry, P. and Veitch, D. (1998). Wavelet analysis of long-range-dependent traffic. *IEEE Trans. Inform. Theory* **44** 2–15. MR1486645

[2] Arcones, M.A. (1994). Limit theorems for nonlinear functionals of a stationary Gaussian sequence of vectors. *Ann. Probab.* **22** 2242–2274. MR1331224

[3] Beran, J. (1992). Statistical methods for data with long-range dependence. *Statist. Sci.* **7** 404–427.

[4] Beran, J. (1994). *Statistics for Long-Memory Processes. Monographs on Statistics and Applied Probability* **61**. New York: Chapman & Hall. MR1304490

[5] Billingsley, P. (1999). *Convergence of Probability Measures*, 2nd ed. *Wiley Series in Probability and Statistics: Probability and Statistics*. New York: Wiley. MR1700749

[6] Breuer, P. and Major, P. (1983). Central limit theorems for nonlinear functionals of Gaussian fields. *J. Multivariate Anal.* **13** 425–441. MR0716933

[7] Brockwell, P.J. and Davis, R.A. (1991). *Time Series: Theory and Methods*, 2nd ed. *Springer Series in Statistics*. New York: Springer. MR1093459

[8] Chareka, P., Matarise, F. and Turner, R. (2006). A test for additive outliers applicable to long-memory time series. *J. Econom. Dynam. Control* **30** 595–621. MR2222294

[9] Coeurjolly, J.F. (2008). Bahadur representation of sample quantiles for functional of Gaussian dependent sequences under a minimal assumption. *Statist. Probab. Lett.* **78** 2485–2489. MR2462683

[10] Coeurjolly, J.F. (2008). Hurst exponent estimation of locally self-similar Gaussian processes using sample quantiles. *Ann. Statist.* **36** 1404–1434. MR2418662

[11] Csörgő, S. and Mielniczuk, J. (1996). The empirical process of a short-range dependent stationary sequence under Gaussian subordination. *Probab. Theory Related Fields* **104** 15–25. MR1367664

[12] Donoho, D.L. and Johnstone, I.M. (1994). Ideal spatial adaptation by wavelet shrinkage. *Biometrika* **81** 425–455. MR1311089

[13] Faÿ, G., Moulines, E., Roueff, F. and Taqqu, M.S. (2009). Estimators of long-memory: Fourier versus wavelets. *J. Econometrics* **151** 159–177. MR2559823

[14] Geweke, J. and Porter-Hudak, S. (1983). The estimation and application of long memory time series models. *J. Time Ser. Anal.* **4** 221–238. MR0738585

[15] Johnstone, I.M. and Silverman, B.W. (1997). Wavelet threshold estimators for data with correlated noise. *J. Roy. Statist. Soc. Ser. B* **59** 319–351. MR1440585

[16] Lévy-Leduc, C., Boistard, H., Moulines, E., Taqqu, M.S. and Reisen, V.A. (2011). Robust estimation of the scale and of the autocovariance function of Gaussian short- and long-range dependent processes. *J. Time Series Anal.* **32** 135–156. MR2807883

[17] Ma, Y. and Genton, M.G. (2000). Highly robust estimation of the autocovariance function. *J. Time Ser. Anal.* **21** 663–684. MR1801704

[18] Molinares, F.F., Reisen, V.A. and Cribari-Neto, F. (2009). Robust estimation in long-memory processes under additive outliers. *J. Statist. Plann. Inference* **139** 2511–2525. MR2523644

[19] Moulines, E., Roueff, F. and Taqqu, M.S. (2007). Central limit theorem for the log-regression wavelet estimation of the memory parameter in the Gaussian semi-parametric context. *Fractals* **15** 301–313. MR2396718

[20] 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

[21] Moulines, E. and Soulier, P. (2003). Semiparametric spectral estimation for fractional processes. In *Theory and Applications of Long-Range Dependence* 251–301. Boston, MA: Birkhäuser. MR1956053
[22] Park, C., Godtliebsen, F., Taqqu, M., Stoev, S. and Marron, J.S. (2007). Visualization and inference based on wavelet coefficients, SiZer and SiNos. *Comput*. *Statist*. *Data Anal*. **51** 5994–6012. MR2407694
[23] Park, J. and Park, C. (2009). Robust estimation of the Hurst parameter and selection of an onset scaling. *Statist*. *Sinica* **19** 1531–1555. MR2589196
[24] Percival, D.B. and Walden, A.T. (2006). *Wavelet Methods for Time Series Analysis*. *Cambridge Series in Statistical and Probabilistic Mathematics* **4**. Cambridge: Cambridge Univ. Press. MR2218866
[25] Robinson, P.M. (1995). Gaussian semiparametric estimation of long range dependence. *Ann*. *Statist*. **23** 1630–1661. MR1370301
[26] Rosenblatt, M. (1985). *Stationary Sequences and Random Fields*. Boston, MA: Birkhäuser. MR0885090
[27] Rousseeuw, P.J. and Croux, C. (1993). Alternatives to the median absolute deviation. *J. Amer. Statist*. *Assoc*. **88** 1273–1283. MR1245360
[28] Shen, H., Zhu, Z. and Lee, T. (2007). Robust estimation of the self-similarity parameter in network traffic using wavelet transform. *Signal Process*. **87** 2111–2124.
[29] Stoev, S., Taqqu, M., Park, C. and Marron, J.S. (2005). On the wavelet spectrum diagnostic for Hurst parameter estimation in the analysis of internet traffic. *Comput*. *Netw*. **48** 423–445.
[30] Stoev, S. and Taqqu, M.S. (2005). Asymptotic self-similarity and wavelet estimation for long-range dependent fractional autoregressive integrated moving average time series with stable innovations. *J*. *Time Ser. Anal*. **26** 211–249. MR2122896
[31] Stoev, S., Taqqu, M.S., Park, C., Michailidis, G. and Marron, J.S. (2006). LASS: A tool for the local analysis of self-similarity. *Comput*. *Statist*. *Data Anal*. **50** 2447–2471. MR2225578
[32] van der Vaart, A.W. (1998). *Asymptotic Statistics*. *Cambridge Series in Statistical and Probabilistic Mathematics* **3**. Cambridge: Cambridge Univ. Press. MR1652247

Received November 2010 and revised May 2011