Asymptotic normality of wavelet estimators of the memory parameter for linear processes

François Roueff, Murad S. Taqqu

ASYMPTOTIC NORMALITY OF WAVELET ESTIMATORS OF THE MEMORY PARAMETER FOR LINEAR PROCESSES

F. ROUEFF AND M.S. TAQQU

TELECOM ParisTech and Boston University

Abstract. We consider linear processes, not necessarily Gaussian, with long, short or negative memory. The memory parameter is estimated semi-parametrically using wavelets from a sample $X_1, \dots, X_n$ of the process. We treat both the log-regression wavelet estimator and the wavelet Whittle estimator. We show that these estimators are asymptotically normal as the sample size $n \to \infty$ and we obtain an explicit expression for the limit variance. These results are derived from a general result on the asymptotic normality of the empirical scalogram for linear processes, conveniently centered and normalized. The scalogram is an array of quadratic forms of the observed sample, computed from the wavelet coefficients of this sample. In contrast with quadratic forms computed on the Fourier coefficients such as the periodogram, the scalogram involves correlations which do not vanish as the sample size $n \to \infty$.

Contents

Date: May 6, 2008.

1991 *Mathematics Subject Classification.* Primary 62M10, 62M15, 62G05 Secondary: 60G18.

Key words and phrases. Spectral analysis, Wavelet analysis, long range dependence, semiparametric estimation.

Murad S. Taqqu would like to thank l'École Normale Supérieure des Télecommunications in Paris for their hospitality. This research was partially supported by the NSF Grants DMS–0505747 and DMS–0706786 at Boston University.

Corresponding author: F. Roueff, TELECOM ParisTech, 46, rue Barrault, 75634 Paris Cédex 13, France.

1. Introduction

We consider a real-valued process $X \stackrel{\text{def}}{=} \{X_k\}_{k\in\mathbb{Z}}$, not necessarily stationary and for any positive integer k, let $\mathbf{\Delta}^k X$ denote its k-th order difference. The first order difference is $[\mathbf{\Delta} X]_t \stackrel{\text{def}}{=} X_t - X_{t-1}$ and $\mathbf{\Delta}^k$ is defined recursively.

Definition 1 ($M(d)$ processes)**.** *The process X is said to have* memory parameter *d, $d \in \mathbb{R}$ (in short, is an M(d) process) and* short-range *spectral density f^* if for any integer $k > d - 1/2$, the k-th order difference process $\mathbf{\Delta}^k X$ is weakly stationary with spectral density function*

$$f_{\mathbf{\Delta}^k X}(\lambda) \stackrel{\text{def}}{=} |1 - e^{-i\lambda}|^{2(k-d)}\, f^*(\lambda) \quad \lambda \in (-\pi, \pi), \tag{1}$$

where f^* is a non-negative symmetric function which is continous and non-zero at the origin.

M(d) processes encompass both stationary and non-stationary processes, depending on the value of the memory parameter d. The function

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

is called the *generalized spectral density* of X. It is a proper spectral density function when $d < 1/2$. In this case, the process X is covariance stationary with spectral density function f. The process X is said to have long-memory if $0 < d < 1/2$, short-memory if $d = 0$ and negative memory if $d < 0$; the process is not invertible if $d < -1/2$. The factor f^* is a nuisance function which determine the "short-range" dependence.

In a typical *semiparametric* estimation setting (see for instance [9, 3, 8]), the following additional assumption is often considered.

Assumption 1. There exists $\beta \in (0, 2]$, $\gamma > 0$ and $\varepsilon \in (0, \pi]$ such that for all $\lambda \in [-\varepsilon, \varepsilon]$,

$$|f^*(\lambda) - f^*(0)| \le L\, f^*(0)\, |\lambda|^\beta\ . \tag{3}$$

Moreover, $f^*(0) > 0$.

We consider an $M(d)$ process satisfying the following *linear assumption*.

Assumption 2. There exists a non-negative integer K such that

$$[\mathbf{\Delta}^K X]_k = \sum_{t\in\mathbb{Z}} a^{(K)}(k-t)\, \xi_t\ , \tag{4}$$

where $\{a^{(K)}(t),\, t \in \mathbb{Z}\}$ is a real-valued sequence satisfying $\sum_t (a^{(K)}(t))^2 < \infty$ and

(A-1) $\{\xi_l,\, l \in \mathbb{Z}\}$ is a sequence of independent and identically distributed real-valued random variables such that $\mathbb{E}[\xi_0] = 0$, $\mathbb{E}[\xi_0^2] = 1$ and $\kappa_4 \stackrel{\text{def}}{=} \mathbb{E}[\xi_0^4] - 3$ is finite.

Here the linear assumption may only apply to a K-order increment of X to allow X to be non-stationary.

Our goal is to estimate d by using a Discrete Wavelet Transform (DWT) of X. In order to study the asymptotic properties of the estimator, we use a central limit theorem for an array of squares of decimated linear processes, established in [10], see Theorems 1 and 2 in this reference. Using this result, we extend to the non-Gaussian linear processes setting, asymptotic normality results for wavelet estimation of the memory parameter d, that have been obtained so far for Gaussian processes (see [4, Thoerem 1] and [7, Theorem 5]). We treat both the log-regression wavelet estimator and the wavelet Whittle estimator.

In Section 3, we give a simplified formulation of the central limit theorem [10, Theorem 2] and apply it to the Discrete Wavelet Transform setting, obtaining a result on the asymptotic distribution of the scalogram of a linear memory process as the scale index and the number of observed wavelet coefficients both tend to infinity, see Theorem 2. We then consider two estimators of the memory parameter d, the log–regression wavelet estimator in Section 4 and the wavelet Whittle estimator in Section 5. Using Theorem 2, we show that both these estimators are asymptotically normal.

2. Definition of the empirical scalogram of a finite sample

We now introduce the wavelet setting and recall the definition of the *scalogram* and the empirical scalogram. Introduce the functions $\phi(t)$, $t \in \mathbb{R}$, and $\psi(t)$, $t \in \mathbb{R}$, which will play the role of the father and mother wavelets respectively, and let $\hat{\phi}(\xi) \stackrel{\text{def}}{=} \int_{\mathbb{R}} \phi(t) \mathrm{e}^{-\mathrm{i}\xi t} \, \mathrm{d}t$ and $\hat{\psi}(\xi) \stackrel{\text{def}}{=} \int_{\mathbb{R}} \psi(t) \mathrm{e}^{-\mathrm{i}\xi t} \, \mathrm{d}t$ denote their Fourier transforms. We suppose that the wavelets ϕ and ψ satisfy the following assumptions :

- (W-1) ϕ and ψ are integrable and have compact supports, $\hat{\phi}(0) = \int_{\mathbb{R}} \phi(x) \mathrm{d}x = 1$ and $\int_{\mathbb{R}} \psi^2(x) \mathrm{d}x = 1$.
- (W-2) There exists $\alpha > 1$ such that $\sup_{\xi \in \mathbb{R}} |\hat{\psi}(\xi)| \, (1 + |\xi|)^{\alpha} < \infty$,
- (W-3) The function ψ has M vanishing moments, *i.e.* $\int_{\mathbb{R}} t^l \psi(t) \, \mathrm{d}t = 0$ for all $l = 0, \ldots, M-1$
- (W-4) The function $\sum_{k \in \mathbb{Z}} k^l \phi(\cdot - k)$ is a polynomial of degree l for all $l = 0, \ldots, M-1$.

We now define what we call the DWT of X in discrete time. Define the family $\{\psi_{j,k}, j > 0, k \in \mathbb{Z}\}$ of translated and dilated functions

$$\psi_{j,k}(t) = 2^{-j/2} \, \psi(2^{-j} t - k) \; . \tag{5}$$

Using the scaling function ϕ, we first define the functions

$$\mathbf{X}_n(t) \stackrel{\text{def}}{=} \sum_{k=1}^{n} X_k \, \phi(t-k) \quad \text{and} \quad \mathbf{X}(t) \stackrel{\text{def}}{=} \sum_{k \in \mathbb{Z}} X_k \, \phi(t-k) \tag{6}$$

The (details) wavelet coefficients are then defined as follows, for all $j \geq 0, k \in \mathbb{Z}$,

$$W_{j,k} \stackrel{\text{def}}{=} \int_{-\infty}^{\infty} \mathbf{X}(t)\psi_{j,k}(t)\,\mathrm{d}t. \tag{7}$$

These wavelet coefficients are the DWT of X. If the support of the scaling function ϕ is included in $(-\mathrm{T}, 0)$ for some integer $\mathrm{T} \geq 1$, then $\mathbf{x}_n(t) = \mathbf{x}(t)$ for all $t = 0, \ldots, n - \mathrm{T} + 1$. If the support of the wavelet function ψ is included in $(0, \mathrm{T})$, then, the support of $\psi_{j,k}$ is included in the interval $(2^j k, 2^j(k + \mathrm{T}))$. Hence

$$W_{j,k} = \int_{-\infty}^{\infty} \mathbf{X}_n(t)\psi_{j,k}(t)\,\mathrm{d}t\ , \tag{8}$$

for all $(j, k) \in \mathcal{I}_n$, where

$$\mathcal{I}_n \stackrel{\text{def}}{=} \{(j, k) :\ j \geq 0, 0 \leq k \leq 2^{-j}(n - \mathrm{T} + 1) - \mathrm{T}\}\ . \tag{9}$$

For any j, the wavelet coefficients $\{W_{j,k}\}_{k\in\mathbb{Z}}$ are obtained by discrete convolution and downsampling. More precisely, under (W-1), for all $j \geq 0$, $k \in \mathbb{Z}$,

$$W_{j,k} = \sum_{l\in\mathbb{Z}} x_l\, h_{j,2^j k-l} = (h_{j,\cdot} \star X)_{2^j k} = (\downarrow^j [h_{j,\cdot} \star X])_k, \tag{10}$$

where $h_{j,l} \stackrel{\text{def}}{=} 2^{-j/2} \int_{-\infty}^{\infty} \phi(t + l)\psi(2^{-j}t)\,\mathrm{d}t$, $\star$ denotes the convolution of discrete sequences and, for any sequence $\{c_k\}_{k\in\mathbb{Z}}$, $(\downarrow^j c)_k = c_{k2^j}$. For all $j \geq 0$, $H_j(\lambda) \stackrel{\text{def}}{=} \sum_{l\in\mathbb{Z}} h_{j,l}\mathrm{e}^{-\mathrm{i}\lambda l}$ denotes the discrete Fourier transform of $\{h_{j,l}\}_{l\in\mathbb{Z}}$,

$$H_j(\lambda) \stackrel{\text{def}}{=} 2^{-j/2} \int_{-\infty}^{\infty} \sum_{l\in\mathbb{Z}} \phi(t + l)\mathrm{e}^{-\mathrm{i}\lambda l}\psi(2^{-j}t)\,\mathrm{d}t. \tag{11}$$

For all $j \geq 0$ and all $m = 0, \ldots, M - 1$,

$$\sum_{l\in\mathbb{Z}} h_{j,l}\, l^m = 2^{-j/2} \int_{-\infty}^{\infty} \psi(2^{-j}t) \sum_{l\in\mathbb{Z}} \phi(t + l)l^m \mathrm{d}t\ .$$

Under assumption (W-4), $t \mapsto \sum_{l\in\mathbb{Z}} \phi(t+l)l^m$ is a polynomial of degree m and (W-3) therefore implies that $\sum_{l\in\mathbb{Z}} h_{j,l}\, l^m = 0$; equivalently, the trigonometric polynomial H_j satisfies $\frac{\mathrm{d}^m H_j(\lambda)}{\mathrm{d}\lambda^m}\Big|_{\lambda=0} = 0$, $m = 0, \ldots, M - 1$ and thus admits a zero at 0 of degree at least equal to M. Therefore, $H_j(\lambda)$ can be factorized as

$$H_j(\lambda) = (1 - \mathrm{e}^{\mathrm{i}\lambda})^M \tilde{H}_j(\lambda)\ , \tag{12}$$

where $\tilde{H}_j(\lambda)$ is a trigonometric polynomial. Hence, the wavelet coefficient (10) may be computed as

$$W_{j,k} = (\downarrow^j [\tilde{h}_{j,\cdot} \star \boldsymbol{\Delta}^M X])_k \tag{13}$$

where $\{\tilde{h}_{j,l}\}_{l\in\mathbb{Z}}$ are the coefficients of the trigonometric polynomial $\tilde{H}_j$.

Let then $\{\phi, \psi\}$ be a pair of scale function and wavelet satisfying (W-1)–(W-4). Let $X = \{X_k,\ k \in \mathbb{Z}\}$ be a process such that $\mathbf{\Delta}^M X$ is weakly stationary and define the DWT $\{W_{j,k},\ j \geq 0,\ k \in \mathbb{Z}\}$ of X by (7). By (13), $\{W_{j,k},\ k \in \mathbb{Z}\}$ is a weakly stationary process for all scales $j \geq 0$.

Definition 2. *The* scalogram *of X is the non-negative sequence $\{\sigma_j^2,\ j \geq 0\}$ of variances of $\{W_{j,k},\ k \in \mathbb{Z}\}$, namely*

$$\sigma_j^2 \stackrel{\text{def}}{=} \mathrm{Var}[W_{j,0}] = \mathbb{E}\left[W_{j,0}^2\right], \quad j \geq 0\,. \tag{14}$$

Remark 1. Observe that, under Assumption 2, if $M \geq K$, then $\mathbf{\Delta}^M X$ is a centered weak stationary process and the scalogram of X is well defined.

Wavelet estimators of the memory parameter d are typically based on quadratic forms of the wavelet coefficients. This is reasonable because, for large scale j, $\log \sigma_j^2$ is approximately an affine function of j with slope $(2 \log 2)\, d$ (see [6]) and, given n observations $X_1, \ldots, X_n$, σ_j^2 can be estimated by the empirical second moment

$$\hat{\sigma}_j^2 \stackrel{\text{def}}{=} n_j^{-1} \sum_{k=0}^{n_j - 1} W_{j,k}^2\,, \tag{15}$$

which is a quadratic form on the wavelet coefficients. Here we denote by n_j the number of available wavelet coefficients at scale index j, namely, from (9),

$$n_j = [2^{-j}(n - \mathrm{T} + 1) - \mathrm{T} + 1]\,, \tag{16}$$

where T is the size of the time series and $[x]$ denotes the integer part of x. It is important to note that although the wavelet coefficient $W_{j,k}$ does not depend on n, the empirical second moment $\hat{\sigma}_j^2$ does through n_j.

Definition 3. *Let $\{\phi, \psi\}$ be a pair of scale function and wavelet satisfying (W-1)–(W-4) and $n \geq 1$. Let us denote the maximal scale index $J = J(n)$ by*

$$J \stackrel{\text{def}}{=} \max\{j :\ n_j > 0\} = \left\lceil \log_2 \left(\frac{n - \mathrm{T} + 1}{\mathrm{T}} \right) \right\rceil\,, \tag{17}$$

where n_j is defined by (16). The empirical scalogram *of the sample $\{X_1, \ldots, X_n\}$ is the non-negative process $\{\hat{\sigma}_j^2,\ j \geq 0\}$, where*

- *for all $j = 1, \ldots, J$, $\hat{\sigma}_j^2$ is defined by (15),*
- *and by convention, $\hat{\sigma}_j^2 = 0$ for $j > J$.*

The estimator of the memory parameter d can then be obtained as follows :

(1) by regressing the logarithm of the empirical variance $\log(\hat{\sigma}_i^2)$ for a finite number of scale indices $j \in \{L, \dots, U\}$ where L is the lower scale and $U \le J$ is the upper scale in the regression, see Section 4.
(2) by minimizing a contrast derived from the likelihood of an array of independent Gaussian random variables each row $j \in \{L, \dots, U\}$ of which having empirical variance $\log(\hat{\sigma}_i^2)$, see Section 5.

3. Joint weak convergence of the empirical scalogram of a linear process

We let $\xrightarrow{\mathcal{L}}$ denote the convergence in law. For convenience, we first state a CLT based on results of [10].

Theorem 1. *Let $\{v_{i,j}(t),\ t \in \mathbb{Z}\}$ be real-valued sequences satisfying $\sum_{t\in\mathbb{Z}} v_{i,j}^2(t) < \infty$ for all $i = 1, \dots, N$ and $j \ge 0$. Suppose that there exist $\delta > 1/2$, $\varepsilon \in (0, \pi]$, a sequence of $[-\pi, \pi)$-valued functions $\Phi_j(\lambda)$ defined on $\lambda \in \mathbb{R}$ and continuous functions $v_{i,\infty}^*$, $i = 1, \dots, N$, defined on $\mathbb{R}$ such that*

$$\sup_{j\ge 0} \sup_{|\lambda|\le\varepsilon} 2^{-j/2} |v_{i,j}^*(\lambda)| (1 + 2^j |\lambda|)^\delta < \infty \,, \tag{18}$$

$$\lim_{j\to\infty} 2^{-j/2} v_{i,j}^*(2^{-j}\lambda) \mathrm{e}^{\mathrm{i}\Phi_j(\lambda)} = v_{i,\infty}^*(\lambda) \quad \textit{for all} \quad \lambda \in \mathbb{R} \,, \tag{19}$$

$$n_j^{1/2} \int_0^\pi \mathbb{1}(|\lambda| > \varepsilon)\, |v_{i,j}^*(\lambda)|^2 \,\mathrm{d}\lambda \to 0 \quad \textit{as} \quad j \to \infty \,, \tag{20}$$

where $v_{i,j}^*$ denotes the Fourier series associated to the sequence $\{v_{i,j}(t),\ t \in \mathbb{Z}\}$,

$$v_{i,j}^*(\lambda) = (2\pi)^{-1/2} \sum_{t\in\mathbb{Z}} v_{i,j}(t)\, \mathrm{e}^{-\mathrm{i}\lambda t} \,. \tag{21}$$

Define $\{Z_{i,j,k},\ i = 1, \dots, N, j \ge 0, k \in \mathbb{Z}\}$ as

$$Z_{i,j,k} = \sum_{t\in\mathbb{Z}} v_{i,j}(2^j k - t)\, \xi_t \,, \quad i = 1, \dots, N,\ k \in \mathbb{Z},\ j \ge 0 \,, \tag{22}$$

where $\{\xi_t,\ t \in \mathbb{Z}\}$ satisfies (A-1). Then, for any diverging sequence (n_j), as $j \to \infty$,

$$n_j^{-1/2} \sum_{k=0}^{n_j-1} \begin{bmatrix} Z_{1,j,k}^2 - \mathbb{E}[Z_{1,j,k}^2] \\ \vdots \\ Z_{N,j,k}^2 - \mathbb{E}[Z_{N,j,k}^2] \end{bmatrix} \xrightarrow{\mathcal{L}} \mathcal{N}(0, \Gamma) \,, \tag{23}$$

where Γ is the covariance matrix defined by

$$\Gamma_{i,i'} = 4\pi \int_{-\pi}^{\pi} \left| \sum_{p\in\mathbb{Z}} v_{i,\infty}^* \overline{v_{i',\infty}^*}(\lambda + 2p\pi) \right|^2 \mathrm{d}\lambda \,, \quad 1 \le i, i' \le N \,. \tag{24}$$

Moreover, one has

$$\lim_{j\to\infty} \mathbb{E}\left[Z_{i,j,0}^2\right] = \int_{-\infty}^{\infty} \left|v_{i,\infty}^*(\lambda)\right|^2 \,\mathrm{d}\lambda < \infty \,. \tag{25}$$

Proof. Observe first that we allowed $\varepsilon = \pi$, in which case Condition (20) is always satisfied since the integral vanishes for all $j \geq 0$. The CLT (23) is a strict application of Theorem 1 in [10] for $\varepsilon = \pi$ and Theorem 2 in [10] for $\varepsilon < \pi$. Using the notations of [10] we have here $\gamma_j = 2^j$ for all $j \geq 0$ and $\lambda_{i,\infty} = 0$ for all $i = 1, \ldots, N$. Observe that in this case, in these two theorems, $\mathrm{C}_{i,i'} = 1$ and $w_{i,i'}^* = v_{i,\infty}^* \overline{v_{i',\infty}^*}$ for all $i, i' = 1, \ldots, N$. The limit (25) follows from Relation (15) in Proposition 1 in the same paper, see also Remark 11. □

Remark 2. When the convergence rate to the limit (25) is fast enough, we will be able to replace the expectations in (23) by $\int_{-\infty}^{\infty} \left|v_{i,\infty}^*(\lambda)\right|^2 \,\mathrm{d}\lambda$, $i = 1, \ldots, N$, which does not depend on j.

In the Gaussian case the spectral density of the wavelet coefficients is sufficient for studying the quadratic forms (15) since second order properties fully determine the distribution of the wavelet coefficients. This is not so in the linear case, which we consider in this paper. It is possible, however, to establish a multivariate central limit theorem for the empirical scalogram.

Theorem 2. *Let X be an $M(d)$ process with short-range spectral density f^* and suppose that Assumption 2 holds. Assume that (W-1)–(W-4) hold with*

$$1/2 - \alpha < d \leq M \quad \textit{and} \quad K \leq M \,. \tag{26}$$

Let $L = L(n)$ be a scale index depending on n such that $L(n) \to \infty$ and $n2^{-L(n)} \to \infty$ as $n \to \infty$. Assume that one of the two following conditions hold.

$$\sup_{\lambda\in(-\pi,\pi)} f^*(\lambda) < \infty \tag{27}$$

$$(n2^{-L(n)})^{1/2} 2^{L(n)(1-2\alpha-2d)} \to 0 \quad \textit{as } n \to \infty \,. \tag{28}$$

Then, as $n \to \infty$, one has the folowing central limit:

$$\left\{\sqrt{n2^{-L(n)}} 2^{-2L(n)d} (\hat{\sigma}^2_{L(n)+u} - \sigma^2_{L(n)+u}),\ u \geq 0\right\} \xrightarrow{\mathcal{L}} \left\{Q_u^{(d)},\ u \geq 0\right\} \,, \tag{29}$$

where $Q_\bullet^{(d)}$ denotes a centered Gaussian process defined on $\mathbb{N}$ with covariance function $\mathbf{\Lambda}_{u,u'}(d) = \mathrm{Cov}\left(Q_u^{(d)}, Q_{u'}^{(d)}\right)$, $u, u' \geq 0$, given by

$$\mathbf{\Lambda}_{u,u'}(d) \stackrel{\mathrm{def}}{=} 4\pi \left(f^*(0)\right)^2 2^{4d(u\vee u')+u\wedge u'} \int_{-\pi}^{\pi} \left|\mathbf{D}_{\infty,|u-u'|}(\lambda)\right|^2 \,\mathrm{d}\lambda \,, \tag{30}$$

with, for all $u \geq 0$ *and* $\lambda \in (-\pi, \pi)$,

$$\mathbf{D}_{\infty,u}(\lambda) \stackrel{\text{def}}{=} \sum_{l \in \mathbb{Z}} |\lambda + 2l\pi|^{-2d} \, \mathbf{e}_u(\lambda + 2l\pi) \, \overline{\hat{\psi}(\lambda + 2l\pi)} \hat{\psi}(2^{-u}(\lambda + 2l\pi)). \tag{31}$$

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

Remark 3. We assume $1 - 2\alpha - 2d < 0$ in (26) so that (28) imposes a sufficiently fast growth rate on $L(n)$ as $n \to \infty$. On the other hand this rate has to be slow enough for the assumption $n2^{-L(n)} \to \infty$ to hold.

Proof. In (4), the sequence $\{a^{(K)}(t),\ t \in \mathbb{Z}\}$ depends on K. To define a quantity which does not, we go to the Fourier domain and set

$$a^*(\lambda) \stackrel{\text{def}}{=} (2\pi)^{-1/2} \, (1 - \mathrm{e}^{-\mathrm{i}\lambda})^{-K} \sum_{t \in \mathbb{Z}} a^{(K)}(t) \, \mathrm{e}^{-\mathrm{i}\lambda t} \, ,$$

where the sum over $t \in \mathbb{Z}$ converges in the sense of $L^2(-\pi, \pi)$. This function $a^*(\lambda)$ satisfies

$$|a^*(\lambda)|^2 = |1 - e^{-\mathrm{i}\lambda}|^{-2d} \, f^*(\lambda) = f(\lambda) \, , \tag{32}$$

where f is defined in (2). Moreover, by (13), since $K \leq M$ (see Condition (26)), for all $j \in \mathbb{N}$ and $k \in \mathbb{Z}$, the wavelet coefficients of X can be expressed as

$$W_{j,k} = (\downarrow^j \, [\tilde{h}_{j,\cdot} \star \boldsymbol{\Delta}^{M-K}(a^{(K)} \star \xi)])_k \, .$$

Since $\tilde{h}_{j,\cdot}$ is a finite sequence, we obtain that

$$W_{j,k} = \sum_{t \in \mathbb{Z}} a_j(k2^j - t) \, \xi_t \, , \tag{33}$$

where $\{a_j(t),\, t \in \mathbb{Z}\}$ is the sequence $\tilde{h}_{j,\cdot} \star \boldsymbol{\Delta}^{M-K}(a^{(K)})$ which is characterized by the $L^2(-\pi, \pi)$ converging series

$$a_j^*(\lambda) \stackrel{\text{def}}{=} (2\pi)^{-1/2} \sum_{t \in \mathbb{Z}} a_j(t) \, \mathrm{e}^{-\mathrm{i}\lambda t} = \tilde{H}_j(\lambda)(1 - \mathrm{e}^{-\mathrm{i}\lambda})^M a^*(\lambda) \, , \tag{34}$$

which, in view of (12), can be simply written as

$$a_j^*(\lambda) = H_j(\lambda) a^*(\lambda) \, . \tag{35}$$

To prove the theorem, we need to show that, for any integer $\ell \geq 0$, one has

$$\sqrt{n2^{-L}} 2^{-2Ld} \left(\begin{bmatrix} \hat{\sigma}_L^2 - \sigma_L^2 \\ \hat{\sigma}_{L+1}^2 - \sigma_{L+1}^2 \\ \vdots \\ \hat{\sigma}_{L+\ell}^2 - \sigma_{L+\ell}^2 \end{bmatrix} \right) \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \left[\boldsymbol{\Lambda}_{u,u'}(d),\, u, u' = 0, \ldots, \ell\right]\right) \, . \tag{36}$$

To this end, we will apply Theorem 1 by relating the right-hand side of (36) to the right-hand side of (23) and by expressing the empirical scalogram $\hat{\sigma}^2_{L+u}$, $0 \le u \le \ell$ in terms of $Z_{i,j,k}$ with adapted indices j, k and i.

We let $j = L+\ell$, that is j is the maximal scale in (36). We let k take values $k = 0, \ldots, n_j$, where n_j, given by (16) is the number of wavelet coefficients available at the maximal scale j. In Theorem 1, $Z_{i,j,k}$ is viewed as the i^{th} component of a k-wise stationary vector, with $i = 1, \ldots, N$. In order to recover this stationarity from the set of wavelet coefficients used to compute the empirical variances in (36), we do as follows. We represent i as $i = 2^{\ell-u} + v$ where $u \in \{0, \ldots, \ell\}$ and $v \in \{0, \ldots, 2^{\ell-u} - 1\}$ and let $N = \sum_{u=0}^{\ell} 2^{\ell-u} = 2^{\ell+1} - 1$. For each (u, v), we set $i = 2^{\ell-u} + v$, $j = L + \ell$, and

$$v_{i,j}(t) \stackrel{\text{def}}{=} 2^{-Ld} a_{L+u}(t + v2^{L+u}), \quad t \in \mathbb{Z}\,, \tag{37}$$

where a_{L+u} is defined in (33). Thus if we focus on a scale $j' \ge L$ and express it as $j' = L + u = j - \ell + u$ (see Figure 1), we have

$$v_{i,j}(t) = 2^{-Ld} a_{j'}(t + v2^{j'}), \quad t \in \mathbb{Z}\,. \tag{38}$$

Hence, by definition of $Z_{i,j,k}$ in (22), one has

$$\begin{aligned} Z_{i,j,k} &= \sum_{t\in\mathbb{Z}} v_{i,j}(2^j k - t)\xi_t \\ &= 2^{-Ld} \sum_{t\in\mathbb{Z}} a_{j'}(2^{j'}\{2^{\ell-u}k + v\} - t)\,\xi_t \\ &= 2^{-Ld} W_{j',2^{\ell-u}k+v}\,. \end{aligned} \tag{39}$$

By (16),

$$n_{j'} = 2^{-(j-\ell+u)}(n - \mathrm{T} + 1) - \mathrm{T} + 1 = 2^{\ell-u}\, n_j + (\mathrm{T} - 1)(2^{\ell-u} - 1),$$

and hence by (14), (15) and (39),

$$\begin{aligned} 2^{-2Ld}\left(\hat{\sigma}^2_{j'} - \sigma^2_{j'}\right) &= n_{j'}^{-1} 2^{-2Ld} \sum_{k'=0}^{n_{j'}-1} (W^2_{j',k'} - \mathbb{E}[W^2_{j',k'}]) \\ &= n_{j'}^{-1} \sum_{v=0}^{2^{\ell-u}-1} \left(\sum_{k=0}^{n_j-1} \left\{ Z^2_{i,j,k} - \mathbb{E}[Z^2_{i,j,k}] \right\} \right) + R_{j'}\,, \end{aligned} \tag{40}$$

where $j' = j - \ell + u$, $k' = 2^{\ell-u}k + v$, $i = 2^{\ell-u} + v$ and

$$R_{j'} \stackrel{\text{def}}{=} n_{j'}^{-1} \sum_{v=0}^{(\mathrm{T}-1)(2^{\ell-u}-1)-1} \left\{ Z^2_{i,j,n_j} - \mathbb{E}\left[Z^2_{i,j,n_j} \right] \right\}\,. \tag{41}$$

We then have

$$2^{-2Ld}\begin{bmatrix} \hat\sigma^2_L-\sigma^2_L \\ \hat\sigma^2_{L+1}-\sigma^2_{L+1} \\ \vdots \\ \hat\sigma^2_{L+\ell}-\sigma^2_{L+\ell} \end{bmatrix} = A_n \begin{bmatrix} \sum_{k=0}^{n_j-1}\{Z^2_{1,j,k}-\mathbb{E}[Z^2_{1,j,k}]\} \\ \sum_{k=0}^{n_j-1}\{Z^2_{2,j,k}-\mathbb{E}[Z^2_{1,j,k}]\} \\ \vdots \\ \sum_{k=0}^{n_j-1}\{Z^2_{N,j,k}-\mathbb{E}[Z^2_{N,j,k}]\} \end{bmatrix} + \begin{bmatrix} R_L \\ R_{L+1} \\ \vdots \\ R_{L+\ell} \end{bmatrix}, \tag{42}$$

where

$$A_n \stackrel{\text{def}}{=} \begin{bmatrix} & & \overbrace{\phantom{n_L^{-1}\dots n_L^{-1}}}^{2^\ell \text{ times}} \\ 0\dots & \dots 0 & n_L^{-1}\dots n_L^{-1} \\ & \overbrace{\phantom{n_{L+1}^{-1}\dots n_{L+1}^{-1}}}^{2^{\ell-1} \text{ times}} & \\ 0\dots 0 & n_{L+1}^{-1}\dots n_{L+1}^{-1} & 0\dots 0 \\ \vdots & \vdots & \vdots \\ n_{L+\ell}^{-1}\ 0\dots 0 & 0\dots 0 & 0\dots 0 \end{bmatrix} \tag{43}$$

is an $(\ell+1)\times N$ matrix. The entries are $n_{j'}^{-1}=n_{L+u}^{-1}$, where u goes from 0 (top line) to ℓ (bottom line), see Figure 1.

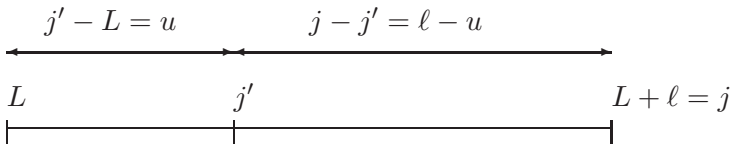

FIGURE 1. This figure indicates the relationship between the various variables.

Let us check that the assumptions of Theorem 1 apply to (37), that is, we show that (18),(20) and (19) are verified for $v_{i,j}(t)$ defined by (37). Using (38), (21), (34) and (35), we get

$$\begin{aligned} v^*_{i,j}(\lambda) &= 2^{-Ld}\,\mathrm{e}^{\mathrm{i}\lambda v 2^{j'}}\ a^*_{j'}(\lambda) \\ &= 2^{-Ld}\,\mathrm{e}^{\mathrm{i}\lambda v 2^{j'}}\ H_{j'}(\lambda)a^*(\lambda)\,. \end{aligned}$$

By continuity of f^* at the origin we have $\sup_{\lambda\in(-\varepsilon,\varepsilon)}\sqrt{f^*(\lambda)}\le C$ for some $C>0$ and $\varepsilon\in(0,\pi]$. Morever, under (27), we may set $\varepsilon=\pi$. By (32) and (2) we get, for all $\lambda\in(-\varepsilon,\varepsilon)$,

$$|a^*(\lambda)|\le|1-\mathrm{e}^{-\mathrm{i}\lambda}|^{-d}\,\sqrt{C}\le\sqrt{C}\,|\lambda|^{-d}\,.$$

By [6, Proposition 3], we have, for all $\lambda\in(-\pi,\pi)$,

$$\left|H_{j'}(\lambda)-2^{j'/2}\hat\phi(\lambda)\overline{\hat\psi(2^{j'}\lambda)}\right|\le C\,2^{j'(1/2-\alpha)}|\lambda|^M$$

and

$$\left|H_{j'}(\lambda)\right| \le C\ 2^{j'/2}\ |2^{j'}\lambda|^M\ (1+2^{j'}|\lambda|)^{-\alpha-M}\ ,$$

where C is a positive constant and α and M are defined in (W-2) and (W-3), respectively. Using that $j = L+\ell$ and $j' = L+u$ for some u only depending on i and L, $j' = j + O(1)$, the last 4 displays and Condition (26) easily yield

$$|v^*_{i,j}(\lambda)| \le C\ 2^{j/2}\ (1+2^j|\lambda|)^{-\alpha-d} \quad \text{for all } \lambda \in (-\varepsilon,\varepsilon) \tag{44}$$

$$\int_\varepsilon^\pi |v^*_{i,j}(\lambda)|^2\ \mathrm{d}\lambda \le C\ 2^{L(1-2\alpha-2d)} \int_\varepsilon^\pi f^*(\lambda)\ \mathrm{d}\lambda\ , \tag{45}$$

and the bound

$$\begin{aligned}\Big|2^{-j/2}v^*_{i,j}(2^{-j}\lambda) - 2^{(u-\ell)/2-Ld}\ \mathrm{e}^{\mathrm{i}\lambda v 2^{u-\ell}}\ \hat\phi(2^{-j}\lambda)\overline{\hat\psi(2^{u-\ell}\lambda)}a^*(2^{-j}\lambda)\Big| \\ = 2^{-j/2-Ld}\ \left|a^*(2^{-j}\lambda)\right|\ \left|H_{j'}(2^{-j}\lambda) - 2^{j'/2}\hat\phi(2^{-j}\lambda)\overline{\hat\psi(2^{u-\ell}\lambda)}\right| \\ \le C\ 2^{-j(\alpha+M)}\ |\lambda|^{M-d}\ ,\end{aligned} \tag{46}$$

valid for $2^{-j}|\lambda| \le \varepsilon$ with C denoting some positive constant depending neither on λ nor on $j \ge 0$. Relation (44) is (18) with $\delta = \alpha + d > 1/2$. Under (27), $\varepsilon = \pi$ and (20) trivially holds (see the proof of Theorem 1). Otherwise, since (16) and $n2^{-L} \to \infty$ imply $n_j \sim n2^{-j} = n2^{-L-\ell}$, Relations (45), (28) and the fact that f^* is always integrable away of the origin (since $|1-\mathrm{e}^{-\mathrm{i}\lambda}|^{K-d}f^*(\lambda)$ is a spectral density and $|1-\mathrm{e}^{-\mathrm{i}\lambda}|^{K-d}$ is lower bounded for λ away of zero) imply (20). By (W-1), $\hat\phi$ is continuous at the origin where it takes value 1 and using (46), (32), (2) and the continuity of $f^*(\lambda)$ at $\lambda = 0$, we have, for all $\lambda \in \mathbb{R}$,

$$\hat\phi(2^{-j}\lambda)2^{-Ld}|a^*(2^{-j}\lambda)| \to 2^{\ell d}\ \sqrt{f^*(0)}\ |\lambda|^{-d} \quad \text{as} \quad j \to \infty\ .$$

Hence we obtain (19) with

$$v^*_{i,\infty}(\lambda) = 2^{(u-\ell)/2+\ell d}\ \mathrm{e}^{\mathrm{i}\lambda v 2^{u-\ell}}\ \sqrt{f^*(0)}\ |\lambda|^{-d}\ \overline{\hat\psi(2^{u-\ell}\lambda)} \tag{47}$$

for all $i = 2^{\ell-u} + v$ such that $u \in \{0,\dots,\ell\}$ and $v \in \{0,\dots,2^{\ell-u}-1\}$, and

$$\mathrm{e}^{\mathrm{i}\Phi_j(\lambda)} \stackrel{\text{def}}{=} \frac{\overline{a^*(\lambda)}}{|a^*(\lambda)|}\ .$$

Since (18),(20) and (19) hold and $n_j \to \infty$, we may apply Theorem 1 and obtain (23).

Observe that (16) and $n2^{-L} \to \infty$ imply, for $j' = L+u \in \{L,\dots,L+\ell\}$,

$$n_{j'} = n2^{-j'} + O(1) \sim n2^{-L-u} \to \infty \quad \text{as} \quad n \to \infty\ .$$

Since $(n2^{-L})^{1/2} n_{j'}^{-1} \sim (n2^{-L})^{1/2}(n2^{-L-u})^{-1} = (n2^{-j})^{-1/2} 2^{-\ell/2} 2^u \sim n_j^{-1/2} 2^{-\ell/2} 2^u$, (43) yields, as $n \to \infty$,

$$\sqrt{n2^{-L}}\, A_n \sim n_j^{-1/2}\, 2^{-\ell/2}\, A_\infty \quad \text{with} \quad A_\infty \stackrel{\text{def}}{=} \begin{bmatrix} & & \overbrace{}^{2^\ell \text{ times}} \\ 0\dots & \dots 0 & 1\dots1 \\ & \overbrace{}^{2^{\ell-1} \text{ times}} & \\ 0\dots0 & 2\dots2 & 0\dots0 \\ \vdots & \vdots & \vdots \\ 2^\ell\ 0\dots0 & 0\dots0 & 0\dots0 \end{bmatrix} .$$

The general term is 2^u for $u \in \{0, \dots, \ell\}$. Relations (25) and (41) give that, for $j' = L+u \in \{L, \dots, L+\ell\}$,

$$\sqrt{n2^{-L}}\mathbb{E}[R_{j'}] = O\left((n2^{-L})^{-1/2}\right) \to 0 \quad \text{as} \quad n \to \infty\,.$$

Applying (23), (42), the two last displays and Slutsky's lemma, we get (36) with

$$\mathbf{\Lambda}(d) = 2^{-\ell}\, A_\infty \Gamma A_\infty^T = 2^{-\ell} \left[2^{u+u'} \sum_{v=0}^{2^{\ell-u}-1} \sum_{v'=0}^{2^{\ell-u'}-1} \Gamma_{2^{\ell-u}+v, 2^{\ell-u'}+v'} \right]_{0 \le u,u' \le \ell} ,$$

where the indices (u, u') run from $(0,0)$ (top left corner) to $(u,u') = (\ell,\ell)$ (bottom right corner) and $\Gamma_{i,i'}$ is defined by (24) with $v^*_{i,\infty}$ and $v^*_{i',\infty}$ defined by (47) for $i, i' \in 1, \dots, N = 2^{\ell+1} - 1$. To conclude the proof, it remains to check that the entries of $\mathbf{\Lambda}_{u,u'}(d)$ as defined above are equal to those given in (30). We shall do that for $u' \ge u$ since the alternative case is obtained by observing that $\mathbf{\Lambda}_{u,u'}(d) = \mathbf{\Lambda}_{u',u}(d)$. Replacing $\Gamma_{i,i'}$ and then $v^*_{i,\infty}$ and $v^*_{i',\infty}$ by these expressions and denoting

$$\lambda_p \stackrel{\text{def}}{=} \lambda + 2p\pi, \quad \lambda \in \mathbb{R},\ p \in \mathbb{Z}\,,$$

we get, for $0 \le u, u' \le \ell$,

$$\begin{aligned} \mathbf{\Lambda}_{u,u'}(d) &= 2^{-\ell+u+u'} \sum_{v=0}^{2^{\ell-u}-1} \sum_{v'=0}^{2^{\ell-u'}-1} 4\pi\, (f^*(0))^2 \\ &\quad \times \int_{-\pi}^{\pi} 2^{u+u'-2\ell+4\ell d} \left| \sum_{p\in\mathbb{Z}} \mathrm{e}^{\mathrm{i}\lambda_p (v 2^{u-\ell} - v' 2^{u'-\ell})}\ |\lambda_p|^{-2d}\, \overline{\hat\psi(2^{u-\ell}\lambda_p)} \hat\psi(2^{u'-\ell}\lambda_p) \right|^2 \mathrm{d}\lambda \\ &= (f^*(0))^2\, 4\pi\, 2^{2(u+u')+\ell(4d-3)} \sum_{v=0}^{2^{\ell-u}-1} \int_{-\pi}^{\pi} G_{u,u',v}(\lambda)\, \mathrm{d}\lambda\,, \end{aligned} \tag{48}$$

where $G_{u,u',v}$ is a (2π)-periodic function defined by

$$G_{u,u',v}(\lambda) \stackrel{\text{def}}{=} \sum_{v'=0}^{2^{\ell-u'}-1} \left| \sum_{p\in\mathbb{Z}} \mathrm{e}^{\mathrm{i}\,\lambda_p (v 2^{u-\ell} - v' 2^{u'-\ell})} g_{u,u'}(2^{u'-\ell}\lambda_p) \right|^2 ,$$

with

$$g_{u,u'}(\lambda) \stackrel{\text{def}}{=} |2^{\ell-u'}\lambda|^{-2d} \overline{\hat{\psi}(2^{u-u'}\lambda)} \hat{\psi}(\lambda), \quad \lambda \in \mathbb{R} .$$

Writing $p = 2^{\ell-u'}q + r$ with $q \in \mathbb{Z}$ and $r \in \{0, \ldots, 2^{\ell-u'} - 1\}$ and transforming a sum over p into a sum over q and r, we get

$$\begin{aligned} G_{u,u',v}(\lambda) &= \sum_{v'=0}^{2^{\ell-u'}-1} \left| \sum_{r=0}^{2^{\ell-u'}-1} \mathrm{e}^{\mathrm{i}\,\lambda_r (v 2^{u-\ell} - v' 2^{u'-\ell})} \sum_{q\in\mathbb{Z}} \mathrm{e}^{\mathrm{i}\,2^{u-u'} v 2q\pi} g_{u,u'}(2^{u'-\ell}\lambda_r + 2q\pi) \right|^2 \\ &= \sum_{v'=0}^{2^{\ell-u'}-1} \left| \sum_{r=0}^{2^{\ell-u'}-1} \mathrm{e}^{-\mathrm{i}\,\lambda_r v' 2^{u'-\ell}} h_{u,u',v}(2^{u'-\ell}\lambda_r) \right|^2 , \end{aligned}$$

where

$$h_{u,u',v}(\lambda) \stackrel{\text{def}}{=} \sum_{q\in\mathbb{Z}} \mathrm{e}^{\mathrm{i}\,2^{u-u'} v \lambda_q} g_{u,u'}(\lambda_q) .$$

Hence

$$G_{u,u',v}(\lambda) = \sum_{v'=0}^{2^{\ell-u'}-1} \sum_{r=0}^{2^{\ell-u'}-1} \sum_{r'=0}^{2^{\ell-u'}-1} \mathrm{e}^{-\mathrm{i}\,2\pi(r-r')v'2^{u'-\ell}} h_{u,u',v}(2^{u'-\ell}\lambda_r) \overline{h_{u,u',v}(2^{u'-\ell}\lambda_{r'})} .$$

In the last display, observe that v' only appear in the complex exponential argument. Moreover we have $\sum_{v'=0}^{2^{\ell-u'}-1} \mathrm{e}^{-\mathrm{i}\,2\pi(r-r')v'2^{u'-\ell}} = 0$ except for $r = r'$ in which case it equals $2^{\ell-u'}$. Hence,

$$G_{u,u',v}(\lambda) = 2^{\ell-u'} \sum_{r=0}^{2^{\ell-u'}-1} \left| h_{u,u',v}(2^{u'-\ell}\lambda_r) \right|^2 .$$

Applying [10, Lemma 1] with $g = \left|h_{u,u',v}\right|^2$ and $\gamma = 2^{\ell-u'}$ gives

$$\begin{aligned} \int_{-\pi}^{\pi} G_{u,u',v}(\lambda)\,\mathrm{d}\lambda &= 2^{\ell-u'} \int_{-\pi}^{\pi} \sum_{r=0}^{2^{\ell-u'}-1} \left| h_{u,u',v}(2^{u'-\ell}\lambda_r) \right|^2 \,\mathrm{d}\lambda \\ &= 2^{2\ell-2u'} \int_{-\pi}^{\pi} \left| h_{u,u',v}(\lambda) \right|^2 \,\mathrm{d}\lambda . \end{aligned}$$

Inserting this equality in (48) and using that $\mathbf{\Lambda}_{u,u'}(d) = \mathbf{\Lambda}_{u',u}(d)$, we get, for all $0 \le u \le u' \le \ell$,

$$\mathbf{\Lambda}_{u,u'}(d) = (f^*(0))^2\ 4\pi\ 2^{2u+4du'-\ell} \times \sum_{v=0}^{2^{\ell-u}-1} \int_{-\pi}^{\pi} \left| \sum_{q\in\mathbb{Z}} |\lambda_q|^{-2d} \mathrm{e}^{\mathrm{i}\, 2^{-(u'-u)} v\lambda_q} \overline{\hat\psi(2^{-(u'-u)}\lambda_q)} \hat\psi(\lambda_q) \right|^2 \mathrm{d}\lambda\ .$$

For $v \in \{0,\dots,2^{\ell-u}-1\}$, we write $v = v' + k2^{u'-u}$ with $v' \in \{0,\dots,2^{u'-u}-1\}$ and $k \in \{0,\dots,2^{\ell-u'}-1\}$ and transform the sum over v into a sum over v' and k. But since $\exp\left\{\mathrm{i}2^{-(u'-u)}v\lambda_q\right\} = \exp\left\{\mathrm{i}2^{-(u'-u)}v'\lambda_q\right\}\exp\{\mathrm{i}k\lambda\}$, and $\sum_{k=0}^{2^{\ell-u'}-1}\left|\mathrm{e}^{\mathrm{i}k\lambda}\right|^2 = 2^{\ell-u'}$, we obtain

$$\mathbf{\Lambda}_{u,u'}(d) = (f^*(0))^2\ 4\pi\ 2^{4du'+u}2^{-(u'-u)} \times \sum_{v'=0}^{2^{u'-u}-1} \int_{-\pi}^{\pi} \left| \sum_{q\in\mathbb{Z}} |\lambda_q|^{-2d} \mathrm{e}^{\mathrm{i}\, 2^{-(u'-u)} v'\lambda_q} \overline{\hat\psi(2^{-(u'-u)}\lambda_q)} \hat\psi(\lambda_q) \right|^2 \mathrm{d}\lambda\ .$$

Relation (30) finally follows by observing that the vector with entries

$$2^{-(u'-u)/2}\left\{\mathrm{e}^{\mathrm{i}\, 2^{-(u'-u)}v'\lambda_q},\ v' = 0,\dots,2^{u'-u}-1\right\}$$

is precisely $\mathbf{e}_{u'-u}(\lambda_q)$. □

To obtain a result valid for an asymptotically infinite weighted sum of the empirical scalogram $\{\hat\sigma^2_{L+u} - \sigma^2_{L+u},\ u \ge 0\}$ as in Theorem 3 below, we need a bound for the covariance $\mathbf{\Lambda}_{u,u'}(d)$ defined in (30) and a bound for the centered empirical scalogram. The two following results provide the bounds.

Lemma 1. *Suppose that ψ satisfies (W-1)–(W-3) and let $d \in (1/2-\alpha, M]$. Then, there exists C only depending on d and ψ such that, for all $u \ge 0$,*

$$\int_{-\pi}^{\pi} |\mathbf{D}_{\infty,u}(\lambda)|^2\ \mathrm{d}\lambda \le C\ 2^{u(1/2-2d)}\ .$$

Proof. See Relation (72) in [7]. An alternative is to use that $\mathbf{\Lambda}_{0,u}(d) = \mathrm{Cov}\left(Q_0^{(d)}, Q_u^{(d)}\right)$ and thus the Cauchy-Schwarz Inequality yields $|\mathbf{\Lambda}_{0,u}(d)|^2 \le |\mathbf{\Lambda}_{0,0}(d)||\mathbf{\Lambda}_{u,u}(d)|$. Using (30), we get, setting $f^*(0) = 1$

$$4\pi 2^{4du} \int_{-\pi}^{\pi} |\mathbf{D}_{\infty,u}(\lambda)|^2\ \mathrm{d}\lambda \le 4\pi 2^{(2d+1/2)u} \int_{-\pi}^{\pi} |\mathbf{D}_{\infty,0}(\lambda)|^2\ \mathrm{d}\lambda\ .$$

The results follows from the fact that $|\mathbf{D}_{\infty,0}(\lambda)|$ is bounded for $d \in (1/2-\alpha, M]$ under (W-1)–(W-3), see Remark 1 in [5]. □

Lemma 2. *Let X be an $M(d)$ process with short-range spectral density f^* and suppose that Assumption 2 holds. Assume that (W-1)–(W-4) hold with Condition (26) on M and α. Then, there exists a positive constant C such that, for all $n \geq 1$ and $j \in \{0, 1, \ldots, J\}$,*

$$\mathbb{E}\left[\left|\hat{\sigma}_j^2 - \sigma_j^2\right|\right] \leq C\left\{2^{(1/2+2d)j}\, n^{-1/2} + 2^{j(1-2\alpha)}\right\} . \tag{49}$$

If moreover Condition (27) holds on f^*, one has

$$\mathrm{Var}\left(\hat{\sigma}_j^2\right) \leq C^2\, 2^{(1+4d)j}\, n^{-1} . \tag{50}$$

Proof. We use the same notations as in Theorem 2 to express $\hat{\sigma}_j^2$ in terms of a decimated linear process, but since here only one scale needs to be considered, we take $L = j = j'$ (hence $u = \ell = 0$ and $i = 1$). In this case (40) reads as

$$2^{-2jd}(\hat{\sigma}_j^2 - \sigma_j^2) = n_j^{-1} \sum_{k=0}^{n_j-1} \left\{Z_{1,j,k}^2 - \mathbb{E}[Z_{1,j,k}^2]\right\} , \tag{51}$$

where $Z_{1,j,k} = \sum_{t\in\mathbb{Z}} v_{1,j}(t)\xi_t$ with $v_{1,j}(t) = 2^{-jd}a_j(t)$. If (27) holds, then $v_{1,j}$ satisfies (18) with $\varepsilon = \pi$ and (19) (see the proof of Theorem 2) and, by Lemmas 5 and 6 in [10], we get

$$\sup_{j,n} \mathrm{Var}\left(n_j^{-1/2} \sum_{k=0}^{n_j-1} Z_{1,j,k}^2\right) < \infty . \tag{52}$$

Since $n_j \asymp n2^{-j}$ for $j \in \{0, 1, \ldots, J\}$, (50) follows.

If (27) does not hold, $v_{1,j}$ satisfies (18) for some $\varepsilon > 0$ which may no longer be taken equal to π (as a consequence of (44) in the proof of Theorem 2) and, applying [10, Proposition 4] with $\lambda_{1,j} = 0$, we get

$$2^{-2jd}(\hat{\sigma}_j^2 - \sigma_j^2) = n_j^{-1/2}\left[n_j^{-1/2} \sum_{k=0}^{n_j-1} \{\widehat{Z}_{1,j,k}^2 - \mathbb{E}[\widehat{Z}_{1,j,k}^2]\} + R_j\right] ,$$

where $\widehat{Z}_{1,j,k}$ satisfies (18) with $\varepsilon = \pi$ and (19) and hence (52) and R_j satisfies, for some positive constant C not depending on j,

$$\mathbb{E}\left[|R_j|\right] \leq C\, \left[n_j^{1/2} I_j + I_j^{1/2}\right] , \tag{53}$$

where

$$I_j \stackrel{\text{def}}{=} \int_0^\pi \mathbb{1}(|\lambda - \lambda_{1,\infty}| > \varepsilon)\, \left|v_{1,j}^*(\lambda)\right|^2\, \mathrm{d}\lambda . \tag{54}$$

Since f^* is always integrable away of the origin, the bound (45) implies (recall that here $L = j$)

$$\int_\varepsilon^\pi |v_{1,j}^*(\lambda)|^2\, \mathrm{d}\lambda \leq C\, 2^{(1-2\alpha-2d)j} .$$

Hence, we obtain, for some constant C not depending on j nor n,

$$\begin{aligned}\mathbb{E}\left[2^{-2jd}|\hat{\sigma}_j^2-\sigma_j^2|\right] &\le C\, n_j^{-1/2}\left[1+n_j^{1/2}2^{(1-2\alpha-2d)j}+2^{(1-2\alpha-2d)j/2}\right]\\ &\le C\, n_j^{-1/2}\left[2+n_j^{1/2}2^{(1-2\alpha-2d)j}\right]\ ,\end{aligned}$$

where we used $1-2\alpha-2d<0$ in Condition (26). Since $n_j \asymp n2^{-j}$ Relation (49) follows. □

We now prove the main result of this section.

Theorem 3. *Let $\{w_u(n), n, u \ge 0\}$ be an array of real numbers such that $w_u(n) \to w_u$ for all $u \ge 0$ as $n \to \infty$ and*

$$\lim_{\ell\to\infty}\sum_{u>\ell}\sup_{n\ge 0}|w_u(n)|2^{(1/2+2d)u}=0\ . \tag{55}$$

Then, under the assumptions of Theorem 2, as $n \to \infty$,

$$\sqrt{n2^{-L}}2^{-2Ld}\sum_{u=0}^{J-L}w_u(n)\{\hat{\sigma}_{L+u}^2-\sigma_{L+u}^2\}\xrightarrow{\mathcal{L}}\mathcal{N}\left(0,\sum_{u,u'\ge 0}w_u\mathbf{\Lambda}_{u,u'}(d)w_{u'}\right)\ , \tag{56}$$

where J is defined in (17).

Proof. We denote, for all $\ell \ge 0$,

$$S_{n,\ell}=\sqrt{n2^{-L}}2^{-2Ld}\sum_{u=0}^{\ell}w_u(n)\{\hat{\sigma}_{L+u}^2-\sigma_{L+u}^2\}$$

and

$$\tilde{S}_{n,\ell}=\sqrt{n2^{-L}}2^{-2Ld}\sum_{u=0}^{\ell}w_u\{\hat{\sigma}_{L+u}^2-\sigma_{L+u}^2\}\ .$$

Theorem 2 then gives that, for any $\ell \ge 0$, as $n \to \infty$,

$$\tilde{S}_{n,\ell}\xrightarrow{\mathcal{L}}\mathcal{N}\left(0,\sum_{0\le u,u'\le\ell}w_u\mathbf{\Lambda}_{u,u'}(d)w_{u'}\right)\ .$$

Note that (55) implies $\sum_{u>\ell}|w_u|2^{(1/2+2d)u}\to 0$ as $\ell\to\infty$, hence, using Lemma 1 and (30), we have

$$\begin{aligned}\sum_{\ell<u,u'}\left|w_u\mathbf{\Lambda}_{u,u'}(d)w_u\right| &\le C\sum_{\ell<u\le u'}|w_uw_u|\,2^{4du'+u}2^{(u'-u)(1/2-2d)}\\ &\le C\sum_{\ell<u}|w_u|2^{(1/2+2d)u}\times\sum_{\ell<u'}|w_{u'}|2^{(1/2+2d)u'}\\ &\to 0\quad\text{as }\ell\to\infty.\end{aligned}$$

The left-hand side of (56) is $S_{n,J-L}$. We decompose it as

$$S_{n,J-L}=[S_{n,J-L}-S_{n,\ell}]+\left[S_{n,\ell}-\tilde{S}_{n,\ell}\right]+\tilde{S}_{n,\ell}$$

From the last 3 displays and applying [1, Theorem 3.2], it is sufficient to prove that

$$\lim_{\ell\to\infty}\limsup_{n\to\infty}\mathbb{E}\left[|S_{n,J-L}-S_{n,\ell}|+\left|S_{n,\ell}-\tilde{S}_{n,\ell}\right|\right]=0\,. \tag{57}$$

To obtain this limit, we need to separate the case where Condition (27) holds from the one where it is replaced by Condition (28). Under Condition (27), we apply (50); under (28), we apply (49). Let us for instance check the second case (the first one is similar, although simpler). The bound (49) implies

$$\mathbb{E}\left[\left|S_{n,\ell}-\tilde{S}_{n,\ell}\right|\right]\le C\sum_{u=0}^{\ell}|w_u-w_u(n)|(2^{(1/2+2d)u}+2^{L(1-2\alpha-2d)}\sqrt{n2^{-L}}\,2^{(1-2\alpha)u})\,,$$

which, using $w_u\to w_u(n)$ and (28), tends to 0 as $n\to\infty$ for all $\ell\ge 0$, and

$$\begin{aligned}\mathbb{E}\left[|S_{n,J-L}-S_{n,\ell}|\right]&\\ &\le C\left[\sum_{u>\ell}|w_u(n)|2^{(1/2+2d)u}+2^{L(1-2\alpha-2d)}\sqrt{n2^{-L}}\sum_{u>\ell}|w_u(n)|2^{u(1-2\alpha)}\right]\\ &\le C\left[1+2^{L(1-2\alpha-2d)}\sqrt{n2^{-L}}\right]\sum_{u>\ell}|w_u(n)|2^{(1/2+2d)u}\,,\end{aligned}$$

(since $1-2\alpha<2d$ in Condition (26)) which tends to 0 as $n\to\infty$ followed by $\ell\to\infty$ by (28) and (55). This yields (57), which achieves the proof. □

4. The log-regression estimation of the memory parameter

The wavelet-based regression estimator of the memory parameter d involves regressing the scale spectrum estimator $\hat{\sigma}_j^2$, defined in (15), with respect to the scale index j. More precisely, an estimator of the memory parameter d is obtained by regressing the logarithm of the empirical variance $\log(\hat{\sigma}_i^2)$ for a finite number of scale indices $j\in\{L,\dots,L+\ell\}$ where $L=L(n)\ge 0$ is the lower scale and $1+\ell\ge 2$ is the number of scales used in the regression. For a sample size equal to n, this estimator is well defined for L and ℓ such that $\ell\ge 1$ and

$$L+\ell\le[\log_2(n-\mathrm{T}+1)-\log_2(\mathrm{T})]\,, \tag{58}$$

where the right-hand side of this inequality is the maximal index j such that $n_j\ge 1$. The regression estimator can be expressed formally as

$$\hat{d}_n(L,\mathbf{w})\stackrel{\text{def}}{=}\sum_{j=L}^{L+\ell}w_{j-L}\log\left(\hat{\sigma}_j^2\right)\,, \tag{59}$$

where the vector $\mathbf{w}\stackrel{\text{def}}{=}[w_0,\dots,w_\ell]^T$ of weights satisfies

$$\sum_{i=0}^{\ell}w_i=0\quad\text{and}\quad 2\log(2)\sum_{i=0}^{\ell}iw_i=1\,. \tag{60}$$

One may choose, for example, $\mathbf{w}$ corresponding to the weighted least-squares regression vector, defined by

$$\mathbf{w} = DB(B^T DB)^{-1}\mathbf{b}\ ,$$

where $B \stackrel{\text{def}}{=} \begin{bmatrix} 1 & 1 & \dots & 1 \\ 0 & 2 & \dots & \ell \end{bmatrix}^T$ is the so-called design matrix, D is a definite positive matrix and

$$\mathbf{b} \stackrel{\text{def}}{=} [0\ (2\log(2))^{-1}]^T. \tag{61}$$

Ordinary least square regression corresponds to the case where D is the identity matrix.

In [6], the process X was assumed Gaussian and a bound for the mean square error and an asymptotic equivalent to the variance of $\hat{d}_n(L, \mathbf{w})$ were obtained. The asymptotic normality is established in [5], also under the Gaussian assumption. Here we show that the asymptotic normality holds under the weaker linear assumption.

Theorem 4. *Let X be an $M(d)$ process with short-range spectral density f^* and suppose that Assumptions 1 and 2 hold. Under (W-1)–(W-4)with*

$$(1+\beta)/2 - \alpha < d \le M \quad \textit{and} \quad K \le M\ , \tag{62}$$

if, as $n \to \infty$, $L(n)$ is such that

$$(n2^{-L(n)})^{-1} + n2^{-(1+2\beta)L(n)} \to 0\ , \tag{63}$$

then one has the following central limit:

$$\sqrt{n2^{-L(n)}}\left(\hat{d}_n(L, \mathbf{w}) - d\right) \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \mathbf{w}^T \mathbf{V}(d)\mathbf{w}\right)\ , \tag{64}$$

where

$$\mathrm{K}(d) \stackrel{\text{def}}{=} \int_{-\infty}^{\infty} |\xi|^{-2d}\, |\hat{\psi}(\xi)|^2\, \mathrm{d}\xi\ , \tag{65}$$

and $\mathbf{V}(d, \psi)$ is the $(1+\ell)\times(1+\ell)$ matrix defined as

$$\mathbf{V}_{i,j}(d, \psi) \stackrel{\text{def}}{=} \frac{4\pi 2^{2d|j-i|} 2^{i\wedge j}}{\mathrm{K}(d)^2} \int_{-\pi}^{\pi} \left|\mathbf{D}_{\infty,|j-i|}(\lambda)\right|^2\, \mathrm{d}\lambda \qquad 0 \le i, j \le \ell\ , \tag{66}$$

Proof. The only assumptions of Theorem 2 that are not included in our set of assumptions here are (27) and (28). The former is verified if $\varepsilon = \pi$ in Assumption 1. If $\varepsilon < \pi$, (28) holds as a consequence of (62) and (63). Hence Theorem 2 applies (note that (62) implies (26) since $\beta > 0$). Applying [7, Theorem 1], under (W-1)–(W-4), we have the following approximation:

$$\left|\sigma_j^2 - f^*(0)\,\mathrm{K}(d)\,2^{2jd}\right| \le C\, f^*(0)\, L\, 2^{(2d-\beta)j} \tag{67}$$

where σ_j^2 is defined in (14) and $\mathrm{K}(d)$ in (65). This, with Theorem 2, [5, Relation (39)] and [5, Proposition 3], gives the result. □

Remark 4. The centering in (29) in Theorem 2 involved the expected value whereas the centering in (64) in Theorem 4 does not. In order to deal with the corresponding bias, Condition (26) is strengthened by (62).

5. The wavelet Whittle estimation of the memory parameter

We now consider the semi-parametric estimator introduced in [7]. As the log-regression estimation, this estimator is also based on the scalogram but is defined as the maximizer of a Whittle type contrast function (see [7, Eq. (20)]),

$$\tilde{d}_n(L,U) \stackrel{\text{def}}{=} \underset{d'\in\mathbb{R}}{\text{Argmin}} \left[\log\left(\sum_{j=L}^{U} 2^{-2d'j} n_j \hat{\sigma}_j^2 \right) + 2d' \log(2) \mathcal{J} \right] \quad \text{with} \quad \mathcal{J} \stackrel{\text{def}}{=} \frac{\sum_{j=L}^{U} j\, n_j}{\sum_{j=L}^{U} n_j} .$$

$\tilde{d}_n$, which involves the scales $L \le j \le U$, is a wavelet analog of the local Whittle Fourier estimator studied in [9] (often referred to as *semiparametric Gaussian estimator*) and is therefore called the *local Whittle wavelet* estimator. To prove the asymptotic normality of $\tilde{d}_n$, we will use Theorem 3.

We denote, for all integer $\ell \ge 1$,

$$\eta_\ell \stackrel{\text{def}}{=} \sum_{j=0}^{\ell} j \frac{2^{-j}}{2-2^{-\ell}} \quad \text{and} \quad \kappa_\ell \stackrel{\text{def}}{=} \sum_{j=0}^{\ell} (j-\eta_\ell)^2 \frac{2^{-j}}{2-2^{-\ell}} , \tag{68}$$

$$\rho^2(d,\ell) \stackrel{\text{def}}{=} \frac{\pi}{(2-2^{-\ell})\kappa_\ell (\log(2)\mathrm{K}(d))^2} \times \left\{ \mathrm{I}_0(d) + \frac{2}{\kappa_\ell} \sum_{u=1}^{\ell} \mathrm{I}_u(d)\, 2^{(2d-1)u} \sum_{i=0}^{\ell-u} \frac{2^{-i}}{2-2^{-\ell}} (i-\eta_\ell)(i+u-\eta_\ell) \right\} , \tag{69}$$

$$\rho^2(d,\infty) \stackrel{\text{def}}{=} \frac{\pi}{[2\log(2)\mathrm{K}(d)]^2} \left\{ \mathrm{I}_0(d) + 2\sum_{u=1}^{\infty} \mathrm{I}_u(d)\, 2^{(2d-1)u} \right\} , \tag{70}$$

where $\mathrm{K}(d)$ is defined in (65).

Theorem 5. *Let X be an $M(d)$ process with short-range spectral density f^* and suppose that Assumptions 1 and 2 hold. Under (W-1)–(W-4) with Condition (62) on α and M, if, as $n \to \infty$, the lower scale $L(n)$ is such that*

$$L(n)(n2^{-L(n)})^{-1/8} + n2^{-(1+2\beta)L(n)} \to 0 , \tag{71}$$

and the upper scale $U(n)$ is such that

$$U(n) - L(n) \to \ell \in \{1, 2 \ldots, \infty\},$$

then one has the following central limit:

$$\sqrt{n2^{-L(n)}} \left(\tilde{d}_n(L(n),U(n)) - d \right) \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \rho^2(d,\ell)\right) , \tag{72}$$

where $\rho^2(d,\ell)$ for $\ell < \infty$ and $\ell = \infty$ are defined in (69) and (70) respectively.

Remark 5. Condition (71) is similar to (63) but $n2^{-L(n)} \to \infty$ is replaced by the stronger condition $L(n)(n2^{-L(n)})^{-1/8} \to 0$, which holds for example, if $n2^{-L(n)}/n^\gamma$ has a positive limit for some $\gamma > 0$ as usually verified.

Proof. Assume $f^*(0) = 1$ without loss of generality. The proof is the same as that of [7, Theorem 5] until Eq. (66),

$$(n2^{-L})^{1/2}\,(\hat{d}_n - d) = \frac{(n2^{-L})^{-1/2}\,\widehat{\mathbf{S}}_n}{2\,\log(2)\,\mathrm{K}(d)\,(2-2^{-(U-L)})\,\kappa_{U-L}}\,(1+o_{\mathbb{P}}(1))\;, \tag{73}$$

where

$$\widehat{\mathbf{S}}_n \stackrel{\text{def}}{=} \sum_{j=L}^{U} [j-\mathcal{J}]\,2^{-2jd}\,n_j\hat{\sigma}_j^2\;.$$

Thus, using $\sum_{j=L}^{U}[j-\mathcal{J}]\,n_j = 0$ and (14), we get

$$\mathbb{E}\left[\widehat{\mathbf{S}}_n\right] = \sum_{j=L}^{U}[j-\mathcal{J}]\,n_j\,\left(2^{-2jd}\,\sigma_j^2 - f^*(0)\,\mathrm{K}(d)\right) = O(n2^{-(1+\beta)L}) = o\left((n2^{-L})^{1/2}\right)\;,$$

where the O-term follows from (67), the fact that $L < \mathcal{J} < L+1$ (see [7, Eq (61)]) and $n_j \le n2^{-j}$ and the o-term follows from (71).

Hence it only remains to establish a CLT for $\widehat{\mathbf{S}}_n$ similar to that of [7, Proposition 10] but under the assumptions of Theorem 5. This is obtained by observing that

$$(n2^{-L})^{-1/2}\,\widehat{\mathbf{S}}_n = (n2^{-L})^{1/2}\,2^{-2Ld}\sum_{u=0}^{U-L} w_u(n)\hat{\sigma}_j^2$$

with $u = j - L$ and

$$w_u(n) \stackrel{\text{def}}{=} [u-(\mathcal{J}-L)]\,2^{-2ud}\,\frac{n_{L+u}}{n2^{-L}}, \quad u \in \{0,\dots,U-L\}\;,$$

which satisfies

$$\sup_n |w_u(n)| \le C\;u2^{-(2d+1)u}, \quad u \in \{0,\dots,U-L\}\;,$$

and by applying Theorem 3. □

Appendix A. Wavelet coefficients linear filters

Assumption (W-1) implies that $\hat{\phi}$ and $\hat{\psi}$ are everywhere infinitely differentiable. When (W-1) holds, Assumptions (W-3) and (W-4) can be expressed in different ways. (W-3) is equivalent to asserting that the first $M-1$ derivative of $\hat{\psi}$ vanish at the origin and hence

$$|\hat{\psi}(\lambda)| = O(|\lambda|^M) \quad \text{as} \quad \lambda \to 0. \tag{74}$$

And, by [2, Theorem 2.8.1, Page 90], (W-4) is equivalent to

$$\sup_{k\neq 0} |\hat{\phi}(\lambda + 2k\pi)| = O(|\lambda|^M) \quad \text{as} \quad \lambda \to 0. \tag{75}$$

Many authors suppose that the $\psi_{j,k}$ are orthogonal and even that they are generated by a multiresolution analysis (MRA). Assumptions (W-1)–(W-4) in Section 3 are satisfied in these cases, ϕ being the scaling function and ψ is the associated wavelet. In this paper, however, we do not assume that wavelets are orthonormal nor that they are associated to a multiresolution analysis. We may therefore work with other convenient choices for ϕ and ψ as long as (W-1)–(W-4) are satisfied. A simple example is to set, for some positive integer N,

$$\phi(x) \stackrel{\text{def}}{=} \mathbb{1}_{[0,1]}^{\otimes N}(x) \quad \text{and} \quad \psi(x) \stackrel{\text{def}}{=} C_N \, \frac{\mathrm{d}^N}{\mathrm{d}x^N} \mathbb{1}_{[0,1]}^{\otimes 2N}(x),$$

where $\mathbb{1}_A$ is the indicator function of the set A, $f^{\otimes N}$ denotes the N-th self-convolution of a function f and C_N is a normalizing constant such that $\int_{-\infty}^{\infty} \psi^2(x)\mathrm{d}x = 1$. It follows that

$$|\hat{\phi}(\xi)| = |2\sin(\xi/2)/\xi|^N \quad \text{and} \quad |\hat{\psi}(\xi)| = C_N \, |\xi|^N |2\sin(\xi/2)/\xi|^{2N}.$$

Using (74) and (75), one easily checks that (W-1)–(W-4) are satisfied with M and α equal to N. Of course the family of functions $\{\psi_{j,k}\}$ are not orthonormal for this choice of the wavelet function ψ (and the function ϕ is not associated to a MRA).

References

[1] P. Billingsley. *Convergence of probability measures*. Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons Inc., New York, second edition, 1999. A Wiley-Interscience Publication.

[2] A. Cohen. *Numerical analysis of wavelet methods*, volume 32 of *Studies in Mathematics and its Applications*. North-Holland Publishing Co., Amsterdam, 2003.

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

[4] E. Moulines, F. Roueff, and M. S. Taqqu. Central Limit Theorem for the log-regression wavelet estimation of the memory parameter in the Gaussian semi-parametric context. To appear in Fractals, 2006.

[5] E. Moulines, F. Roueff, and M.S. Taqqu. Central Limit Theorem for the log-regression wavelet estimation of the memory parameter in the Gaussian semi-parametric context. To appear, 2007.

[6] E. Moulines, F. Roueff, and M.S. Taqqu. On the spectral density of the wavelet coefficients of long memory time series with application to the log-regression estimation of the memory parameter. *J. Time Ser. Anal.*, 28(2), 2007. [DOI].

[7] E. Moulines, F. Roueff, and M.S. Taqqu. A wavelet Whittle estimator of the memory parameter of a non-stationary Gaussian time series. *Ann. Statist.*, 2007. To appear, [arXiv].

[8] E. Moulines and P. Soulier. Long-range dependence: Theory and applications. In P. Doukhan, G. Oppenheim, and M. Taqqu, editors, *Theory and applications of long-range dependence*. Birkhaüser, Boston, 2002.

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

[10] F. Roueff and M. S. Taqqu. Central limit theorems for arrays of decimated linear processes. Technical report, HAL/arxiv, 2008.

TELECOM ParisTech, INSTITUT Télécom, CNRS LTCI, 46, rue Barrault, 75634 Paris Cédex 13, France.

E-mail address: roueff@tsi.enst.fr

Department of Mathematics and Statistics, Boston University Boston, MA 02215, USA.

E-mail address: murad@math.bu.edu