Electronic Journal of Statistics
Vol. 9 (2015) 153–203
ISSN: 1935-7524
DOI: 10.1214/15-EJS987

# Large scale reduction principle and application to hypothesis testing

**Marianne Clausel**

*Laboratoire Jean Kuntzmann, Université de Grenoble, CNRS*
*F38041 Grenoble Cedex 9*
*e-mail:* marianne.clausel@imag.fr

**François Roueff**

*Institut Mines–Telecom, Telecom ParisTech, CNRS LTCI, 46 rue Barrault*
*75634 Paris Cedex 13, France*
*e-mail:* roueff@telecom-paristech.fr

**and**

**Murad S. Taqqu**

*Departement of Mathematics and Statistics, Boston University, Boston, MA 02215, USA*
*e-mail:* murad@bu.edu

**Abstract:** Consider a non–linear function $G(X_t)$ where $X_t$ is a stationary Gaussian sequence with long–range dependence. The usual reduction principle states that the partial sums of $G(X_t)$ behave asymptotically like the partial sums of the first term in the expansion of $G$ in Hermite polynomials. In the context of the wavelet estimation of the long–range dependence parameter, one replaces the partial sums of $G(X_t)$ by the wavelet scalogram, namely the partial sum of squares of the wavelet coefficients. Is there a reduction principle in the wavelet setting, namely is the asymptotic behavior of the scalogram for $G(X_t)$ the same as that for the first term in the expansion of $G$ in Hermite polynomial? The answer is negative in general. This paper provides a minimal growth condition on the scales of the wavelet coefficients which ensures that the reduction principle also holds for the scalogram. The results are applied to testing the hypothesis that the long-range dependence parameter takes a specific value.

**MSC 2010 subject classifications:** Primary 42C40, 60G18, 62M15; Secondary 60G20, 60G22.
**Keywords and phrases:** Long-range dependence, long memory, self-similarity, wavelet transform, estimation, hypothesis testing.

Received July 2014.

## Contents

## 1. Introduction

Let $X = \{X_t\}_{t\in\mathbb{Z}}$ be a centered stationary Gaussian process with unit variance and spectral density $f(\lambda), \lambda \in (-\pi, \pi)$. Such a stochastic process is said to have *short memory* or *short–range dependence* if $f(\lambda)$ is bounded around $\lambda = 0$ and *long memory* or *long–range dependence* if $f(\lambda) \to \infty$ as $\lambda \to 0$. We will suppose that $\{X_t\}_{t\in\mathbb{Z}}$ has long memory with memory parameter $0 < d < 1/2$, that is,

$$f(\lambda) \sim |\lambda|^{-2d} f^*(\lambda) \text{ as } \lambda \to 0 \tag{1.1}$$

where the *short range part* $f^*$ of the spectral density is a bounded spectral density which is continuous and positive at the origin. The parameter $d$ is also called the long-range dependence parameter.

A standard assumption in the semi-parametric setup is

$$|f^*(\lambda) - f^*(0)| \leq C f^*(0)\, |\lambda|^{\beta} \quad \lambda \in (-\pi, \pi), \tag{1.2}$$

where $\beta$ is some smoothness exponent in $(0,2]$. This hypothesis is semi–parametric in nature because the function $f^*$ plays the role of a "nuisance function". It is convenient to set

$$f(\lambda) = |1-\mathrm{e}^{-\mathrm{i}\lambda}|^{-2d} f^*(\lambda), \quad \lambda \in (-\pi, \pi]. \tag{1.3}$$

Consider now a process $\{Y_t\}_{t\in\mathbb{Z}}$, such that

$$\left(\Delta^K Y\right)_t = G(X_t), \quad t \in \mathbb{Z}, \tag{1.4}$$

for $K \geq 0$, where $(\Delta Y)_t = Y_t - Y_{t-1}$, $\{X_t\}_{t\in\mathbb{Z}}$ is Gaussian with spectral density $f$ satisfying (1.3) and where $G$ is a function such that $\mathbb{E}[G(X_t)] = 0$ and $\mathbb{E}[G(X_t)^2] < \infty$. While the process $\{Y_t\}_{t\in\mathbb{Z}}$ is not necessarily stationary, its $K$-th difference $\Delta^K Y_t$ is stationary. Nevertheless, as in [31] one can speak of the "generalized spectral density" of $\{Y_t\}_{t\in\mathbb{Z}}$, which we denote $f_{G,K}$. It is defined as

$$f_{G,K}(\lambda) = |1-e^{-i\lambda}|^{-2K} f_G(\lambda), \tag{1.5}$$

where $f_G$ is the spectral density of $\{G(X_t)\}_{t\in\mathbb{Z}}$.

Note that $G(X_t)$ is the output of a non–linear filter $G$ with Gaussian input. According to the Hermite expansion of $G$ and the value $d$, the time series $Y$ may be long–range dependent (see [9] for more details). We aim at developing efficient estimators of the memory parameter of such non–linear time series.

Since the 80's many methods for the estimation of the memory parameter have been developed. Let us cite the Fourier methods developed by Fox and Taqqu [18] and Robinson [23, 22]. Since the 90's, wavelet methods have become very popular. The idea of using wavelets to estimate the memory parameter of a time series goes back to [30] and [13, 14, 15, 17]. See also [2, 3, 6, 4, 7]. As shown in [16, 2, 29] and [5] in a parametric context, the memory parameter of a time series can be estimated using the normalized limit of its *scalogram* (2.15), that is the average of squares of its wavelet coefficients computed at a given scale. It is well–known that, when considering Gaussian or linear time series, the wavelet–based estimator of the memory parameter is consistent and asymptotically Gaussian (see [20] for a general framework in the Gaussian case and [25] for the linear case). This result is particulary important for statistical purpose since it provides confidence intervals for the wavelet–based estimator of the memory parameter.

The application of wavelet–based methods for the estimation of the memory parameter of non-Gaussian stochastic processes has been much less treated in the literature. See [1] for some empirical studies. In [8] is considered the case of the Rosenblatt process which is a non-Gaussian self-similar process with stationary increments living in the second Wiener chaos, that is, it can be expressed as a double iterated integral with respect to the Wiener process. In this case, the wavelet–based estimator of the memory parameter is consistent but satisfies a non–central limit theorem. More precisely, conveniently renormalized, the scalogram which is a sum of squares of wavelet coefficients converges to a Rosenblatt variable and thus admits a non–Gaussian limit. This result, surprisingly, also holds for a time series of the form $H_{q_0}(X_t)$ where $X_t$ is Gaussian with unit variance and $H_{q_0}$ denotes the $q_0$-th Hermite polynomial with $q_0 \geq 2$ (see [11]).

The general case $G(X_t)$ is expected to derive from the case $G = H_{q_0}$. Namely, one could expect that some "reduction theorem" analog to the one of [26] holds. Recall that the classical reduction theorem of [26] states that if $G(X)$ is long–range dependent then the limit in the sense of finite–dimensional distributions of $\sum_{k=1}^{[nt]} G(X_k)$ adequately normalized, depends only on the first term $c_{q_0} H_{q_0}/q_0!$ in the Hermite expansion of $G$. The reduction principle then states that there exist normalization factors $a_n \to \infty$ as $n \to \infty$ such that

$$\frac{1}{a_n}\sum_{k=1}^{[nt]} G(X_k) \qquad \text{and} \qquad \frac{1}{a_n}\sum_{k=1}^{[nt]} \frac{c_{q_0}}{q_0!} H_{q_0}(X_k),$$

have the same non–degenerate limit as $n \to \infty$. A reduction principle was established in [9], Theorem 5.1 for the *wavelet coefficients* of a non–linear time series of the form $G(X_t)$. In applications, the wavelet coefficients are not used directly but only through the scalogram. For example, [12] use the scalogram to compare Fourier and wavelet estimation methods of the memory parameter. The difficulty is that the scalogram is a quadratic function of the wavelet coefficients involving not only the number of observations but also the scale at which the wavelet coefficients are computed. In practice, however, the scalogram is easy to obtain and one can take advantage of the structure of sample moments to investigate statistical properties. Its use is well–illustrated numerically in [1] who consider a number of statistical applications.

The following is a natural question:

Does a reduction principle hold for the scalogram?

In [10] we illustrated through different large classes of examples, that the reduction principle for the scalogram does not necessarily hold and that the asymptotic limit of the scalogram may even be Hermite process of order greater than 2. It is then important to find sufficient conditions for the reduction principle to hold. In this case, the normalized limit of the scalogram of the time series $G(X_t)$ would be the same as the time series $c_{q_0} H_{q_0}(X)/q_0!$ studied in [11] and therefore will be asymptotically Gaussian if $q_0 = 1$ and a Rosenblatt random variable if $q_0 \geq 2$. In Theorem 3.2, we prove that the reduction principle holds *at large scales*, namely if

$$n_j \ll \gamma_j^{\nu_c} \text{ as } j \to \infty, \tag{1.6}$$

that is, if the number of wavelet coefficients $n_j$ at scale $j$ (typically $N2^{-j}$, where $N$ is the sample size) does not grow as fast as the scale factor $\gamma_j$ (typically $2^j$) to the power $\nu_c$, as the sample size $N$ and the scale index $j$ go to infinity. The critical exponent $\nu_c$ depends on the function $G$ under consideration and may take the value $\nu_c = \infty$ for some functions, in which case the reduction principle holds without any particular growth condition on $\gamma_j$ and $n_j$ besides $n_j \to \infty$ and $\gamma_j \to \infty$ as $j \to \infty$.

The paper is organized as follows. In Section 2, we introduce long–range dependence and the scalogram. The main Theorem 3.2, which states that under Condition (1.6) the reduction principle holds is stated in Section 3 with the critical exponent $\nu_c$ given in Section 4 and examples provided in Section 5. Section 6 contains statistical applications. Numerical experiments can be found in Section 7. The decomposition of the scalogram in Wiener chaos is described in Section 8. That section contains Theorem 8.2 on which Theorem 3.2 is based. Several proofs are in Section 9. Section 10 contains technical lemmas. The integral representations are described in Appendix A and the wavelet filters are given in Appendix B. Appendix C depicts the multiscale wavelet inference setting.

For the convenience of the reader, in addition to providing a formal proof of a given result, we sometimes describe in a few lines the idea behind the proof.

## 2. Long–range dependence and the multidimensional wavelet scalogram

The centered Gaussian sequence $X = \{X_t\}_{t\in\mathbb{Z}}$ with unit variance and spectral density (1.3) is long–range dependent because $d > 0$ and hence its spectrum explodes at $\lambda = 0$.

The long–memory behavior of a time series $Y$ of the form (1.4) is well–known to depend on the expansion of $G$ in Hermite series. Recall that if $\mathbb{E}[G(X_0)] = 0$ and $\mathbb{E}[G(X_0)^2] < \infty$ for $X_0 \sim \mathcal{N}(0,1)$, $G(X)$ can be expanded in Hermite polynomials, that is,

$$G(X) = \sum_{q=1}^{\infty} \frac{c_q}{q!} H_q(X). \tag{2.1}$$

One sometimes refer to (2.1) as an expansion in Wiener chaos. The convergence of the infinite sum (2.1) is in $L^2(\Omega)$,

$$c_q = \mathbb{E}[G(X)H_q(X)], \quad q \geq 1, \tag{2.2}$$

and

$$H_q(x) = (-1)^q e^{\frac{x^2}{2}} \frac{d^q}{dx^q}\left(e^{-\frac{x^2}{2}}\right),$$

are the Hermite polynomials. These Hermite polynomials satisfy $H_0(x) = 1$, $H_1(x) = x, H_2(x) = x^2 - 1$ and one has

$$\mathbb{E}[H_q(X)H_{q'}(X)] = \int_{\mathbb{R}} H_q(x)H_{q'}(x)\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-x^2/2}\mathrm{d}x = q!\mathbb{1}_{\{q=q'\}}.$$

Observe that the expansion (2.1) starts at $q = 1$, since

$$c_0 = \mathbb{E}[G(X)H_0(X)] = \mathbb{E}[G(X)] = 0, \tag{2.3}$$

by assumption. Denote by $q_0 \geq 1$ the *Hermite rank* of $G$, namely the index of the first non–zero coefficient in the expansion (2.1). Formally,

$$q_0 = \min\{q \geq 1,\, c_q \neq 0\}. \tag{2.4}$$

One has then

$$\sum_{q=q_0}^{+\infty} \frac{c_q^2}{q!} = \mathbb{E}[G(X)^2] < \infty. \tag{2.5}$$

In the special case where $G = H_q$, whether $\{H_q(X_t)\}_{t\in\mathbb{Z}}$ is also long–range dependent depends on the respective values of $q$ and $d$. We show in [9], that the spectral density of $\{H_q(X_t)\}_{t\in\mathbb{Z}}$ behaves like $|\lambda|^{-2\delta_+(q)}$ as $\lambda \to 0$, where

$$\delta_+(q) = \max(\delta(q), 0) \quad \text{where} \quad \delta(q) = qd - (q-1)/2. \tag{2.6}$$

We will also let $\delta_+(0) = \delta(0) = 1/2$. For $q \geq 1$, $\delta_+(q)$ is the memory parameter of $\{H_q(X_t)\}_{t\in\mathbb{Z}}$. It is a non–increasing function of $q$. Therefore, since $0 < d < 1/2$, $\{H_q(X_t)\}_{t\in\mathbb{Z}}$, $q \geq 1$, is long–range dependent[1] if and only if

$$\delta(q) > 0 \Longleftrightarrow d > \frac{1}{2}(1 - 1/q), \tag{2.7}$$

that is, $d$ must be sufficiently close to 1/2. Specifically, for long–range dependence,

$$q = 1 \Rightarrow d > 0, \quad q = 2 \Rightarrow d > 1/4, \quad q = 3 \Rightarrow d > 1/3, \quad q = 4 \Rightarrow d > 3/8. \tag{2.8}$$

From another perspective,

$$\delta(q) > 0 \Longleftrightarrow 1 \leq q < 1/(1 - 2d), \tag{2.9}$$

and thus $\{H_q(X_t)\}_{t\in\mathbb{Z}}$ is short–range dependent if $q \geq 1/(1 - 2d)$.

Recall that the Hermite rank of $G$ is $q_0 \geq 1$, that is the expansion of $G(X_t)$ starts at $q_0$. We always assume that $\{H_{q_0}(X_t)\}_{t\in\mathbb{Z}}$ has long memory, that is,

$$q_0 < 1/(1 - 2d). \tag{2.10}$$

The condition (2.10), with $q_0$ defined as the Hermite rank (2.4), ensures such that $\{Y_t\}_{t\in\mathbb{Z}} = \{\Delta^{-K} G(X_t)\}_{t\in\mathbb{Z}}$ is long-range dependent with long memory parameter

$$d_0 = K + \delta(q_0) \in (K, K + 1/2). \tag{2.11}$$

More precisely, we have the following result which also determines a Hölder condition on the short-range part of the spectral density. This condition shall involve $q_0$ defined in (2.11), and, if $G$ is not reduced to $c_{q_0} H_{q_0}/(q_0!)$, it also involves the index of the second non-vanishing Hermite coefficient denoted by

$$q_1 = \inf\{q > q_0 : c_q \neq 0\}.$$

If there is no such $q_1$ we let $\delta_+(q_1) = 0$ in (2.12).

[1] In our context, the values $d = 1/2 - 1/(2q)$, $q \geq 1$, constitute boundary values which introduce logarithmic terms and will be omitted for simplicity. See Remark 3.3.

**Theorem 2.1.** *Let $Y$ be defined as above. Then the generalized spectral density $f_{G,K}$ of $Y$ can be written as*

$$f_{G,K}(\lambda) = |1 - \mathrm{e}^{-\mathrm{i}\lambda}|^{-2d_0} f_G^*(\lambda),$$

*where $d_0$ is defined by (2.11) and $f_G^*$ is bounded, continuous and positive at the origin. Moreover, for any $\zeta > 0$ satisfying*

$$\zeta \leq \min(\beta, 2(\delta(q_0) - \delta_+(q_1)) \quad \textit{and, if } q_0 \geq 2, \quad \zeta < 2\delta(q_0), \tag{2.12}$$

*there exists a constant $C > 0$ such that*

$$|f_G^*(\lambda) - f_G^*(0)| \leq C f_G^*(0)\, |\lambda|^{\zeta}, \quad \lambda \in (-\pi, \pi). \tag{2.13}$$

*Proof.* See Section 9.1. □

*Idea behind the proof of Theorem 2.1.* Starting with the regularity of the nuisance function $f^*$ in (1.1), one derives that of $f_{H_q}^*$ and, more generally, that of $f_G^*$, taking advantage of the fact that the terms in the expansion of $G(X)$ in Hermite polynomials are uncorrelated.

**Remark 2.1.** The exponent $\zeta$ in (2.12) will affect the bias of the mean of the scalogram (see (6.6)). The higher $\zeta$, the lower the bias. Since in (2.12), $\zeta$ is required to satisfy a non–strict and a strict inequality (if $q_0 \geq 2$), we cannot provide an explicit expression for $\zeta$. However, in most cases one has $q_0 = 1$ or $\delta_+(q_1) > 0$ and hence one can set $\zeta = \min(\beta, 2(\delta(q_0) - \delta_+(q_1)))$ which then satisfies both inequalities in (2.12).

Our estimator of the long memory parameter of $Y$ is defined from its wavelet coefficients, denoted by $\{W_{j,k},\ j \geq 0,\ k \in \mathbb{Z}\}$, where $j$ indicates the scale index and $k$ the location. These wavelet coefficients are defined by

$$W_{j,k} = \sum_{t \in \mathbb{Z}} h_j(\gamma_j k - t) Y_t, \tag{2.14}$$

where $\gamma_j \uparrow \infty$ as $j \uparrow \infty$ is a sequence of non–negative decimation factors applied at scale index $j$. The properties of the memory parameter estimator are directly related to the asymptotic behavior of the scalogram $S_{n_j,j}$, defined by

$$S_{n_j,j} = \frac{1}{n_j} \sum_{k=0}^{n_j - 1} W_{j,k}^2, \tag{2.15}$$

as $n_j \to \infty$ (large sample behavior) and $j \to \infty$ (large scale behavior). More precisely, we will study the asymptotic behavior of the sequence

$$\overline{S}_{n_{j+u},j+u} = S_{n_{j+u},j+u} - \mathbb{E}(S_{n_{j+u},j+u}) = \frac{1}{n_{j+u}} \sum_{k=0}^{n_{j+u}-1} \left(W_{j+u,k}^2 - \mathbb{E}(W_{j+u,k}^2)\right), \tag{2.16}$$

adequately normalized as $j, n_j \to \infty$.

There are two perspectives. One can consider, as in [9], that the wavelet coefficients $W_{j+u,k}$ are processes indexed by $u$ taking a finite number of values. A second perspective consists in replacing the filter $h_j$ in (2.14) by a multidimensional filter $h_{\ell,j}, \ell = 1, \ldots, m$ and thus replacing $W_{j,k}$ in (2.14) by

$$W_{\ell,j,k} = \sum_{t\in\mathbb{Z}} h_{\ell,j}(\gamma_j k - t) Y_t,\ \ell = 1, \ldots, m,$$

(see Appendix C for more details). We adopted this second perspective in [11, 10] and we also adopt it here since it allows us to compare our results to those obtained in [25] in the Gaussian case.

We use bold faced symbols $\mathbf{W}_{j,k}$ and $\mathbf{h}_j$ to emphasize the multivariate setting and let

$$\mathbf{h}_j = \{h_{\ell,j},\ \ell = 1, \ldots, m\}, \qquad \mathbf{W}_{j,k} = \{W_{\ell,j,k},\ \ell = 1, \ldots, m\},$$

with

$$\mathbf{W}_{j,k} = \sum_{t\in\mathbb{Z}} \mathbf{h}_j(\gamma_j k - t) Y_t = \sum_{t\in\mathbb{Z}} \mathbf{h}_j(\gamma_j k - t) \Delta^{-K} G(X_t),\ j \geq 0, k \in \mathbb{Z}. \quad (2.17)$$

We then will study the asymptotic behavior of the sequence

$$\overline{\mathbf{S}}_{n_j,j} = \frac{1}{n_j} \sum_{k=0}^{n_j-1} \left(\mathbf{W}_{j,k}^2 - \mathbb{E}[\mathbf{W}_{j,k}^2]\right), \quad (2.18)$$

adequately normalized as $j \to \infty$, where, by convention, in this paper,

$$\mathbf{W}_{j,k}^2 = \{W_{\ell,j,k}^2,\ \ell = 1, \ldots, m\}. \quad (2.19)$$

The squared Euclidean norm of a vector $\mathbf{x} = [x_1, \ldots, x_m]^T$ will be denoted by $|\mathbf{x}|^2 = x_1^2 + \cdots + x_m^2$ and the $L^2$ norm of a random vector $\mathbf{X}$ is denoted by

$$\|\mathbf{X}\|_2 = \left(\mathbb{E}\left[|\mathbf{X}|^2\right]\right)^{1/2}. \quad (2.20)$$

We now summarize the main assumptions of this paper in the following set of conditions.

**Assumptions A.** $\{\mathbf{W}_{j,k},\ j \geq 1, k \in \mathbb{Z}\}$ are the multidimensional wavelet coefficients defined by (2.17), where

(i) $\{X_t\}_{t\in\mathbb{Z}}$ is a stationary Gaussian process with mean 0, variance 1 and spectral density $f$ satisfying (1.3).

(ii) $G$ is a real-valued function whose Hermite expansion (2.1) satisfies condition (2.10), namely $q_0 < 1/(1-2d)$, and whose coefficients in the Hermite expansion satisfy the following condition: for any $\lambda > 0$

$$c_q = O((q!)^d \mathrm{e}^{-\lambda q}) \quad \text{as } q \to \infty. \quad (2.21)$$

(iii) the wavelet filters $(\mathbf{h}_j)_{j\geq 1}$ and their asymptotic Fourier transform $\widehat{\mathbf{h}}_\infty$ satisfy the standard conditions (W-1)–(W-3) with $M$ vanishing moments. See details in Appendix B.

We shall prove that, provided that the number of vanishing moments of the wavelet is large enough, these assumptions yield the following general bound for the centered scalogram.

**Theorem 2.2.** *Suppose that Assumptions A hold with $M \geq K + \delta(q_0)$. Then for any two diverging sequences $(\gamma_j)$ and $(n_j)$, we have, as $j \to \infty$,*

$$\left\|\overline{\mathbf{S}}_{n_j,j}\right\|_2 = O\left(\gamma_j^{2d_0} n_j^{-(1/2-d)}\right). \tag{2.22}$$

*Proof.* Theorem 2.2 is proved in Section 9.2. □

*Idea behind the proof of Theorem 2.2.* One decomposes $\overline{\mathbf{S}}_{n_j,j}$ further in terms $\mathbf{S}_{n_j,j}^{(q,q',p)}$ as in (8.3) and applies the bounds obtained in part in Proposition 8.1.

It is important to note that Theorem 2.2 holds whatever the relative growth of $(\gamma_j)$ and $(n_j)$ but it only provides a bound. This bound will be sufficient to derive a consistent estimator of the long memory parameter $K + \delta(q_0)$, see Theorem 6.1 below.

Obtaining a sharp rate of convergence of the centered scalogram and its asymptotic limit is of primary importance in statistical applications but this can be quite a complicated task. We exhibit several cases in [11, 10] that underline the wild diversity of the asymptotic behavior of the centered scalogram. In general the nature of the limit depends on the relative growth of $(\gamma_j)$ and $(n_j)$. We will show, however, that if $n_j \ll \gamma_j^{\nu_c}$, where $\nu_c$ is a critical exponent, then the reduction principle holds. In this case, the limit will be either Gaussian or expressed in terms of the Rosenblatt process which is defined as follows.

**Definition 2.1.** The Rosenblatt process of index $d$ with

$$1/4 < d < 1/2, \tag{2.23}$$

is the continuous time process

$$Z_d(t) = \int_{\mathbb{R}^2}'' \frac{\mathrm{e}^{\mathrm{i}(u_1+u_2)\,t} - 1}{\mathrm{i}(u_1+u_2)} |u_1|^{-d}|u_2|^{-d}\,\mathrm{d}\widehat{W}(u_1)\mathrm{d}\widehat{W}(u_2),\ t \in \mathbb{R}. \tag{2.24}$$

The multiple integral (2.24) with respect to the complex-valued Gaussian random measure $\widehat{W}$ is defined in Appendix A. The symbol $\int_{\mathbb{R}^2}''$ indicates that one does not integrate on the diagonal $u_1 = u_2$. The integral is well-defined when (2.23) holds because then it has finite $L^2$ norm. This process is self–similar with self-similarity parameter

$$H = 2d \in (1/2, 1),$$

that is for all $a > 0$, $\{Z_d(at)\}_{t\in\mathbb{R}}$ and $\{a^H Z_d(t)\}_{t\in\mathbb{R}}$ have the same finite–dimensional distributions, see [27]. When $t = 1$, $Z_d(1)$ is said to have the Rosenblatt distribution. This distribution is tabulated in [28].

## 3. Reduction principle at large scales

We shall now state the main results and discuss them. They are proved in the following sections. We use $\xrightarrow{\mathcal{L}}$ to denote convergence in law.

The following result involving the case

$$G = \frac{c_{q_0}}{q_0!} H_{q_0},\ c_{q_0} \neq 0,\ q_0 \geq 1,$$

is proved in Theorem 3.2 of [11] and will serve as reference:

**Theorem 3.1.** *Suppose that Assumptions A(i) and A(iii) hold with* $M \geq K + \delta(q_0)$*, where* $\delta(\cdot)$ *is defined in (2.6). Assume that* $Y$ *is a non–linear time series such that* $\Delta^K Y = \frac{c_{q_0}}{q_0!} H_{q_0}(X)$*, with* $q_0 \geq 1$ *and* $q_0 < 1/(1-2d)$*. Define the centered multivariate scalogram* $\overline{\mathbf{S}}_{n,j}$ *related to* $Y$ *by (2.16) and let* $(n_j)$ *and* $(\gamma_j)$ *be any two diverging sequences of integers.*

*(a) Suppose* $q_0 = 1$ *and that* $(\gamma_j)$ *is a sequence of even integers. Then, as* $j \to \infty$*,*

$$n_j^{1/2} \gamma_j^{-2(d+K)} \overline{\mathbf{S}}_{n_j,j} \xrightarrow{\mathcal{L}} c_1^2 \mathcal{N}(0, \Gamma), \tag{3.1}$$

*where* $\Gamma$ *is the* $m \times m$ *matrix with entries*

$$\Gamma_{\ell,\ell'} = 4\pi (f^*(0))^2 \int_{-\pi}^{\pi} \left| \sum_{p \in \mathbb{Z}} |\lambda + 2p\pi|^{-2(K+d)} [\widehat{h}_{\ell,\infty} \overline{\widehat{h}_{\ell',\infty}}](\lambda + 2p\pi) \right|^2 \mathrm{d}\lambda, \quad 1 \leq \ell, \ell' \leq m. \tag{3.2}$$

*(b) Suppose* $q_0 \geq 2$*. Then as* $j \to \infty$*,*

$$n_j^{1-2d} \gamma_j^{-2(\delta(q_0)+K)} \overline{\mathbf{S}}_{n_j,j} \xrightarrow{\mathcal{L}} \frac{c_{q_0}^2}{(q_0-1)!} f^*(0)^{q_0}\, \mathbf{L}_{q_0-1}\, Z_d(1), \tag{3.3}$$

*where* $Z_d(1)$ *is the Rosenblatt process in (2.24) evaluated at time* $t = 1$*,* $f^*(0)$ *is the* short-range spectral density *at zero frequency in (1.1) and where for any* $p \geq 1$*,* $\mathbf{L}_p$ *is the deterministic* $m$*-dimensional vector* $[L_p(\widehat{h}_{\ell,\infty})]_{\ell=1,\ldots,m}$ *with finite entries defined by*

$$L_p(g) = \int_{\mathbb{R}^p} \frac{|g(u_1 + \cdots + u_p)|^2}{|u_1 + \cdots + u_p|^{2K}} \prod_{i=1}^{p} |u_i|^{-2d}\, \mathrm{d}u_1 \cdots \mathrm{d}u_p, \tag{3.4}$$

*for any* $g : \mathbb{R} \to \mathbb{C}$*.*

Thus Theorem 3.1 states that in the case $G = H_{q_0}$, $q_0 \geq 1$ the limit of the scalogram is either Gaussian or has a Rosenblatt distribution[2]. Our main result Theorem 3.2 states that beyond this simple case, the limits continue to be either Gaussian or Rosenblatt under fairly general conditions, involving $n_j$ and

[2] This case corresponds to $\mathcal{L} = \{0\}$ using the notation introduce in (4.1) below.

$\gamma_j$, namely that $n_j \ll \gamma_j^{\nu_c}$ as $j \to \infty$ where $\nu_c$ is a positive (possibly infinite) *critical exponent* given in Definition 4.1, see Section 4 for details.

**Theorem 3.2.** *Suppose that Assumptions A hold with $M \geq K + \delta(q_0)$, where $\delta(\cdot)$ is defined in (2.6) and that*

$$d \notin \{1/2 - 1/(2q) : q = 1, 2, 3, \dots\}. \tag{3.5}$$

*Define the centered multivariate scalogram $\overline{\mathbf{S}}_{n,j}$ related to $Y$ by (2.16). Let $(n_j)$ be any diverging sequence of integers such that, as $j \to \infty$,*

$$n_j \ll \gamma_j^{\nu_c}, \tag{3.6}$$

*where $\nu_c$ is given in Definition 4.1 below. Then, the following limits hold depending on the value of $q_0$.*

*(a) If $q_0 = 1$ and $\gamma_j$ even, then, the convergence (3.1) holds.*
*(b) If $q_0 \geq 2$, then, the convergence (3.3) holds.*

*Proof.* We shall prove in Theorem 8.2, see (8.9), that, under Conditions (3.5) and (3.6), $\overline{\mathbf{S}}_{n_j,j}$ can be *reduced* to a dominating term $\mathbf{S}_{n_j,j}^{(q_0,q_0,q_0-1)}$ in the sense of the $L^2$ norm (2.20). This dominating term depends only on the term $c_{q_0} H_{q_0}(X)/(q_0!)$ of the expansion of $G(X)$. We can then apply Theorem 3.1 to conclude. □

This result extends Theorem 3.1 stated above, where $G$ was restricted to $G = \frac{c_{q_0}}{q_0!} H_{q_0}$. While extending the result to a much more general function $G$, Theorem 3.2 involves two additional conditions. Condition (3.5) is merely here to avoid logarithmic corrections, see Remark 3.3 below. Condition (3.6) is restrictive only when $\nu_c$ is finite, in which case it imposes a minimal growth of the analyzing scale $\gamma_j$ with respect to that of $n_j$. We say that the *reduction principle holds at large scales*. The main interest of having a reduction principle is to conclude that the same asymptotic analysis is valid as in the case $G = \frac{c_{q_0}}{q_0!} H_{q_0}$.

**Remark 3.1.** In practice such a result can be used as follows: If $d$, $G$ are both known, $\nu_c$ can be evaluated numerically. We then get a practical condition, albeit asymptotic, for the reduction principle. See Section 6.3 for an application.

**Remark 3.2.** The case where G is unknown is much more complicated. In this case, there is to our knowledge, no practical way to estimate $\nu_c$ nor even the Hermite rank of the time series. There is, nevertheless, a situation where one can obtain easily the associated critical index $\nu_c$. This is when the time series $Y$ is stationary with $K = 0$ and $G$ is even. In this case the Hermite rank is greater or equal to 2 and $\nu_c = \infty$ (see Section 5 for more details).

**Remark 3.3.** The values $d = 1/2 - 1/(2q)$, $q \geq 1$, constitute boundary values which already appear in the classical reduction theorem, see [26]. These boundary values also exist in our context. If $d = 1/2 - 1/(2q)$, $q \geq 1$, one gets similar results but with logarithm terms. In fact, one can show that if one drops the restriction (3.5), then the conclusion of Theorem 3.2 holds if

1) $n_j \ll \gamma_j^{\nu_c} (\log \gamma_j)^{-4}$.
2) For any $\varepsilon > 0$, $\log \gamma_j = o(n_j^{\varepsilon})$ as $j \to \infty$.

The technical condition 2) is very weak and condition 1) is the same as (3.6) up to a logarithmic correction. We assume (3.5) for simplicity of the exposition.

**Remark 3.4.** We provided in [10] several examples for which different limits are obtained. In these examples one does not have (3.6) and consequently different terms in the decomposition in Wiener chaos of the scalogram dominate and provide different limits. Since the limits are not the same as when $G = H_{q_0}$, the reduction principle does not hold in these cases.

## 4. Critical exponent

The precise description of the critical exponent given below involves a number of sequences, in particular, the subsequence of Hermite coefficients $c_q$, $q \geq 1$ that are *non-vanishing*. We denote this subsequence by $\{c_{q_\ell}\}_{\ell\in\mathcal{L}}$ where $(q_\ell)_{\ell\in\mathcal{L}}$ is a (finite of infinite) increasing sequence of integers such that

$$q_\ell = \text{ index of the } (\ell+1)\text{th non–zero coefficient}, \quad \ell \in \mathcal{L}. \tag{4.1}$$

Thus the indexing set $\mathcal{L}$ is a set of consecutive integers starting at 0 with same cardinality as the set of non-vanishing coefficients. We set

$$I_0 = \{\ell \in \mathcal{L}\ : \ell+1 \in \mathcal{L}, q_{\ell+1} - q_\ell = 1\}, \tag{4.2}$$

that is, $q_\ell$ and $q_{\ell+1}$ take consecutive values when $\ell \in I_0$. The set $I_0$ could be either empty (there are no consecutive values of $q_\ell$) or not empty. Then we set

$$\ell_0 = \begin{cases} \min(I_0) \geq 0, & \text{when } I_0 \text{ is not empty},\\ \infty, & \text{when } I_0 \text{ is empty}. \end{cases} \tag{4.3}$$

When $\ell_0$ is finite (that is, $I_0$ is not empty), $q_{\ell_0}$ is the smallest index $q$ such that two Hermite coefficients $c_q$, $c_{q+1}$ are non–zero.

We define similarly for any $r \geq 0$

$$I_r = \{\ell \in \mathcal{L} : q_{\ell+1} = q_\ell + r + 1\}. \tag{4.4}$$

which involves the terms distant by $r+1$. Finally, we extend the definition of $\ell_0$ in (4.3) to any $r \geq 0$ by

$$\ell_r = \min(I_r). \tag{4.5}$$

We also define

$$\mathcal{R} = \{r \geq 0 : I_r \neq \emptyset \text{ and } \delta(r+1) > 0\}. \tag{4.6}$$

Thus $r \in \mathcal{R}$ describes the gaps $r+1$ where $H_{r+1}(X_t)$ is long-range dependent. Since by (2.6), $\delta(r+1) > 0$ is equivalent to $r+1 < 1/(1-2d)$, we have

$$\mathcal{R} \subset \{0, 1, \ldots, [1/(1-2d)] - 1\}. \tag{4.7}$$

Finally, let

$$J_d = \{\ell \in \mathcal{L} : \delta(q_{\ell+1} - q_\ell) > 0\} = \left\{\ell \in \mathcal{L} : q_{\ell+1} < q_\ell + (1-2d)^{-1}\right\}, \tag{4.8}$$

where we used the expression for $\delta(q)$ in (2.6). Note that

$$J_d = \bigcup_{r\in\mathcal{R}} I_r, \tag{4.9}$$

and thus

$$J_d \neq \emptyset \Longleftrightarrow \mathcal{R} \neq \emptyset. \tag{4.10}$$

We illustrate these quantities in the following example.

**Illustration** Suppose

$$G(x) = c_1\, H_1(x) + \frac{c_3}{3!}\, H_3(x) + \frac{c_4}{4!}\, H_4(x) + \frac{c_5}{5!}\, H_5(x) + \frac{c_{24}}{24!}\, H_{24}(x),$$

where $c_1$, $c_3$, $c_4$, $c_5$ and $c_{24}$ are non-zero constants. Then

$$\begin{aligned}
&q_0 = 1,\ q_1 = 3,\ q_2 = 4,\ q_3 = 5,\ q_4 = 24 \text{ and } \mathcal{L} = \{0,1,2,3,4\},\\
&I_0 = \{1,2\},\ I_1 = \{0\},\ I_2 = \cdots = I_{17} = \emptyset,\ I_{18} = \{3\},\\
&\ell_0 = 1, \ell_1 = 0, \ell_{18} = 3.
\end{aligned}$$

To determine $\mathcal{R}$ we need to involve $d$. Here $q_0 = 1$ so $d$ can take any value in $(0, 1/2)$ to satisfy Condition (2.10) which guarantees that $G(X)$ is long-range dependent. We need to consider the gaps of size 1, 2 and 19, namely, $r = 0, 1$ and 18. Consequently, by (2.6) and using the fact that $\delta(q)$ is decreasing,

a) If $d \in (0, 1/4]$, or equivalently $\delta(2) \leq 0$, then $\mathcal{R} = \{0\}$.
b) If $d \in (1/4, 9/19]$, or equivalently $\delta(2) > 0$ and $\delta(19) \leq 0$, then $\mathcal{R} = \{0, 1\}$.
c) If $d \in (9/19, 1/2)$, or equivalently $\delta(19) > 0$, then $\mathcal{R} = \{0, 1, 18\}$.

Finally, by (4.9), we get for $J_d$ the following subsets of $\mathcal{L}$. In Case a): $J_d = I_0 = \{1, 2\}$, Case b): $J_d = I_0 \cup I_1 = \{0, 1, 2\}$ and Case c): $J_d = I_0 \cup I_1 \cup I_{18} = \{0, 1, 2, 3\}$.

These sets and indices enter in the following definition.

**Definition 4.1.** The critical exponent is

$$\nu_c = \begin{cases}
\infty,\ \text{if } \mathcal{L} = \{0\},\\[1ex]
\infty,\ \text{if } q_0 = 1,\ d \leq 1/4 \text{ and } I_0 = \emptyset,\\[1ex]
\frac{d+1/2-2\delta_+(q_{\ell_0})}{d},\ \text{if } q_0 = 1,\ d \leq 1/4 \text{ and } I_0 \neq \emptyset,\\[1ex]
\frac{1-2\delta_+(q_1-1)}{2d-1/2},\ \text{if } q_0 = 1,\ d > 1/4,\ 1 \in \mathcal{L} \text{ and } J_d = \emptyset,\\[1ex]
\min\left(\frac{1-2\delta_+(q_1-1)}{2d-1/2}, \frac{2d+1/2-2\delta_+(q_{\ell_r})-\delta(r+1)}{\delta(r+1)} : r \in \mathcal{R}\right),\\
\quad \text{if } q_0 = 1,\ d > 1/4 \text{ and } J_d \neq \emptyset,\\[1ex]
\infty,\ \text{if } q_0 \geq 2 \text{ and } I_0 = \emptyset,\\[1ex]
1 + \frac{4(\delta(q_0)-\delta_+(q_{\ell_0}))}{1-2d},\ \text{if } q_0 \geq 2 \text{ and } I_0 \neq \emptyset.
\end{cases}$$

The exponent $\nu_c$ depends on $d$ and on the function $G$ through the expansion coefficient indices $(q_\ell)_{\ell\in\mathcal{L}}$ defined in (4.1). In fact one has

**Proposition 4.1.** *Every possible sequence $(q_\ell)_{\ell\in\mathcal{L}}$ and every value of $d$ satisfying (2.10) give rise to a $\nu_c \in (0,\infty]$.*

*Proof.* See Section 9.3. □

The value $\nu_c = \infty$ is the simplest case since then the reduction principle holds whatever the respective growth rates of the diverging sequences $(n_j)$ and $(\gamma_j)$ are. This happens for instance when there are no consecutive non–zeros coefficients ($I_0 = \emptyset$) and either $q_0 = 1$ and $d \le 1/4$ or $q_0 \ge 2$ (which implies $d > 1/4$).

## 5. Examples

In this section, we examine some specific cases of functions $G$. We always assume that $G$ satisfies Assumption A(ii).

### 5.1. *G is even*

If $G$ is an even function then $q_0 \ge 2$ and $I_0 = \emptyset$ because the Hermite expansion has only even terms. Hence $\nu_c = \infty$ and the reduction principle applies for any diverging sequences $(n_j)$ and $(\gamma_j)$.

### 5.2. *G is odd*

If $G$ is an odd function then we have again $I_0 = \emptyset$ since the Hermite expansion has no even terms. But unlike the even case, we may have $q_0 = 1$. If it is not the case, then $q_0 \ge 3$ so that $\nu_c = \infty$ and the reduction principle applies for any diverging sequences $(n_j)$ and $(\gamma_j)$. If $q_0 = 1$ and $d \le 1/4$, we find again $\nu_c = \infty$. If $q_0 = 1$ and $d > 1/4$, the formula of the exponent $\nu_c$ is more involved and takes various possible forms, see Section 5.4 for one of the possible cases, namely $I_0 = \emptyset$, $q_0 = 1$ and $\delta(q_1) > 0$.

### 5.3. $I_0 \ne \emptyset$ *and* $q_0 \ge 2$

This corresponds to the class studied in Section 3.1 of [10] with the additional condition $\delta(q_{\ell_0} + 1) > 0$ (see (3.3) in this reference). Using this additional condition, we have $\delta(q_{\ell_0}) > 0$ since $\delta(q)$ is decreasing. Hence $\delta_+(q_{\ell_0}) = \delta(q_{\ell_0})$ and

$$\nu_c = 1 + \frac{4(\delta(q_0) - \delta_+(q_{\ell_0}))}{1-2d} = 1 + \frac{4(\delta(q_0) - \delta(q_{\ell_0}))}{1-2d} = 1 + 2(q_{\ell_0} - 2q_0).$$

This value of $\nu_c$ corresponds to the exponent $\nu$ defined in (3.4) and appearing in Theorem 3.1 of [10]. This theorem shows that if the opposite condition to (3.6) holds, namely, $\gamma_j^{\nu_c} \ll n_j$, then the reduction principle does not apply since the limit is Gaussian instead of Rosenblatt. We say that the *reduction principle does not apply at small scales*. In Theorem 3.2, the reduction principle is proved even when $\delta_+(q_{\ell_0}+1) = 0$, but whether the reduction principle does apply or not at small scales, namely if $\gamma_j^{\nu_c} \ll n_j$, remains an open question.

### *5.4.* $I_0 = \emptyset$, $q_0 = 1$ *and* $\delta(q_1) > 0$

The expansion of $G$ contains $H_1$ but does not contain any two consecutive polynomials. This corresponds to the class studied in Section 3.2 of [10] (see (3.8) in this reference). The exponent $\nu_c$ simplifies as follows. First observe that $\delta(q_1) > 0$ implies $\delta_+(q_1 - 1) > 0$, so that $q_1 \in \mathcal{R}$, and also $\delta(2) > 0$ and hence $d > 1/4$. We thus need to focus on the term of $\nu_c$ in Definition 4.1 involving min. Using (2.6), for the first term in the min

$$\frac{1 - 2\delta_+(q_1 - 1)}{2d - 1/2} = \frac{(q_1 - 1)(1 - 2d)}{\delta(2)}, \tag{5.1}$$

which corresponds to the exponent $\nu_2$ in (3.10) of [10]. Now focus on the second term in the min. Take any $r \in \mathcal{R}$ and consider $\ell_r$ defined in (4.5). Note that $q_{\ell_r}$ is the smallest Hermite polynomial index of the expansion of $G$ such that the next one appears after a gap equal to $r+1$. There are only two possibilities: (a) either $q_{\ell_r} = q_0 = 1$, (b) or $q_{\ell_r} \geq q_1$. In case (a), we have $r + 1 = q_{\ell_r+1} - q_{\ell_r} = q_1 - 1$ and thus

$$\frac{2d + 1/2 - 2\delta_+(q_{\ell_r}) - \delta(r+1)}{\delta(r+1)} = \frac{1/2 - \delta(q_1 - 1)}{\delta(q_1 - 1)}, \tag{5.2}$$

which corresponds to the exponent $\nu_1$ in (3.10) of [10]. In case (b), using $r+1 \geq 2$ (since $I_0 = \emptyset$) and $q_{\ell_r} \geq q_1$, we get

$$\begin{aligned}\frac{2d + 1/2 - 2\delta_+(q_{\ell_r}) - \delta(r+1)}{\delta(r+1)} &\geq \frac{2d + 1/2 - 2\delta_+(q_1) - \delta(2)}{\delta(2)} \\ &= \frac{q_1(1 - 2d)}{\delta(2)} > \frac{(q_1 - 1)(1 - 2d)}{\delta(2)},\end{aligned}$$

which already appeared in (5.1). Therefore with (5.1) and (5.2) and Definition 4.1 of $\nu_c$ for $q_0 = 1$ and $d > 1/4$, we get

$$\nu_c = \min\left(\frac{(q_1 - 1)(1 - 2d)}{\delta(2)}, \frac{1/2 - \delta(q_1 - 1)}{\delta(q_1 - 1)}\right),$$

which corresponds to $\min(\nu_1, \nu_2)$ using the definitions in (3.10) of [10]. Hence the reduction principle established in Theorem 3.2 under the condition $n_j \ll \gamma_j^{\nu_c}$ corresponds to the cases $n_j \ll \gamma_j^{\nu_1}$ and $n_j \ll \gamma_j^{\nu_2}$ of Theorems 3.3 and 3.5 in [10], respectively. These two theorems further show that when the additional condition $\delta(q_1) > 0$ holds the reduction principle does not hold under the opposite condition $\gamma_j^{\nu_c} \ll n_j$, illustrating the fact that the reduction principle may not hold at small scales.

## 6. Application to wavelet statistical inference

### 6.1. Wavelet inference setting

Suppose that we observe a sample $Y_1, \ldots, Y_N$ of $Y$. Recall that $Y$ has long memory parameter $d_0 = K + \delta(q_0)$. In this section, we assume that we are given an unidimensional wavelet filter $g_j$ satisfying Assumptions (W-1)–(W-3) in Appendix B (see also (C.1) and (C.6)). Then one can derive the wavelet estimator

$$\hat{d}_0 = \sum_{i=0}^{p} w_i \log \hat{\sigma}^2_{j+i}, \tag{6.1}$$

where $w_0, \ldots, w_p$ are weights satisfying

$$\sum_{i=0}^{p} w_i = 0 \qquad \text{and} \qquad \sum_{i=0}^{p} i w_i = 1/(2\log 2), \tag{6.2}$$

and $(\hat{\sigma}^2_j)_{i \le j \le i+p}$ denotes the multiscale scalogram obtained from $Y_1, \ldots, Y_N$,

$$(\hat{\sigma}^2_j)_{i \le j \le i+p} = \mathbf{S}_{n_j, j} = \frac{1}{n_j} \sum_{k=0}^{n_j - 1} \mathbf{W}^2_{j,k}, \tag{6.3}$$

(see Appendix C for more details). In this setting, $(\gamma_j)$ and $(n_j)$ are specified as follows

$$\gamma_j = 2^j \quad \text{and} \quad n_j = N 2^{-j} + O(1). \tag{6.4}$$

As usual in this setting the asymptotics are to be understood as $N \to \infty$ with a well chosen diverging sequence $j = j_N$ such that

$$\lim_{N \to \infty} N 2^{-j} = \infty, \tag{6.5}$$

and thus $(n_j)$ diverge as $N \to \infty$. We refer to [20, Theorem 1] for the asymptotic behavior of the mean of the scalogram

$$\mathbb{E}\left[\hat{\sigma}^2_j\right] = C\, 2^{2 d_0 j} \, \left(1 + O(2^{-\zeta j})\right), \tag{6.6}$$

where $C$ is a positive constant and $\zeta$ is an exponent satisfying the conditions of Theorem 2.1. This relation follows from Theorem 2.1, provided that $M \ge d_0 - 1/2$. Choosing weights such that the conditions in (6.2) hold then yields

$$\sum_{i=0}^{p} w_i \log \mathbb{E}\left[\hat{\sigma}^2_{j+i}\right] = d_0 + O(2^{-\zeta j}). \tag{6.7}$$

### 6.2. Consistency

We now state a consistency result.

**Theorem 6.1.** *Consider the wavelet estimation setting (6.1)–(6.5) and suppose that Assumptions A hold with $M \ge K + \delta(q_0)$. Then, as $N \to \infty$, $\hat{d}_0$ converges to $d_0$ in probability.*

*Proof.* By (6.1), we have

$$\hat{d}_0 - \sum_{i=0}^{p} w_i \log \mathbb{E}\left[\hat{\sigma}^2_{j+i}\right] = \sum_{i=0}^{p} w_i \log\left(1 + \frac{\hat{\sigma}^2_{j+i} - \mathbb{E}\left[\hat{\sigma}^2_{j+i}\right]}{\mathbb{E}\left[\hat{\sigma}^2_{j+i}\right]}\right). \tag{6.8}$$

The numerators in the last ratio are the components of $\overline{\mathbf{S}}_{n_j,j}$ by (6.3). By Theorem 2.2 and (6.4), we have

$$\overline{\mathbf{S}}_{n_j,j} = O_P\left(\gamma_j^{2d_0}\, n_j^{-(1/2-d)}\right) = O_P\left(2^{2d_0 j}\,(N2^{-j})^{-(1/2-d)}\right).$$

Hence, with (6.8) and (6.6), we get that

$$\hat{d}_0 - \sum_{i=0}^{p} w_i \log \mathbb{E}\left[\hat{\sigma}^2_{j+i}\right] = O_P((N2^{-j})^{-(1/2-d)}).$$

Applying (6.7) then yields

$$\hat{d}_0 = d_0 + O_P\left((N2^{-j})^{-(1/2-d)}\right) + O(2^{-\zeta j}). \tag{6.9}$$

The result then follows from (6.5). □

**Remark 6.1.** We note that this consistency result applies without any knowledge of $G$ or $\beta$.

### *6.3. Hypothesis testing*

Consider again a sample $Y_1, \ldots, Y_N$ of $Y$ and suppose now that $G$ is known and has Hermite rank $q_0$.

Denote by $\widetilde{d}_0$ the estimator that would be obtained instead of $\hat{d}_0$ if we had $G$ replaced by $c_{q_0} H_{q_0}/(q_0!)$. We shall apply Theorem C.1 and Theorem C.2 of Appendix C. Theorem C.1 (case $q_0 = 1$) derives from Theorem 2 of [24] and Theorem C.2 (case $q_0 \geq 2$) derives from Theorem 4.1 of [11]. We obtain the following: for conveniently chosen diverging sequences $j = (j_N)$, there exists some renormalization sequence $(u_N)$ such that as $N \to \infty$,

$$u_N(\widetilde{d}_0 - d_0) \overset{(\mathcal{L})}{\to} U(d, K, q_0), \tag{6.10}$$

with

$$u_N = \begin{cases} (N2^{-j})^{1/2} \text{ if } q_0 = 1, \\ (N2^{-j})^{1-2d} \text{ if } q_0 \geq 2, \end{cases} \tag{6.11}$$

and where $U(d, K, q_0)$ is a centered Gaussian random variable if $q_0 = 1$ and a Rosenblatt random variable if $q_0 \geq 2$. The precise distribution of $U(d, K, q_0)$ is given in Theorems C.1 and C.2. Beside the chosen wavelet, the distribution of $U$ only depends on $d$, $K$ and $q_0$.

As application of the reduction principle in this setting, we use (6.10) to define a statistical test procedure which applies to a general $G$. Let $d_0^*$ be a given possible value for the true unknown memory parameter $d_0$ of $Y$ and consider the hypotheses

$$H_0 : d_0 = d_0^* \quad \text{against} \quad H_1 : d_0 \in (0, \bar{K} + 1/2) \setminus \{d_0^*\}.$$

Here $\bar{K}$ denotes a known *maximal value* for the true (possibly unknown) integration parameter $K$. So to insure that the number $M$ of vanishing moments satisfies $M \geq d_0$, it suffices to impose $M > \bar{K}$. Since $G$ is assumed to be known, for the given value $d_0^*$, one can define the parameters $d^*$, $K^*$ and $\nu_c^*$ defined as $d$, $K$ and $\nu_c$ by replacing $d_0$ by $d_0^*$.

Let $\alpha \in (0,1)$ be a significance level. Define the statistical test

$$\delta_s = \begin{cases} 1 & \text{if } |\hat{d}_0 - d_0^*| > s_N(\alpha), \\ 0 & \text{otherwise.} \end{cases} \tag{6.12}$$

where $s_N(\alpha)$ is the $(1-\alpha/2)$ quantile of $U(d^*, K^*, q_0)/u_N$.

The following theorem provides conditions for the test $\delta_s$ to be consistent with asymptotic significance level $\alpha$, namely, that its power goes to 1 and its first type error goes to $\alpha$ as $N$ goes to $\infty$.

**Theorem 6.2.** *Suppose that Assumptions A(i), (ii) hold with $M > \bar{K}$ and that the unidimensional wavelet filter $g_j$ satisfies Assumptions (W-1)–(W-3). Assume additionally that (3.5) holds. Let $j = (j_N)$ be a diverging sequence such that (6.5) holds. Suppose moreover that, as $N \to \infty$,*

$$N 2^{-j} \ll 2^{j\nu_c^*}, \tag{6.13}$$

*and that there exists a positive exponent $\zeta$ satisfying (2.12) and*

$$2^{-\zeta j} \ll u_N^{-1}, \tag{6.14}$$

*with $u_N$ defined as in (6.11). Then, if (3.6) is satisfied, $\delta_s$ is a consistent test with asymptotic significance level $\alpha$.*

**Remark 6.2.** Observe that the different conditions that have to be simultaneously satisfied by $(j_N)$ can be reformulated as follows:

- $\lim_{N\to\infty} j_N = \infty$ and $\lim_{N\to\infty} N 2^{-j_N} = \infty$.
- $N 2^{-j_N} \ll 2^{j_N \zeta'}$ with

$$\zeta' = \begin{cases} \min(\nu_c^*, 2\zeta) \text{ if } q_0 = 1, \\ \min(\nu_c^*, \zeta/(1-2d)) \text{ otherwise.} \end{cases}$$

In particular, one can easily check that since $\nu_c^*$ and $\zeta$ are both positive so is $\zeta'$. Hence these conditions are not incompatible.

*Proof.* See Section 9.5. □

*Idea behind the proof of Theorem 6.2.* Condition (6.13) states that $n_j \ll \gamma_j^{\nu_c^*}$ and will insure that the reduction principle holds under $H_0$. Condition (6.14) will ensure that the bias is negligible under $H_0$. These conditions will allow us through Relation (9.26) to transfer the problem to the case $G(x) = \frac{c_{q_0}}{q_0!} H_{q_0}(x)$ which was treated in [11].

## 7. Numerical experiments

### *7.1. General setting*

We now investigate the performance of the proposed estimation and test procedures based on Monte Carlo simulations. The raw samples, wavelet transforms and wavelet estimators are generated and computed using the LRD toolbox documented in [12]. The quantiles of the Rosenblatt distribution are computed using [28]. Each Monte Carlo simulation mentioned below is based on 1000 independent draws of samples $(Y_t)_{1\leq t\leq N}$.

We consider two models, both obtained by non-linear transformation of a sample $(X_t)_{1\leq t\leq N}$ generated according to a Gaussian ARFIMA$(0, d, 0)$ distribution with zero mean and unit variance. In both models we take $K = 0$ to simplify the setting.

*Model 1* In this model, we set $Y_t = H_1(X_t) + 1/(2\sqrt{3})H_3(X_t)$. This model is a particular instance of the example of Section 5.2 with $q_0 = 1$, hence $d_0 = d$. Let us restrict our experiments to $d \in (1/4, 1/2)$ so that we fall in the case investigated in Section 5.4 and thus, using $q_1 = 3$ and the computations of Section 5.4, we find $\nu_c = (1 - 2d)/(2d - 1/2)$, which is decreasing from $\infty$ to 0 as $d$ goes from 1/4 to 1/2.

*Model 2* Here we set $Y_t = 1/\sqrt{2}H_2(X_t) + 1/(2\sqrt{3})H_3(X_t)$. This model is a particular instance of the example described in Section 5.3 with $q_0 = q_{\ell_0} = 2$. Then $d \in (1/4, 1/2)$ ensures that $d_0 = \delta(2) = 2d - 1/2 > 0$. Moreover, by Definition 4.1, $\nu_c = 1$.

For each model and each sample length $N$, any statistical procedure based on the statistic $\hat{d}_0$ defined in (6.1) raises the question of the choice of the minimal scale index $j$ and of the weights $w_i$, $i = 0, \ldots, p$. We first explain how we choose $j$ and then the weights.

### *7.2. Choice of the scale index j*

The asymptotic behaviors (Gaussian or Rosenblatt) derived in Section 6.3 hold only if the bias is negligible (Condition (6.14)) and if the reduction principle hold (Condition (6.13)), which is summarized in Remark 6.2 using the exponent $\zeta'$. We chose to set $j = [\log_2 N/(\zeta' + 1)]$ in our experiments, which neither guarantees that the bias is negligible nor that the reduction principle holds but which corresponds to a limit value, sufficiently largely above which both conditions hold. In Tables 1 and 2, we display the obtained values of $j$ corresponding to particular $d$'s and $N = 2^{12}$ or $N = 2^{15}$.

### *7.3. Choice of the weights $w_i$, $i = 0, \ldots, p$*

Recall that the weights $w_i$ in 6.2 influence the asymptotic variance of $\hat{d}_0$. We compare two possible choices. First the two-scales regression, for which $p = 1$ and

Table 1
*Model 1: scale index j depending on d and N*

| | d=0.3 | d=0.35 | d=0.4 |
|---|---|---|---|
| $N = 2^{12}$ | 5 | 5 | 7 |
| $N = 2^{15}$ | 6 | 6 | 9 |

Table 2
*Model 2: scale index j depending on d and N*

| | d=0.3 | d=0.35 | d=0.4 | d=0.45 |
|---|---|---|---|---|
| $N = 2^{12}$ | 8 | 6 | 6 | 6 |
| $N = 2^{15}$ | 10 | 7 | 7 | 7 |

Table 3
*Performances of two-scales regression estimator (Model 1, $N = 2^{12}$)*

| | bias | std | MSE |
|---|---|---|---|
| d=0.3 | -0.0555 | 0.1973 | 0.0420 |
| d=0.325 | -0.0650 | 0.1946 | 0.0421 |
| d=0.35 | -0.1190 | 0.2041 | 0.0451 |
| d=0.375 | -0.0693 | 0.2801 | 0.0832 |
| d=0.4 | -0.0884 | 0.3858 | 0.1565 |

Table 4
*Performances of two-scales regression estimator (Model 2, $N = 2^{12}$)*

| | $d_0$ | bias | std | MSE |
|---|---|---|---|---|
| d=0.35 | 0.2 | -0.0464 | 0.3132 | 0.1001 |
| d=0.375 | 0.25 | -0.0682 | 0.3160 | 0.1044 |
| d=0.4 | 0.3 | -0.0615 | 0.3041 | 0.0961 |
| d=0.425 | 0.35 | -0.0640 | 0.3099 | 0.1001 |
| d=0.45 | 0.4 | -0.0638 | 0.3083 | 0.0991 |

Table 5
*Performances of two-scales regression estimator (Model 1, $N = 2^{15}$)*

| | bias | std | MSE |
|---|---|---|---|
| d=0.3 | -0.0340 | 0.0881 | 0.0089 |
| d=0.325 | -0.0885 | 0.1374 | 0.0267 |
| d=0.35 | -0.0418 | 0.1019 | 0.0121 |
| d=0.375 | -0.0460 | 0.1452 | 0.0232 |
| d=0.4 | -0.0412 | 0.2703 | 0.0747 |

Table 6
*Performances of two-scales regression estimator (Model 2, $N = 2^{15}$)*

| | $d_0$ | bias | std | MSE |
|---|---|---|---|---|
| d=0.35 | 0.2 | -0.0272 | 0.1776 | 0.0323 |
| d=0.375 | 0.25 | -0.0335 | 0.1776 | 0.0326 |
| d=0.4 | 0.3 | -0.0316 | 0.1842 | 0.0349 |
| d=0.425 | 0.35 | -0.0874 | 0.4067 | 0.1728 |
| d=0.45 | 0.4 | -0.0428 | 0.1670 | 0.0297 |

Table 7
*Performances of four-scales regression estimator (Model 1, $N = 2^{12}$)*

| | bias | std | MSE |
|---|---|---|---|
| d=0.3 | -0.0542 | 0.0862 | 0.0104 |
| d=0.325 | -0.0593 | 0.0846 | 0.0107 |
| d=0.35 | -0.0588 | 0.0880 | 0.0112 |
| d=0.375 | -0.0813 | 0.1295 | 0.0234 |
| d=0.4 | -0.1298 | 0.2147 | 0.0629 |

Table 8
*Performances of four-scales regression estimator (Model 2, $N = 2^{12}$)*

| | $d_0$ | bias | std | MSE |
|---|---|---|---|---|
| d=0.35 | 0.2 | -0.0653 | 0.1494 | 0.0266 |
| d=0.375 | 0.25 | -0.0862 | 0.1608 | 0.0333 |
| d=0.4 | 0.3 | -0.0809 | 0.1566 | 0.0310 |
| d=0.425 | 0.35 | -0.0920 | 0.1559 | 0.0328 |
| d=0.45 | 0.4 | -0.0867 | 0.1596 | 0.0329 |

the conditions in (6.2) imply $w_1 = 1/(2\log 2)$ and $w_0 = -1/(2\log 2)$. Second, the four-scales regression, which corresponds to $p = 3$, with Abry-Veitch weights. These weights have been introduced in the seminal paper [29] and provides close to optimal choice in the linear case, see [12]. Tables 3, 4, 5 and 6 display the obtained bias, variance, and MSE (mean square error) for *Model 1* and *Model 2* with sample lengths $N = 2^{12}$ and $N = 2^{15} = 32768$ for the two-scales regression. We see that the four-scales regression clearly achieves better performances for both models. Hence, in the following experiments, we set the weights to the four-scales Abry-Veitch weights.

TABLE 9
*Performances of four-scales regression estimator (Model 1, $N = 2^{15}$)*

| | bias | std | MSE |
|---|---|---|---|
| d=0.3 | -0.0338 | 0.0402 | 0.0028 |
| d=0.325 | -0.0878 | 0.1112 | 0.0201 |
| d=0.35 | -0.0368 | 0.0425 | 0.0032 |
| d=0.375 | -0.0363 | 0.0619 | 0.0051 |
| d=0.4 | -0.0513 | 0.1289 | 0.0192 |

TABLE 10
*Performances of four-scales regression estimator (Model 2, $N = 2^{15}$)*

| | $d_0$ | bias | std | MSE |
|---|---|---|---|---|
| d=0.35 | 0.2 | -0.0302 | 0.0811 | 0.0075 |
| d=0.375 | 0.25 | -0.0722 | 0.1026 | 0.0157 |
| d=0.4 | 0.3 | -0.0462 | 0.0891 | 0.0101 |
| d=0.425 | 0.35 | -0.0455 | 0.0831 | 0.0090 |
| d=0.45 | 0.4 | -0.0409 | 0.0856 | 0.0090 |

### *7.4. Finite-sample performances for testing $d_0 = d_0^*$ against $d_0 \neq d_0^*$*

We now evaluate the performances of the statistical test defined in (6.12) and studied in Section 6.3 for testing $H_0 : d_0 = d_0^*$ against $H_1 : d_0 \neq d_0^*$. Recall that the test statistic is $|\widehat{d}_0 - d_0^*|$. The threshold $s_N(\alpha)$ yielding the asymptotic significance level $\alpha$ is derived in Theorem 6.2 under some condition on the scale index $j$ (see Remark 6.2). In the case of *Model 1*, the asymptotic distribution is Gaussian, the rate is $u_N = (N2^{-j})^{1/2}$ and the asymptotic variance is the same as for Gaussian processes, so can be computed by relying on the Toolbox detailed in [12]. We display in Tables 11 and 12 the finite-sample rejection rate for standard values of the asymptotic level $\alpha$. We can see that the asymptotic analysis provides an underestimated significance level, which amounts to say that the variance of the estimator $\hat{d}_0$ is underestimated by its theoretical asymptotic value. This can be explained by the fact that the asymptotic one is based on the reduction principle and thus approximates the variance of all non-Gaussian terms of the Wiener chaos decomposition of the scalogram by zero. These experiments show that this approximation is not so sharp for a finite sample, although it slightly improves as $N$ grows. The case of *Model 2* is more complicated as it involves a Rosenblatt distribution and more complicated constants, namely the $L_q(\widehat{g}_\infty)$ for $q = q_0, q_0 - 1$ in (C.13) and (C.14). To circumvent the numerical computation of the asymptotic variance we rely on the following Bootstrap procedure.

TABLE 11
*Rejection rates under $H_0$ for different values of $d_0^*$ and $\alpha$ for Model 1 with $N = 2^{12}$*

| $d_0^*$ | 0.3 | 0.325 | 0.35 | 0.375 | 0.4 |
|---|---|---|---|---|---|
| $\alpha = 0.01$ | 0.0700 | 0.2730 | 0.0750 | 0.0830 | 0.1290 |
| $\alpha = 0.05$ | 0.1770 | 0.3540 | 0.2320 | 0.1800 | 0.2340 |
| $\alpha = 0.1$ | 0.2420 | 0.4290 | 0.3130 | 0.2670 | 0.3090 |

TABLE 12
*Rejection rates under $H_0$ for different values of $d_0^*$ and $\alpha$ for Model 1 with $N = 2^{15}$*

| $d_0^*$ | 0.3 | 0.325 | 0.35 | 0.375 | 0.4 |
|---|---|---|---|---|---|
| $\alpha = 0.01$ | 0.0730 | 0.0700 | 0.0930 | 0.0730 | 0.0590 |
| $\alpha = 0.05$ | 0.1780 | 0.1840 | 0.1830 | 0.1590 | 0.1530 |
| $\alpha = 0.1$ | 0.2630 | 0.2460 | 0.2620 | 0.2390 | 0.2150 |

TABLE 13
*Rejection rates under $H_0$ for different values of $d_0^*$ and different levels $\alpha$ for Model 2 and $N = 2^{12}$)*

| $d_0^*$ | 0.1 | 0.15 | 0.2 | 0.25 | 0.3 | 0.35 | 0.4 |
|---|---|---|---|---|---|---|---|
| $\alpha = 0.01$ | 0.3130 | 0.664 | 0.679 | 0.695 | 0.633 | 0.669 | 0.616 |
| $\alpha = 0.05$ | 0.407 | 0.689 | 0.707 | 0.721 | 0.667 | 0.7 | 0.637 |
| $\alpha = 0.1$ | 0.480 | 0.726 | 0.738 | 0.75 | 0.694 | 0.735 | 0.669 |

TABLE 14
*Rejection rates under $H_0$ for different values of $d_0^*$ and different levels $\alpha$ for Model 2 $N = 2^{15}$)*

| $d_0^*$ | 0.1 | 0.15 | 0.2 | 0.25 | 0.3 | 0.35 | 0.4 |
|---|---|---|---|---|---|---|---|
| $\alpha = 0.01$ | 0.318 | 0.4210 | 0.3230 | 0.3860 | 0.4470 | 0.4250 | 0.4250 |
| $\alpha = 0.05$ | 0.435 | 0.5110 | 0.4280 | 0.4770 | 0.4840 | 0.4580 | 0.4450 |
| $\alpha = 0.1$ | 0.508 | 0.5760 | 0.4970 | 0.5290 | 0.5430 | 0.4980 | 0.4820 |

Step 1 We pick $m$ sub-samples of the original time series of $2^{N-L}$ consecutive observations, randomly with replacement.

Step 2 For $\ell = 1, \ldots, m$, we compute an estimator $\hat{d}_0(\ell)$ based on the $\ell$th sub–sample (with the same $j$ and weights $w_i$ as for $\hat{d}_0$).

Step 3 We compute the empirical variance $\hat{v}_L$ of the sample $\hat{d}_0(\ell)$, $\ell = 1, \ldots, m$ obtained in Step 2 and set the empirical variance of the full sample estimate $\hat{d}_0$ to $\hat{v} = 2^{-L(1-2d_0^*)}\hat{v}_L$.

In our experiments we chose $L = 3$ and $m = 50$. The factor $2^{-L(1-2d^*)}$ relating $\hat{v}$ to $\hat{v}_L$ is inherited from the rate $u_N$ defined in (6.11) in the case $q_0 = 2$ under $H_0$ (since we are in the case of *Model 2*). The threshold associated to the significance level $\alpha$ is then set by approximating $(\hat{d}_0 - d_0^*)/\sqrt{\hat{v}}$ as a Rosenblatt random variable with unit variance. We display a four–scales regression (see Tables 13 and 14). The obtained rejection rate are again larger than $\alpha$ because of the approximation of the scalogram by its sole Rosenblatt term in its Wiener chaos decomposition.

Next we evaluate the power of the statistical test. We focus on the power function, that is, the power as a function of the true $d_0$ for the test corresponding to the null hypothesis $H_0$ with $d_0^* = 0.4$ for *Model 1* and $d_0^* = 0.3$ for *Model 2*, data length $N = 2^{15}$ and significance level $\alpha = 0.05$. Recall that the test is based on the reduction principle. In particular, its definition only depends on the the first term of the Hermite expansion of $G$. Hence we found it interesting to compare the powers of the test procedure, computed for *Model 1* and *Model 2*, with the powers obtained with data generated by changing the function $G$ of each of these models into its first Hermite expansion term. So for *Model 1*, we compare the power function with the one obtained by directly applying the test to $(H_{q_0}(X_t))_{t=1,\ldots,N}$, that is to $(X_t)_{t=1,\ldots,N}$ and for *Model 2*, we compare the power function with the one obtained by applying the test to $(H_2(X_t))_{t=1,\ldots,N}$. In each case the data length is $N = 2^{15}$. The results are displayed in Figures 1 and 2. We also display some ROC curves in Figure 3 for *Model 1* and Figure 4

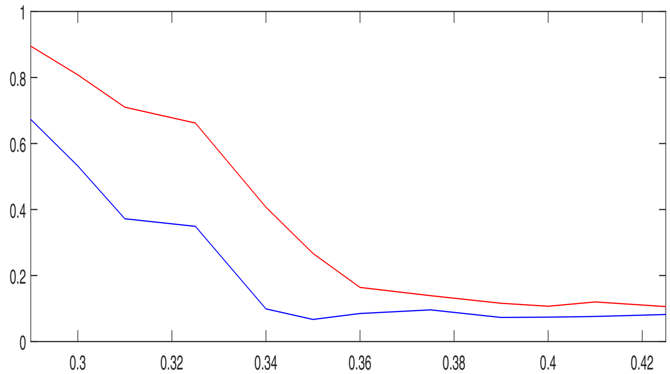


FIG 1. *Rejection rates as a function of* $d_0$ *for two data sets:* $(X_t)_{1\le t\le N}$ *(blue bottom curve), Model 1 (red top curve),* $d_0^* = 0.4$, $N = 2^{15}$.

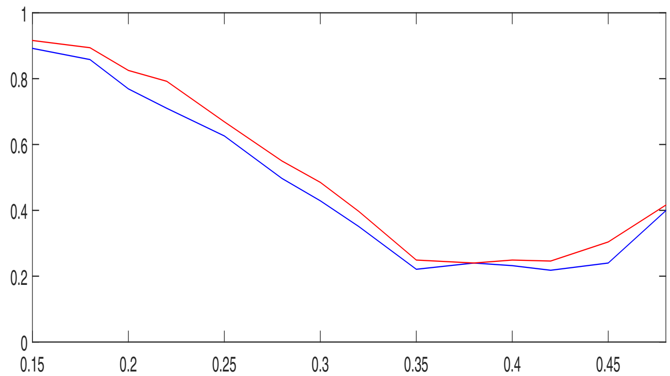


FIG 2. *Rejection rates as a function of* $d_0$ *for two data sets:* $(H_2(X_t))_{1\le t\le N}$ *(blue bottom curve), Model 2 (red top curve),* $d_0^* = 0.3$, $N = 2^{15}$.

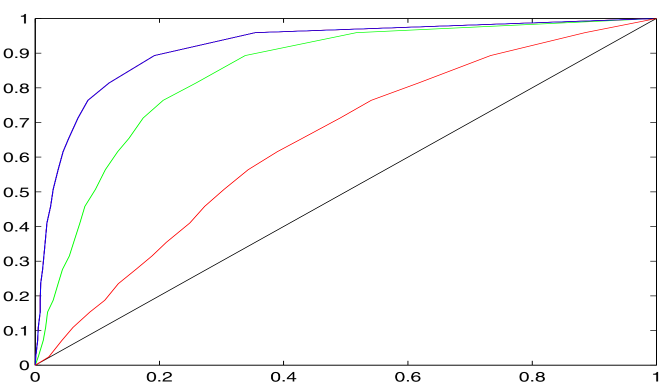


FIG 3. *ROC curves for Model 1 and* $d_0^* = 0.4$ *for three data sets:* $d_0 = 0.3$ *(blue top curve),* $d_0 = 0.325$ *(green middle curve),* $d_0 = 0.35$ *(red bottom curve).* $N = 2^{15}$.

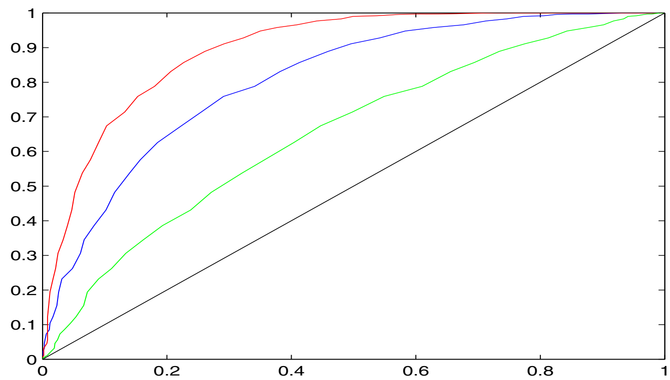


Fig 4. *ROC curves for Model 2 and* $d_0^* = 0.3$ *for three data sets:* $d_0 = 0.15$ *(blue top curve),* $d_0 = 0.2$ *(green middle curve),* $d_0 = 0.25$ *(red bottom curve).* $N = 2^{15}$.

for *Model 2*. The data length $N$ also equals in each case $2^{15}$. The higher the curve the better. We see that, as expected, the power increases, as $d_0$ goes away from $d_0^*$. Moreover, comparing red and blue plots, we see that the power deteriorates when additional terms are added in the Hermite expansion of $G$.

## 8. Decomposition in Wiener chaos

As in [9] and [10], we need the expansion of the scalogram into Wiener chaos. The wavelet coefficients can be expanded in the following way:

$$\mathbf{W}_{j,k} = \sum_{q=1}^{\infty} \frac{c_q}{q!} \mathbf{W}_{j,k}^{(q)}, \tag{8.1}$$

where $\mathbf{W}_{j,k}^{(q)}$ is a multiple integral of order $q$. Then, using the same convention as in (2.19), we have

$$\mathbf{W}_{j,k}^2 = \sum_{q=1}^{\infty} \left(\frac{c_q}{q!}\right)^2 \left(\mathbf{W}_{j,k}^{(q)}\right)^2 + 2\sum_{q'=2}^{\infty}\sum_{q=1}^{q'-1} \frac{c_q}{q!}\frac{c_{q'}}{q'!}\mathbf{W}_{j,k}^{(q)}\mathbf{W}_{j,k}^{(q')}, \tag{8.2}$$

where the convergence of the infinite sums hold in $L^1(\Omega)$ sense.

Each $\mathbf{W}_{j,k}^{(q)}$ is a multiple integral and consequently so is $\overline{\mathbf{S}}_{n_j,j}$ in (2.18). (Basic facts about Multiple integrals and Wiener chaos are recalled in Appendix A).

In Proposition 4.2 of [10], we gave the following explicit expression of the Wiener chaos expansion of the scalogram.

**Proposition 8.1.** *For all* $j$, $\{\mathbf{W}_{j,k}\}_{k\in\mathbb{Z}}$ *is a weakly stationary sequence. Moreover, for any* $j \in \mathbb{N}$, $\overline{\mathbf{S}}_{n_j,j}$ *can be expanded into Wiener chaos as follows*

$$\overline{\mathbf{S}}_{n_j,j} = \frac{1}{n_j}\sum_{k=0}^{n_j-1} \mathbf{W}_{j,k}^2 - \mathbb{E}[\mathbf{W}_{j,0}^2]$$

$$= \sum_{q=1}^{\infty} \left(\frac{c_q}{q!}\right)^2 \sum_{p=0}^{q-1} p! \binom{q}{p}^2 \mathbf{S}_{n_j,j}^{(q,q,p)}$$
$$+ 2 \sum_{q'=2}^{\infty} \sum_{q=1}^{q'-1} \frac{c_q}{q!} \frac{c_{q'}}{q'!} \sum_{p=0}^{q} p! \binom{q}{p} \binom{q'}{p} \mathbf{S}_{n_j,j}^{(q,q',p)}, \tag{8.3}$$

*where, for all* $q, q' \geq 1$ *and* $0 \leq p \leq \min(q, q')$, $\mathbf{S}_{n_j,j}^{(q,q',p)}$ *is of the form*

$$\mathbf{S}_{n_j,j}^{(q,q',p)} = \widehat{I}_{q+q'-2p} \left( \mathbf{g}_{n_j,j}^{(q,q',p)} \right), \tag{8.4}$$

*and where the infinite sums converge in the* $L^1(\Omega)$ *sense. The function* $\mathbf{g}_{n_j,j}^{(q,q',p)}(\xi)$, $\xi = (\xi_1, \ldots, \xi_{q+q'-2p}) \in \mathbb{R}^{q+q'-2p}$, *in (8.4) is defined as follows:*

$$\begin{aligned} \mathbf{g}_{n_j,j}^{(q,q',p)}(\xi) \quad = \quad & D_{n_j}(\gamma_j\{\xi_1 + \cdots + \xi_{q+q'-2p}\}) \times \textstyle\prod_{i=1}^{q+q'-2p} [\sqrt{f(\xi_i)} \mathbb{1}_{(-\pi,\pi)}(\xi_i)] \\ & \times \widehat{\kappa}_j^{(p)}(\xi_1 + \cdots + \xi_{q-p}, \xi_{q-p+1} + \cdots + \xi_{q+q'-2p}), \end{aligned} \tag{8.5}$$

*where* $f$ *denotes the spectral density of the underlying Gaussian process* $X$ *and for any integer* $n$,

$$D_n(u) = \frac{1}{n_j} \sum_{k=0}^{n_j-1} \mathrm{e}^{\mathrm{i}ku} = \frac{1 - \mathrm{e}^{\mathrm{i}n_j u}}{n_j(1 - \mathrm{e}^{\mathrm{i}u})}, \tag{8.6}$$

*denotes the normalized Dirichlet kernel, and for* $\xi_1, \xi_2 \in \mathbb{R}$, *if* $p \neq 0$,

$$\begin{aligned} \widehat{\boldsymbol{\kappa}}_j^{(p)}(\xi_1, \xi_2) = \int_{(-\pi,\pi)^p} & \left( \prod_{i=1}^{p} f(\lambda_i) \right) \widehat{\mathbf{h}}_j^{(K)}(\lambda_1 + \cdots + \lambda_p + \xi_1) \\ & \times \overline{\widehat{\mathbf{h}}_j^{(K)}(\lambda_1 + \cdots + \lambda_p - \xi_2)} \, \mathrm{d}^p \lambda, \end{aligned} \tag{8.7}$$

*and, if* $p = 0$,

$$\widehat{\boldsymbol{\kappa}}_j^{(p)}(\xi_1, \xi_2) = \widehat{\mathbf{h}}_j^{(K)}(\xi_1) \overline{\widehat{\mathbf{h}}_j^{(K)}(\xi_2)}. \tag{8.8}$$

The random summand $\mathbf{S}_{n_j,j}^{(q,q',p)}$ is expressed in (8.4) as a Wiener–Itô integral of order $q + q' - 2p$ and $q + q' - 2p$ will be called the *order* of $\mathbf{S}_{n_j,j}^{(q,q',p)}$.

The limits involved in Theorem 3.1 are those given by the term $\mathbf{S}_{n_j,j}^{(q_0,q_0,q_0-1)}$ as proved in Propositions 5.3 and 5.4 of [10]. A sufficient condition to get the reduction principle is that the other terms are negligible with respect to this term. Theorem 3.2 is then a direct consequence of the following main result:

**Theorem 8.2.** *Suppose that Assumptions A hold with* $M \geq K + \delta(q_0)$, *where* $\delta(\cdot)$ *is defined in (2.6) and that (3.5) holds. Define the centered multivariate scalogram* $\overline{\mathbf{S}}_{n,j}$ *related to* $Y$ *by (2.16). Suppose that* $(\gamma_j)$ *and* $(n_j)$ *are any diverging sequences of integers. Then Condition (3.6) implies, as* $j \to \infty$,

$$\left\| \overline{\mathbf{S}}_{n_j,j} - \mathbf{S}_{n_j,j}^{(q_0,q_0,q_0-1)} \right\|_2 \ll \| \mathbf{S}_{n_j,j}^{(q_0,q_0,q_0-1)} \|_2. \tag{8.9}$$

*Proof.* Theorem 8.2 is proved in Section 9.4. □

*Idea behind the proof of Theorem 8.2.* One uses the expansion (8.3). The norms of the relevant terms are bounded in Proposition 8.3. We then deduce bounds for the difference $\|\overline{\mathbf{S}}_{n_j,j} - \mathbf{S}_{n_j,j}^{(q_0,q_0,q_0-1)}\|_2$ in Proposition 8.4. The main task in the proof of Theorem 8.2 is to show that these bounds are negligible compared to the leading term $\|\mathbf{S}_{n_j,j}^{(q_0,q_0,q_0-1)}\|_2$ whose asymptotic behavior is also given in Proposition 8.4.

Our results are based on $L^2(\Omega)$ upper bounds of the terms $\|\mathbf{S}_{n_j,j}^{(q,q',p)}\|_2$ established in Proposition 5.1 of [10]. To recall this result, we introduce some notations.

For any $s \in \mathbb{Z}_+$ and $d \in (0, 1/2)$, set

$$\Lambda_s(a) = \prod_{i=1}^{s} (a_i!)^{1-2d}, \quad \forall a = (a_1, \ldots, a_s) \in \mathbb{N}^s. \tag{8.10}$$

For any $q, q', p \geq 0$, define $\alpha, \beta$ and $\beta'$ as follows:

$$\alpha(q, q', p) = \begin{cases} \min\left(1 - \delta_+(q-p) - \delta_+(q'-p), 1/2\right) & \text{if } p \neq 0, \\ \frac{1}{2} & \text{if } p = 0, \end{cases} \tag{8.11}$$

$$\beta(q, p) = \max\left(\delta_+(p) + \delta_+(q-p) - 1/2, 0\right), \tag{8.12}$$

$$\beta'(q, q', p) = \max\left(2\delta_+(p) + \delta_+(q-p) + \delta_+(q'-p) - 1, -1/2\right). \tag{8.13}$$

Notice that for any $q \geq 0$, $\beta(q, 0) = \delta_+(q)$ and that, by definition of $\beta, \beta'$, we have, for all $0 \leq p \leq q \leq q'$, we have

$$\beta'(q, q', p) \leq \beta(q, p) + \beta(q', p). \tag{8.14}$$

Define the function $\varepsilon$ on $\mathbb{Z}_+$ as

$$\varepsilon(p) = \begin{cases} 0 & \text{if for any } s \in \{1, \ldots, p\},\ s(1-2d) \neq 1, \\ 1 & \text{if for some } s \in \{1, \ldots, p\},\ s(1-2d) = 1. \end{cases} \tag{8.15}$$

We first recall Proposition 5.1 of [10] where Part (i) corresponds to $p \geq 1$ and Part (ii) to $p = 0$.

**Proposition 8.3.** *Suppose that Assumptions A hold.*

*(i) There exists $C > 0$ such that for for all $n, \gamma_j \geq 2$ and $1 \leq q \leq q'$ and $1 \leq p \leq \min(q, q'-1)$,*

$$\begin{aligned} \|\mathbf{S}_{n,j}^{(q,q',p)}\|_2 \quad \leq \quad & C^{\frac{q+q'}{2}} \Lambda_2(q-p, p)^{1/2} \Lambda_2(q'-p, p)^{1/2} \gamma_j^{2K} \\ & \times [n_j^{-\alpha(q,q',p)} \gamma_j^{\beta'(q,q',p)} + n_j^{-1/2} \gamma_j^{\beta(q,p)+\beta(q',p)}] \\ & \times (\log n_j)^{\varepsilon(q+q'-2p)} (\log \gamma_j)^{3\varepsilon(q')}. \end{aligned} \tag{8.16}$$

(ii) *Assume that* $M \geq K + \max(\delta_+(q), \delta_+(q'))$. *Then there exists some* $C > 0$ *such that for all* $n, \gamma_j \geq 2$ *and* $1 \leq q \leq q'$,

$$\|\mathbf{S}_{n_j,j}^{(q,q',0)}\|_2 \leq C^{\frac{q+q'}{2}} \Lambda_1(q)^{1/2} \Lambda_1(q')^{1/2} n_j^{-1/2} \gamma_j^{2K+\delta_+(q)+\delta_+(q')} (\log \gamma_j)^{\varepsilon(q')} . \tag{8.17}$$

Note that under Condition (3.5) we have $\varepsilon(p) = 0$ for all $p \geq 1$ in (8.15). Thus the logarithmic terms vanish in (8.16) and (8.17). Moreover, if $p = 0$ then $\Lambda_2(q, 0) = \Lambda_1(q)$, $\alpha(q, q', 0) = 1/2$, $\beta(q, 0) = \delta_+(q)$ and $\beta'(q, q', 0) = \delta_+(q) + \delta_+(q')$. Therefore, if Condition (3.5) holds, the bounds (8.16) and (8.17) imply the following common bound

$$\|\mathbf{S}_{n,j}^{(q,q',p)}\|_2 \leq C^{\frac{q+q'}{2}} \Lambda_2(q-p,p)^{1/2} \Lambda_2(q'-p,p)^{1/2} \gamma_j^{2K} \times [n_j^{-\alpha(q,q',p)} \gamma_j^{\beta'(q,q',p)} + n_j^{-1/2} \gamma_j^{\beta(q,p)+\beta(q',p)}]. \tag{8.18}$$

Consider now the decomposition

$$\overline{\mathbf{S}}_{n_j,j} = \mathbf{S}_{n_j,j}^{(q_0,q_0,q_0-1)} + \left(\overline{\mathbf{S}}_{n_j,j} - \mathbf{S}_{n_j,j}^{(q_0,q_0,q_0-1)}\right).$$

The following result provides the sharp rate of the first term and a bound on the second one, relying on Wiener chaos decomposition (8.3).

**Proposition 8.4.** *Assume that Assumptions A hold with* $M \geq K + \delta_+(q_0)$ *and suppose that Condition (3.5) holds. Let* $(n_j)$ *and* $(\gamma_j)$ *be any diverging sequences. Then, there exists a positive constant* $C$ *such that, for all* $j \geq 1$,

$$\left\|\overline{\mathbf{S}}_{n_j,j} - \mathbf{S}_{n_j,j}^{(q_0,q_0,q_0-1)}\right\|_2 \leq C\ \gamma_j^{2K} \sup_{(q,q',p)\in\mathcal{A}_0} [n_j^{-\alpha(q,q',p)} \gamma_j^{\beta'(q,q',p)} + n_j^{-1/2} \gamma_j^{\beta(q,p)+\beta(q',p)}], \tag{8.19}$$

*where we denote*

$$\mathcal{A}_0 = \{(q, q', p) : 1 \leq q \leq q',\ 0 \leq p \leq \min(q, q'-1),\ c_q \times c_{q'} \neq 0\} \backslash \{(q_0, q_0, q_0-1)\}. \tag{8.20}$$

*Moreover, the two following assertions hold:*

(i) *If* $q_0 = 1$, *as* $j \to \infty$,

$$\|\mathbf{S}_{n_j,j}^{(q_0,q_0,q_0-1)}\|_2 = \|\mathbf{S}_{n_j,j}^{(1,1,0)}\|_2 \sim C\ n_j^{-1/2} \gamma_j^{2(d+K)}, \tag{8.21}$$

*where* $C$ *is a positive constant.*

(ii) *If* $q_0 \geq 2$, *as* $j \to \infty$,

$$\|\mathbf{S}_{n_j,j}^{(q_0,q_0,q_0-1)}\|_2 \sim C n_j^{-1+2d} \gamma_j^{2(\delta(q_0)+K)}, \tag{8.22}$$

*where* $C$ *is a positive constant.*

*Proof.* By Proposition 8.1 and (8.20), applying the Minkowski inequality, we have

$$\left\|\overline{\mathbf{S}}_{n_j,j} - \mathbf{S}_{n_j,j}^{(q_0,q_0,q_0-1)}\right\|_2 \leq 2 \sum_{(q,q',p)\in\mathcal{A}_0} \frac{|c_q|}{q!}\frac{|c_{q'}|}{q'!}\, p! \binom{q}{p}\binom{q'}{p} \|\mathbf{S}_{n_j,j}^{(q,q',p)}\|_2.$$

The bound (8.18) implies that

$$\begin{aligned}&\sum_{(q,q',p)\in\mathcal{A}_0} \frac{|c_q|}{q!}\frac{|c_{q'}|}{q'!}\, p! \binom{q}{p}\binom{q'}{p} \|\mathbf{S}_{n_j,j}^{(q,q',p)}\|_2\\ &\leq \left(\sum_{(q,q',p)\in\mathcal{A}_0} \frac{|c_q|}{q!}\frac{|c_{q'}|}{q'!}\, p! \binom{q}{p}\binom{q'}{p} C^{\frac{q+q'}{2}} \Lambda_2(q-p,p)^{1/2}\Lambda_2(q'-p,p)^{1/2}\right)\\ &\times \gamma_j^{2K} \sup_{(q,q',p)\in\mathcal{A}_0} [n_j^{-\alpha(q,q',p)}\gamma_j^{\beta'(q,q',p)} + n_j^{-1/2}\gamma_j^{\beta(q,p)+\beta(q',p)}].\end{aligned}$$

By Lemma 8.6 of [10], the last two displays yield (8.19).

We now prove (8.21) and (8.22). First consider the case where $q_0 = 1$. This asymptotic equivalence (8.21) is related to the convergence (3.1) and follows from its proof, see e.g. [20]. Since $q_0 = 1$, we have $c_1 \neq 0$. Moreover in Condition (W-3) on the wavelet filters recalled in Appendix B, $\widehat{h}_{\ell,\infty}$ are functions that are non-identically zero and which are continuous as locally uniform limits of continuous functions. Therefore $\sum_\ell \Gamma_{\ell,\ell}^2 > 0$ and we get (8.21).

Now consider the case where $q_0 \geq 2$. The bound (8.22) is then related to Theorem 3.2((b)) where the weak convergence is stated and follows from its proof, see [11]. □

## 9. Proofs

### 9.1. *Proof of Theorem 2.1*

The generalized spectral density $f_{G,K}$ of $Y$ is related to the spectral density $f_G$ of $G(X)$ by (1.5). By definition of $d_0$, the result shall then follow if we prove the existence of a bounded function $f_G^*$ such that

$$f_G(\lambda) = |1 - \mathrm{e}^{-\mathrm{i}\lambda}|^{-2\delta(q_0)}\, f_G^*(\lambda), \tag{9.1}$$

and satisfying all the properties stated in Theorem 2.1.

We now prove (9.1). To this end, we consider the following decomposition of $G(X)$ as the sum of two uncorrelated processes,

$$G(X) = G_1(X) + G_2(X) =: \sum_{1\leq q<1/(1-2d)} \frac{c_q}{q!} H_q(X) + \sum_{q\geq 1/(1-2d)} \frac{c_q}{q!} H_q(X).$$

The proof of Proposition 6.2 in [9] shows that $G_2(X)$ admits a bounded spectral density $f_{G_2}$. We first consider the case where $G_1$ reduces to the term $c_{q_0}H_{q_0}/q_0!$.

Since the two processes $H_{q_0}(X)$ and $G_2(X)$ are uncorrelated, one has

$$f_G(\lambda) = \frac{c_{q_0}^2}{q_0!} f_{H_{q_0}}(\lambda) + f_{G_2}(\lambda).$$

We can then set

$$f_G^*(\lambda) = \frac{c_{q_0}^2}{q_0!} f_{H_{q_0}}^*(\lambda) + |1 - \mathrm{e}^{-\mathrm{i}\lambda}|^{2\delta(q_0)} f_{G_2}(\lambda).$$

Let us check $f_G^*$ has the properties stated in the theorem. Relation (9.1) follows from the definition of $f_G^*$ and $f_{H_{q_0}}^*$. To prove the other properties stated in Theorem 2.1, we distinguish the two cases $q_0 = 1$ and $q_0 \geq 2$. If $q_0 = 1$, $f_{H_{q_0}} = f$, $f_{H_{q_0}}^* = f^*$ and then $\zeta \leq \beta$, one has

$$|f_{H_q}^*(\lambda) - f_{H_q}^*(0)| \leq C|\lambda|^\zeta, \tag{9.2}$$

for some $C > 0$. If $q_0 \geq 2$, Lemma 10.1 yields that there exists a bounded function $f_{H_{q_0}}^*$ such that

$$f_{H_q}(\lambda) = |1 - \mathrm{e}^{-\mathrm{i}\lambda}|^{-2\delta(q)} f_{H_q}^*(\lambda) \tag{9.3}$$

Moreover for any $\zeta \in (0, 2\delta(q))$ such that $\zeta \leq \beta$, one has

$$|f_{H_q}^*(\lambda) - f_{H_q}^*(0)| \leq C|\lambda|^\zeta, \tag{9.4}$$

for some $C > 0$. In any case, the boundedness of $f_{G_2}$ and the properties of $f_{H_{q_0}}^*$ (equation (9.2) if $q_0 = 1$ or (9.3), (9.4) if $q_0 \geq 2$) then imply that (2.13) holds in the case $G_1 = c_{q_0} H_{q_0}/q_0!$, that is if $\delta_+(q_1) = 0$.

We now deal with the case where $H_{q_1}$ has also long memory, namely $\delta_+(q_1) > 0$. Since the terms $H_q(X)$ for $q < 1/(1-2d)$ are all pairwise uncorrelated, the spectral density of long-range dependent part $G_1(X)$ reads as follows

$$f_{G_1}(\lambda) = \sum_{1 \leq q < 1/(1-2d)} \frac{c_q^2}{q!} f_{H_q}(\lambda).$$

We now apply Equation (10.2) of Lemma 10.1 successively to each $q < 1/(1-2d)$. Hence

$$f_{G_1}(\lambda) = |1 - \mathrm{e}^{-\mathrm{i}\lambda}|^{-2\delta(q_0)} \left[ \sum_{1 \leq q < 1/(1-2d)} \frac{c_q^2}{q!} |1 - \mathrm{e}^{-\mathrm{i}\lambda}|^{2\delta(q_0) - 2\delta(q)} f_{H_q}^*(\lambda) \right].$$

Since $f_G = f_{G_1} + f_{G_2}$, we then get (9.1) with

$$f_G^*(\lambda) = \left[ \sum_{1 \leq q < 1/(1-2d)} \frac{c_q^2}{q!} |1 - \mathrm{e}^{-\mathrm{i}\lambda}|^{2\delta(q_0) - 2\delta(q)} f_{H_q}^*(\lambda) \right] + |1 - \mathrm{e}^{-\mathrm{i}\lambda}|^{2\delta(q_0)} f_{G_2}(\lambda).$$

Since $|1-\mathrm{e}^{-\mathrm{i}\lambda}|=0$ for $\lambda=0$, we have $f_G^*(0)=c_{q_0}^2 f_{H_{q_0}}^*(0)/q_0!$. We now prove that under Condition (2.12) on $\zeta$, we get (2.13). Indeed,

$$\begin{aligned}|f_G^*(\lambda)-f_G^*(0)| \le \frac{c_{q_0}^2}{q_0!}&|f_{H_{q_0}}^*(\lambda)-f_{H_{q_0}}^*(0)|\\ &+\left[\sum_{q_1\le q<1/(1-2d)}\frac{c_q^2}{q!}|1-\mathrm{e}^{-\mathrm{i}\lambda}|^{2\delta(q_0)-2\delta(q)} f_{H_q}^*(\lambda)\right]\\ &+|1-\mathrm{e}^{-\mathrm{i}\lambda}|^{2\delta(q_0)} f_{G_2}(\lambda).\end{aligned}$$

Using the boundedness of $f_{H_q}^*$ for any $q\ge q_1$, we deduce that for some $C>0$ and any $q\ge q_1$,

$$|1-\mathrm{e}^{-\mathrm{i}\lambda}|^{2\delta(q_0)-2\delta(q)} f_{H_q}^*(\lambda)\le C|\lambda|^{2\delta(q_0)-2\delta(q_1)}, \tag{9.5}$$

whereas by Lemma 10.1 applied with $q=q_0$, we deduce that for any $\zeta\in(0,2\delta(q_0))$ such that $\zeta\le\beta$, one has

$$|f_{H_{q_0}}^*(\lambda)-f_{H_{q_0}}^*(0)|\le L|\lambda|^{\zeta}. \tag{9.6}$$

We now combine (9.5) and (9.6) and deduce that for any $\zeta\in(0,2\delta(q_0))$ such that $\zeta\le\min(\beta,2\delta(q_0)-2\delta(q_1))$, one has

$$|f_G^*(\lambda)-f_G^*(0)|\le L'|\lambda|^{\zeta}, \tag{9.7}$$

for some $L'>0$.

### 9.2. *Proof of Theorem 2.2*

The bound (2.22) follows the same lines as the proof of Proposition 8.4. It is a consequence of Proposition 8.1, Proposition 8.3, Lemma 8.6 of [10] and of the following bounds:

$$\begin{aligned}\alpha(q,q',p)&\ge 1/2-d\quad\text{and equality implies } q'=q+1 \text{ and } p=q,\\ \beta(q,p)&\le\delta(q_0)\quad\text{and equality implies } q=q_0,\\ \beta'(q,q',p)&\le 2\delta(q_0)\quad\text{and equality implies } q=q'=q_0.\end{aligned}$$

The two first bounds follow from Lemma 8.3 in [10], and the last one from (8.14). The equality cases are used to get rid of the logarithmic corrections appearing in (8.16) and (8.17) since $q'=q+1$ and $p=q$ imply $\varepsilon(q+q'-2p)=\varepsilon(1)=0$ and $q=q'=q_0$ implies $\varepsilon(q')=\varepsilon(q_0)=0$. This concludes the proof.

### 9.3. *Proof of Proposition 4.1*

We want to show we always have $\nu_c>0$. By definition of $\delta$ and $\delta_+$ in (2.6), we have $\delta(q)\le\delta(1)=d$ for all $q\ge 1$, and since $d>0$, $\delta_+(q)\le d$. With $d<1/2$, this implies that $d+1/2-2\delta_+(q_{\ell_0})\ge 1/2-d>0$ and thus the third line of the definition of $\nu_c$ is positive. For the same reason, $1-2\delta_+(q_1-1) \geq 1-2d$ and the fourth line of the definition of $\nu_c$ is positive. In addition, for any $r \geq 0$, $2d+1/2-2\delta_+(q_{\ell_r})-\delta(r+1) \geq 1/2-d$ and $\delta(r+1) < 1/2$ so that, for any $r \in \mathcal{R}$,

$$\frac{2d+1/2-2\delta_+(q_{\ell_r})-\delta(r+1)}{\delta(r+1)} \geq 1-2d > 0,$$

which ensures that the quantities inside the min are uniformly lower-bounded by a positive value. Finally, for the last line, we separate the cases $\delta_+(q_{\ell_0}) = 0$ and $\delta_+(q_{\ell_0}) = \delta(q_{\ell_0})$. In the first case, we have $4(\delta(q_0)-\delta_+(q_{\ell_0})) = 4\delta(q_0) = 2-2q_0(1-2d)$ and so

$$\frac{4(\delta(q_0)-\delta_+(q_{\ell_0}))}{1-2d} = 2(1/(1-2d)-q_0) > 0,$$

as a consequence of (2.10). In the second case, we have $4(\delta(q_0)-\delta_+(q_{\ell_0})) \geq 4(\delta(q_0)-\delta(q_{\ell_0})) = 2(q_{\ell_0}-q_0)(1-2d)$ and so

$$\frac{4(\delta(q_0)-\delta_+(q_{\ell_0}))}{1-2d} \geq 2(q_{\ell_0}-q_0) \geq 0,$$

which is non-negative by definition of $q_{\ell_0}$. Hence the last line defining $\nu_c$ is at least one, hence is positive, which concludes the proof.

### *9.4. Proof of Theorem 8.2*

By Proposition 8.4, it is sufficient to show that the right-hand side of (8.19) is negligible with respect to the right-hand side of (8.21) if $q_0 = 1$ or to the right-hand side of (8.22) if $q_0 \geq 2$, that is, respectively,

$$\lim_{j\to\infty} n_j^{1/2}\gamma_j^{-2d} \sup_{(q,q',p)\in\mathcal{A}_0} [n_j^{-\alpha(q,q',p)}\gamma_j^{\beta'(q,q',p)} + n_j^{-1/2}\gamma_j^{\beta(q,p)+\beta(q',p)}] = 0, \quad (9.8)$$

$$\lim_{j\to\infty} n_j^{1-2d}\gamma_j^{-2\delta(q_0)} \sup_{(q,q',p)\in\mathcal{A}_0} [n_j^{-\alpha(q,q',p)}\gamma_j^{\beta'(q,q',p)} + n_j^{-1/2}\gamma_j^{\beta(q,p)+\beta(q',p)}] = 0. \quad (9.9)$$

We now distinguish the two cases $q_0 = 1$, $q_0 \geq 2$.

#### *9.4.1. Proof of Theorem 8.2 in the case $q_0 = 1$*

In this case, we need to show that Condition (3.6) implies (9.8).

By Lemma 8.3 (4) in [10], we have, for all $0 \leq p \leq q \leq q'$, $\beta(q,p)+\beta(q',p) \leq \delta_+(q)+\delta_+(q')$. We may thus write

$$\begin{aligned}\sup_{(q,q',p)\in\mathcal{A}_0} \gamma_j^{\beta(q,p)+\beta(q',p)} &\leq \sup_{(q,q',p)\in\mathcal{A}_0} \gamma_j^{\delta_+(q)+\delta_+(q')} \\ &\leq \gamma_j^{\sup\left\{\delta^+(q)+\delta^+(q')\colon\, 1\leq q\leq q',\, (q,q')\neq(1,1)\right\}},\end{aligned}$$

since for $q_0 = 1$, the triplet $(q_0, q_0, q_0 - 1) = (1, 1, 0)$ is excluded from $\mathcal{A}_0$. Using Lemma 10.2, we obtain, as $j \to \infty$,

$$\sup_{(q,q',p)\in\mathcal{A}_0} n_j^{-1/2} \gamma_j^{\beta(q,p)+\beta(q',p)} = o\left(n_j^{-1/2}\gamma_j^{2d}\right). \tag{9.10}$$

Inserting this in (9.8), we only need to show that (3.6) implies

$$\lim_{j\to\infty} n_j^{1/2}\gamma_j^{-2d} \sup_{(q,q',p)\in\mathcal{A}_0} n_j^{-\alpha(q,q',p)}\gamma_j^{\beta'(q,q',p)} = 0. \tag{9.11}$$

Observe that, by definition, $\alpha(q, q', p) \leq 1/2$. We shall therefore partition $\mathcal{A}_0$ into $\mathcal{A}_0 = \mathcal{A}_1 \cup \mathcal{A}_2$, where

$$\mathcal{A}_1 = \{(q, q', p) \in \mathcal{A}_0 : \alpha(q, q', p) = 1/2\}$$
$$\mathcal{A}_2 = \{(q, q', p) \in \mathcal{A}_0 : \alpha(q, q', p) < 1/2\}.$$

Since $\alpha(q, q', p) = 1/2$ for $(q, q', p) \in \mathcal{A}_1$, we get with (8.14) that

$$\sup_{(q,q',p)\in\mathcal{A}_1} n_j^{-\alpha(q,q',p)}\gamma_j^{\beta'(q,q',p)} \leq \sup_{(q,q',p)\in\mathcal{A}_0} n_j^{-1/2}\gamma_j^{\beta(q,p)+\beta(q',p)} = o\left(n_j^{-1/2}\gamma_j^{2d}\right),$$

as $j \to \infty$ by (9.10).

If $\mathcal{A}_2 = \emptyset$ we conclude that (9.11) holds. By Lemma 10.4, we note that $\mathcal{A}_2 = \emptyset$ if and only if $d \leq 1/4$ and $I_0$ defined by (4.2) is an empty set. Hence, from now on, we assume that $\mathcal{A}_2 \neq \emptyset$, that is, either $d \leq 1/4$ and $I_0 \neq \emptyset$, or $d > 1/4$. It only remains to show that, under these conditions, (3.6) implies

$$\lim_{j\to\infty} n_j^{1/2}\gamma_j^{-2d} \sup_{(q,q',p)\in\mathcal{A}_2} n_j^{-\alpha(q,q',p)}\gamma_j^{\beta'(q,q',p)} = 0. \tag{9.12}$$

To compute the sup, we first optimize on $p$, then on $q'$, and finally on $q$.

**Optimization on $p$** By Lemma 10.3, if $(q, q', p) \in \mathcal{A}_2$ for a given $(q, q')$, then $\alpha(q, q', p)$ is minimal and $\beta'(q, q', p)$ is maximal for the largest possible $p$, which corresponds to $p = q - 1$ if $q' = q$ and to $p = q$ if $q' > q$. For such a $p$, we have, if $q = q'$,

$$\alpha(q, q', p) = \alpha(q, q, q - 1) = \min(1 - 2d, 1/2),$$

and if $q' > q$,

$$\alpha(q, q', p) = \alpha(q, q', q) = 1/2 - \delta_+(q' - q).$$

Since being in $\mathcal{A}_2$ implies $\alpha(q, q', p) < 1/2$, we must have $1 - 2d < 1/2$ (that is $d > 1/4$) if $q = q'$ and $\delta(q' - q) > 0$ if $q' > q$. To separate the cases $q = q'$ and $q \neq q'$, we define

$$\mathcal{A}_{2,1} = \begin{cases} \emptyset & \text{if } d \leq 1/4, \\ \{(q, q, q-1) : q \geq 2,\ c_q \neq 0\} & \text{if } d > 1/4. \end{cases}$$

and

$$\mathcal{A}_{2,2} = \{(q, q', q) : q' > q \geq 1,\ c_q c_{q'} \neq 0,\ \delta(q' - q) > 0\}.$$

Note that in $\mathcal{A}_{2,1}$ we set $q \geq 2$ to avoid $(q, q, q-1) = (1, 1, 0)$. Recall that the indices of the non-zero coefficients $c_q$ are labeled as $q_\ell$, see (4.1). Then

$$\{(q, q, q-1) : q \geq 2,\ c_q \neq 0\} = \{(q_\ell, q_\ell, q_\ell - 1) : \ell \in \mathcal{L},\ \ell \geq 1\},$$

and, similarly,

$$\mathcal{A}_{2,2} = \{(q_\ell, q_{\ell'}, q_\ell) : \ell, \ell' \in \mathcal{L},\ 0 \leq \ell < \ell',\ \delta(q_{\ell'} - q_\ell) > 0\}.$$

Defining

$$A_j := \sup_{(\ell,\ell')} n_j^{-1/2+\delta(q_{\ell'}-q_\ell)} \gamma_j^{\beta'(q_\ell, q_{\ell'}, q_\ell)}, \tag{9.13}$$

where the $\sup_{\ell,\ell'}$ is taken over $(\ell, \ell') \in \mathcal{L}^2$ such that $\ell < \ell'$ and $\delta(q_{\ell'} - q_\ell) > 0$, and

$$B_j := \sup_\ell \gamma_j^{\beta'(q_\ell, q_\ell, q_\ell - 1)}, \tag{9.14}$$

where the $\sup_\ell$ is taken over all $\ell \in \mathcal{L}$ such that $\ell \geq 1$, we thus obtain the two following assertions.

- If $d \leq 1/4$, the sup over $\mathcal{A}_2$ can be restricted to $\mathcal{A}_{22}$. This gives

$$\sup_{(q,q',p)\in\mathcal{A}_2} n_j^{-\alpha(q,q',p)} \gamma_j^{\beta'(q,q',p)} = A_j, \tag{9.15}$$

- If $d > 1/4$, the sup over $\mathcal{A}_2$ has to be performed over $\mathcal{A}_{21}$ and $\mathcal{A}_{22}$. This gives

$$\sup_{(q,q',p)\in\mathcal{A}_2} n_j^{-\alpha(q,q',p)} \gamma_j^{\beta'(q,q',p)} = \max\left(n_j^{-1+2d} B_j, A_j\right). \tag{9.16}$$

**Optimization on $q'$** We only need to consider $A_j$ since $B_j$ corresponds to $q' = q$. For $A_j$, optimizing on $q'$ means optimizing on $\ell'$ in the sup of (9.15). We know from Lemma 10.3 that, for each $\ell$, $\alpha(q_\ell, q_{\ell'}, q_\ell)$ is non-decreasing and $\beta'(q_\ell, q_{\ell'}, q_\ell)$ is non-increasing as $\ell'$ increases, hence the $\sup_{\ell,\ell'}$ is achieved when $\ell' = \ell + 1$ and thus $\alpha(q_\ell, q_{\ell'}, q_\ell) < 1/2$ implies

$$A_j = \begin{cases} \sup_{\ell \in J_d} n_j^{-1/2+\delta(q_{\ell+1}-q_\ell)} \gamma_j^{\beta'(q_\ell, q_{\ell+1}, q_\ell)} & \text{if } J_d \neq \emptyset, \\ 0 & \text{otherwise}, \end{cases} \tag{9.17}$$

where $J_d$ is defined in (4.8). When $J_d = \emptyset$, the sup in (9.13) is taken over the empty set. We use the convention $\sup_\emptyset(\dots) = 0$.

**Optimization on $q$** We deal separately with the cases

(a) $d \leq 1/4$.
(b) $d > 1/4$.

The case (a) is the simplest since in (9.15), $B_j$ does not appear. Recall also that we have $I_0 \neq \emptyset$ in this case since we assumed $\mathcal{A}_2 \neq \emptyset$. The optimization on $q$ here amounts to optimize $A_j$ on $\ell$ in (9.17) in the case $J_d \neq \emptyset$. Observe that when $d \leq 1/4$, $\delta(r) = 0$ for all $r \geq 2$, see (2.6). Thus the condition $\delta(q_{\ell+1} - q_\ell) > 0$ on $\ell \in J_d$ is equivalent to $q_{\ell+1} = q_\ell + 1$, in which case $\delta(q_{\ell+1} - q_\ell) = \delta(1) = d > 0$ and $J_d = I_0$. Hence,

$$A_j = \sup_{\ell \in I_0} n_j^{-1/2+d} \gamma_j^{\beta'(q_\ell, q_\ell+1, q_\ell)} = \sup_{\ell \in I_0} n_j^{-1/2+d} \gamma_j^{2\delta_+(q_\ell)+d-1/2},$$

where we used that $\beta'(q_\ell, q_\ell+1, q_\ell) = \max(2\delta_+(q_\ell)+d-1/2, -1/2) = 2\delta_+(q_\ell)+d-1/2$, see (8.13). This sup is achieved for the smallest $\ell$ since $\delta_+$ is non-increasing. Recall that the smallest $\ell$ in $I_0$ is denoted by $\ell_0$ in (4.3), thus,

$$A_j = n_j^{-1/2+d} \gamma_j^{2\delta_+(q_{\ell_0})+d-1/2}.$$

Now, we note that in case (a) with $q_0 = 1$, $\nu_c$ in Definition 4.1 takes value

$$\nu_c = \frac{d + 1/2 - 2\delta_+(q_{\ell_0})}{d}.$$

Hence Condition (3.6) implies

$$n_j^{1/2} \gamma_j^{-2d} A_j = n_j^d \gamma_j^{2\delta_+(q_{\ell_0})-d-1/2} = o(1).$$

With (9.15), we obtain (9.12) and case (a) is complete.

We now turn to the case (b), that is, we assume now that $d > 1/4$ and show that (9.12) holds under Condition (3.6). Optimizing $B_j$ on $q$ amounts to optimizing the sup in (9.14) on $\ell \in \mathcal{L}$ with $\ell \geq 1$. Note that $\beta'(q_\ell, q_\ell, q_\ell - 1) = \max(2\delta_+(q_\ell - 1) + 2d - 1, -1/2) = 2\delta_+(q_\ell - 1) + 2d - 1$ which is non-increasing as $\ell$ increases. Hence the sup in (9.14) is achieved for $\ell = 1$ and thus

$$B_j = \gamma_j^{2\delta_+(q_1-1)+2d-1}. \tag{9.18}$$

Note that in this case $\nu_c$ in Definition 4.1 takes value

$$\nu_c = \begin{cases} \frac{1-2\delta_+(q_1-1)}{2d-1/2} & \text{if } J_d = \emptyset, \\ \min\left(\frac{1-2\delta_+(q_1-1)}{2d-1/2}, \min_{r \in \mathcal{R}} \left(\frac{2d + 1/2 - 2\delta_+(q_{\ell_r}) - \delta(r+1)}{\delta(r+1)}\right)\right) & \text{if } J_d \neq \emptyset. \end{cases} \tag{9.19}$$

In both cases, we have $\nu_c \leq (1-2\delta_+(q_1-1))/(2d-1/2)$, and thus Condition (3.6) implies, as $j \to \infty$,

$$\gamma_j^{-2d} n_j^{-1/2+2d} B_j = n_j^{-1/2+2d} \gamma_j^{2\delta_+(q_1-1)-1} = o(1).$$

If $J_d = \emptyset$ so that $A_j = 0$ in (9.17), we thus obtain with (9.16) that (3.6) implies (9.12). Similarly, if $J_d \neq \emptyset$, which we now assume, it only remains to prove that (3.6) implies

$$\lim_{j \to \infty} n_j^{1/2} \gamma_j^{-2d} A_j = 0. \tag{9.20}$$

Using (9.17) and that $\beta'(q_\ell, q_{\ell+1}, q_\ell) = \max(2\delta_+(q_\ell) + 1/2 + \delta(q_{\ell+1} - q_\ell) - 1, 1/2) = 2\delta_+(q_\ell) + \delta(q_{\ell+1} - q_\ell) - 1/2$, we have

$$A_j = \sup_{\ell \in J_d} n_j^{-1/2+\delta(q_{\ell+1}-q_\ell)} \gamma_j^{2\delta_+(q_\ell)+\delta(q_{\ell+1}-q_\ell)-1/2}. \tag{9.21}$$

Optimizing on $q$ here means optimizing this sup on $\ell \in J_d$. To do so, we partition $J_d$ as in (4.9) and, by the definition of $I_r$ in (4.4), we have, for all $\ell \in I_r$,

$$n_j^{-1/2+\delta(q_{\ell+1}-q_\ell)} \gamma_j^{2\delta_+(q_\ell)+\delta(q_{\ell+1}-q_\ell)-1/2} = n_j^{-1/2+\delta(r+1)} \gamma_j^{2\delta_+(q_\ell)+\delta(r+1)-1/2}.$$

Since $\delta$ is non-increasing, we get with Definition (4.5) that, for all $r \in \mathcal{R}$,

$$\sup_{\ell \in I_r} n_j^{-1/2+\delta(q_{\ell+1}-q_\ell)} \gamma_j^{2\delta_+(q_\ell)+\delta(q_{\ell+1}-q_\ell)-1/2} = n_j^{-1/2+\delta(r+1)} \gamma_j^{2\delta_+(q_{\ell_r})+\delta(r+1)-1/2}.$$

Hence, by (9.21) and (4.9), we get that

$$n_j^{1/2} \gamma_j^{-2d} A_j = \max_{r \in \mathcal{R}} n_j^{\delta(r+1)} \gamma_j^{2\delta_+(q_{\ell_r})+\delta(r+1)-1/2-2d}.$$

Now, since we are in the case $J_d \neq \emptyset$, $\nu_c$ in (9.19) satisfies $\nu_c \leq (2d + 1/2 - 2\delta_+(q_{\ell_r}) - \delta(r+1))/(\delta(r+1))$ for all $r \in \mathcal{R}$ and recalling that $\mathcal{R}$ is a finite set (see (4.7)), we see that (3.6) implies (9.20). The proof of the case $q_0 = 1$ is concluded.

### *9.4.2. Proof of Theorem 8.2 in the case $q_0 \geq 2$*

In this case, we need to show that (3.6) implies (9.9).

Recall that in Assumptions A include Condition (2.10) and thus $q_0 \geq 2$ implies $d > 1/4$. Hence we have $1 - 2d < 1/2$ and since moreover for all $q' \geq q \geq q_0$ and $0 \leq p \leq q$, we have $\beta(q,p) + \beta(q',p) \leq 2\delta_+(q_0) = 2\delta(q_0)$, we obtain that

$$\lim_{j\to\infty} n_j^{1-2d} \gamma_j^{-2\delta(q_0)} \sup_{(q,q',p)\in\mathcal{A}_0} n_j^{-1/2} \gamma_j^{\beta(q,p)+\beta(q',p)} = 0.$$

This correspond to the second term between brackets in (9.9) and we thus only need to prove that (3.6) implies

$$\lim_{j\to\infty} n_j^{1-2d} \gamma_j^{-2\delta(q_0)} \sup_{(q,q',p)\in\mathcal{A}_0} n_j^{-\alpha(q,q',p)} \gamma_j^{\beta'(q,q',p)} = 0. \tag{9.22}$$

Let us partition $\mathcal{A}_0$ into $\mathcal{A}_0 = \cup_{i=1}^5 \mathcal{A}_i$, where

$$\begin{aligned}
\mathcal{A}_1 &= \{(q,q',p) \in \mathcal{A}_0,\, q = q' = q_0\},\\
\mathcal{A}_2 &= \{(q,q,p) \in \mathcal{A}_0,\, q = q' > q_0\},\\
\mathcal{A}_3 &= \{(q,q',p) \in \mathcal{A}_0,\, q' \geq q+2\},\\
\mathcal{A}_4 &= \{(q,q',p) \in \mathcal{A}_0,\, q' = q+1,\, p \leq q-1\},\\
\mathcal{A}_5 &= \{(q,q',p) \in \mathcal{A}_0,\, q = p,\, q' = q+1\}.
\end{aligned}$$

We shall prove that for $i = 1, 2, 3, 4$,

$$\lim_{j\to\infty} n_j^{1-2d} \gamma_j^{-2\delta(q_0)} \sup_{(q,q',p)\in\mathcal{A}_i} n_j^{-\alpha(q,q',p)} \gamma_j^{\beta(q,p)+\beta(q',p)} = 0, \tag{9.23}$$

and that, when $I_0$ defined as in (4.2) is not empty, (3.6) implies

$$\lim_{j\to\infty} n_j^{1-2d}\gamma_j^{-2\delta(q_0)} \sup_{\ell\in I_0} n_j^{-\alpha(q_\ell,q_\ell+1,q_\ell)}\gamma_j^{\beta'(q_\ell,q_\ell+1,q_\ell)} = 0. \tag{9.24}$$

Since $\beta'(q,q',p) \le \beta(q,p)+\beta(q',p)$ and $\mathcal{A}_5 = \{(q_\ell,q_\ell+1,q_\ell) : \ell \in I_0\}$, we indeed have that (9.23) and (9.24) imply (9.22) and the proof will be concluded.

The limit (9.23) can be deduced for $i = 1,\ldots,4$ from Lemma 8.3 of [10]. More precisely this lemma implies the following facts (recall that $d > 1/4$).

1. For all $p = 0,\ldots,q_0-2$, we have $\alpha(q_0,q_0,p) > 1-2d$ and $2\beta(q_0,p) \le 2\delta(q_0)$, which implies (9.23) for $i = 1$ since $(q_0,q_0,q_0-1)$ is excluded from $\mathcal{A}_0$.
2. For all $q \ge q_0+1$ and $p = 0,\ldots,q-1$, we have $\alpha(q,q,p) \ge 1-2d$ and $2\beta(q,p) \le 2\delta_+(q_0+1) < 2\delta(q_0)$, which implies (9.23) for $i = 2$.
3. For $q \ge q_0$, $q' \ge q+2$ and $p = 0,\ldots,q$, we have $\alpha(q,q',p) \ge 1-2d$ and $\beta(q,p)+\beta(q',p) \le \delta_+(q_0)+\delta_+(q_0+2) < 2\delta(q_0)$, which implies (9.23) for $i = 3$.
4. For $q \ge q_0$ and $p = 0,\ldots,q-1$, we have $\alpha(q,q+1,p) \ge \min(3/2(1-2d),1/2) > 1-2d$ and $\beta(q,p)+\beta(q+1,p) \le \delta_+(q_0+1)+\delta(q_0) < 2\delta(q_0)$, which implies (9.23) for $i = 4$.

Hence we obtain that (9.23) is valid for $i = 1,\ldots,4$. If $I_0$ is empty, the proof is concluded. We now assume that $I_0$ is not empty, so that $\ell_0$ is finite, and it only remains to show that Condition (3.6) implies (9.24). Observe that, for any $q \ge q_0$, we have $\alpha(q,q+1,q) = 1/2-d$ and $\beta'(q,q+1,q) = 2\delta_+(q)+d-1/2$ is non–increasing as $q$ increases. Hence over $\ell \in I_0$, $\alpha(q_\ell,q_\ell+1,q_\ell)$ is constant and $\beta'(q_\ell,q_\ell+1,q_\ell)$ is maximal at $\ell = \ell_0$, where it takes value $2\delta_+(q_{\ell_0})+d-1/2$. We conclude that

$$n_j^{1-2d}\gamma_j^{-2\delta(q_0)} \sup_{\ell\in I_0} n_j^{-\alpha(q_\ell,q_\ell+1,q_\ell)}\gamma_j^{\beta'(q_\ell,q_\ell+1,q_\ell)} = n_j^{1/2-d}\gamma_j^{d-1/2-2(\delta(q_0)-\delta_+(q_{\ell_0}))}.$$

Note that in this case $\nu_c$ in Definition 4.1 takes value

$$\nu_c = 1 + \frac{4(\delta(q_0)-\delta_+(q_{\ell_0}))}{1-2d}.$$

Thus Condition (3.6) implies (9.24) and the proof is finished.

### *9.5. Proof of Theorem 6.2*

The fact that the test $\delta_s$ is consistent follows directly from the consistency statement in Theorem 6.1 and the fact that $(u_N)$ is diverging.

To show that the test $\delta_s$ has asymptotic confidence level $\alpha$, it suffices to show that when $d_0 = d_0^*$ (null hypothesis), we have

$$u_N(\hat{d}_0 - d_0) \overset{(\mathcal{L})}{\to} U. \tag{9.25}$$

We first observe that under the conditions on $j = (j_N)$ of the theorem, the convergence (6.10) involving $\widetilde{d}_0$ holds, see [11].

The computations of Section 5 in [11] allows us to specify (6.9) as

$$\hat{d}_0 - d_0 = L(2^{-2d_0 j}\overline{\mathbf{S}}_{n_j,j}) + o_P\left(2^{-2d_0 j}\overline{\mathbf{S}}_{n_j,j}\right) + O(2^{-\zeta j}),$$

where $L$ is the linear form

$$L(z_1,\dots,z_p) = \sum_{i=1}^{p} w_i 2^{-2d_0 i} z_i,$$

where the weights $(w_i)$ have been defined in Section 6.1. The same linearization holds for $\widetilde{d}_0 - d_0$ with $\overline{\mathbf{S}}_{n_j,j}$ replaced by $\mathbf{S}^{(q_0,q_0,q_0-1)}_{n_j,j}$, so by subtracting, we get

$$\hat{d}_0 - d_0 = \widetilde{d}_0 - d_0 + 2^{-2d_0 j}\left[O_P\left(\left|\overline{\mathbf{S}}_{n_j,j} - \mathbf{S}^{(q_0,q_0,q_0-1)}_{n_j,j}\right|\right) + o_P\left(\left|\mathbf{S}^{(q_0,q_0,q_0-1)}_{n_j,j}\right|\right)\right] + O(2^{-\zeta j}). \tag{9.26}$$

Using (6.13), which corresponds to (3.6) under $H_0$, we can apply Theorem 8.2 so that

$$\overline{\mathbf{S}}_{n_j,j} - \mathbf{S}^{(q_0,q_0,q_0-1)}_{n_j,j} = o_P\left(\left\|\mathbf{S}^{(q_0,q_0,q_0-1)}_{n_j,j}\right\|_2\right).$$

With (9.26), we get

$$\hat{d}_0 - d_0 = \widetilde{d}_0 - d_0 + o_P\left(2^{-2d_0 j}\left\|\mathbf{S}^{(q_0,q_0,q_0-1)}_{n_j,j}\right\|_2\right) + O(2^{-\zeta j}).$$

Since $u_N(\widetilde{d}_0 - d_0)$ converges in distribution, it remains to check that $u_N\|\mathbf{S}^{(q_0,q_0,q_0-1)}_{n_j,j}\|_2 = O(1)$ and $u_N 2^{-\zeta j} = o(1)$. By the definition of $u_N$ in (6.10) and since $\gamma_j = 2^j$, $d_0 = K + \delta(q_0)$ and $\delta(1) = d$, the asymptotic equivalences (8.21) and (8.22) in Proposition 8.4 can be written as

$$\|S^{(q_0,q_0,q_0-1)}_{n_j,j}\|_2 \sim C\, u_N^{-1} 2^{2d_0 j}.$$

The bound $u_n\|\mathbf{S}^{(q_0,q_0,q_0-1)}_{n_j,j}\|_2 = O(1)$ follows under $H_0$. Finally the bound $u_N 2^{-\zeta j} = o(1)$ follows from the bias negligibility condition (6.14). Hence we get (9.25), which concludes the proof.

## 10. Technical lemmas

The next lemma give an explicit expression of the spectral density of $H_q(X)$ for $q < 1/(1-2d)$ and is a refined version of Lemma 4.1 in [9]. It is used in the proof of Theorem 2.1.

**Lemma 10.1.** *Let $q$ be a positive integer greater than 2. The spectral density of $\{H_q(X_\ell)\}_{\ell\in\mathbb{Z}}$ is*

$$f_{H_q} := q!(f \star \cdots \star f), \tag{10.1}$$

*where $f$ denotes the spectral density of $X = \{X_\ell\}_{\ell \in \mathbb{Z}}$. Moreover if in addition $q < 1/(1-2d)$ the function $f^*_{H_q}$ in*

$$f_{H_q}(\lambda) = |1 - \mathrm{e}^{-\mathrm{i}\lambda}|^{-2\delta(q)} f^*_{H_q}(\lambda). \tag{10.2}$$

*is bounded on $\lambda \in (-\pi, \pi)$ and for any $\zeta \in (0, 2\delta(q))$ such that $\zeta \le \beta$, where $\beta$ has been defined in (1.2), one has*

$$|f^*_{H_q}(\lambda) - f^*_{H_q}(0)| \le L|\lambda|^\zeta, \tag{10.3}$$

*for some $L > 0$.*

*Proof.* The explicit expression (10.1) of $f_{H_q}$ has already been given in Lemma 4.1 in [9]. Moreover in the same lemma, we also already showed that $f^*_{H_q}$ defined by (10.2) is a bounded function. We then only need to prove that (10.3) holds for some $L > 0$. We prove the result by induction on $q$.

Assume first that $q = 2$. By assumption on $f^*$ and definition of $\beta$, we know that for some $C > 0$ and any $\zeta \le \beta$

$$|f^*(\lambda) - f^*(0)| \le C|\lambda|^\zeta. \tag{10.4}$$

Since $f_{H_2} = 2f * f$, we then apply the second part of Lemma 8.2 of [9], with $\beta_1 = \beta_2 = 2d$, $g_1^* = g_2^* = f^*$ (using the notations of that lemma). We see that Condition (66) of Lemma 8.2 of [9] is satisfied provided that $\zeta \le \beta$ and $\zeta < \beta_1 + \beta_2 - 1 = 2d + 2d - 1 = 2\delta(2)$ (which are necessary conditions of the lemma). Hence for some $L > 0$, one has

$$|f^*_{H_2}(\lambda) - f^*_{H_2}(0)| \le L|\lambda|^\zeta.$$

If we now assume that $q > 2$, we can also apply the second part of Lemma 8.2 of [9], with $\beta_1 = 2\delta(q-1)$, $\beta_2 = 2d$, $g_1^* = f^*_{H_{q-1}}$ and $g_2^* = f^*$ which allows us to proceed by induction. □

Lemmas 10.2 to 10.4 are used in the proof of Theorem 8.2.

**Lemma 10.2.** *Let $\delta^+$ be the exponent defined in (2.6). One has*

$$\sup\left\{\delta^+(q) + \delta^+(q') : 1 \le q \le q',\ (q, q') \ne (1,1)\right\} < 2d. \tag{10.5}$$

*Proof.* For any $(q, q')$ in the considered set, one has $q \ge 1$ and $q' \ge 2$. Since $\delta_+$ is non-increasing, we get $\delta^+(q) + \delta^+(q') \le \delta^+(1) + \delta^+(2) = d + (2d - 1/2)_+ < 2d$ since $d < 1/2$. Lemma 10.2 follows. □

**Lemma 10.3.** *Let $\alpha(q, q', p)$ and $\beta'(q, q', p)$ be the exponents defined in (8.11) and (8.13) respectively, for $0 \le p \le q \le q'$. Then the following facts hold:*

*(i) $\alpha(q, q', p)$ is non-decreasing as $q$ or $q'$ increases and is non-increasing as $p$ increases.*
*(ii) $\beta'(q, q', p)$ is non-increasing as $q$ or $q'$ increases.*

*(iii)* *On the set* $\{(q,q',p) : 0 \le p \le q \le q',\ \alpha(q,q',p) < 1/2\}$, $\beta'(q,q',p)$ *is non-decreasing as* $p$ *increases.*

*Proof.* The facts (i) and (ii) directly follow by observing that $\delta_+$ is a non-increasing function. Now suppose that $\alpha(q,q',p) < 1/2$. It follows that $p \neq 0$ and $\delta_+(q-p) > 0$ and $\delta_+(q'-p) > 0$ since otherwise $\delta_+(q-p)+\delta_+(q'-p) \le 1/2$ which implies $\alpha(q,q',p) = 1/2$ in the case $p \neq 0$. Now, when $\delta_+(q-p) > 0$ and $\delta_+(q'-p) > 0$, we have in the definition of $\beta'$ that

$$\beta'(q,q',p) = \begin{cases} \max(\delta(q-p) + \delta(q'-p) - 1, -1/2) & \text{if } \delta_+(p) = 0 \\ \max(\delta(q) + \delta(q'), -1/2) & \text{if } \delta_+(p) > 0. \end{cases}$$

The second line comes from the fact that $2\delta(p) + \delta(q-p) + \delta(q'-p) - 1 = \delta(q) + \delta(q')$. Now it is clear that $\beta'(q,q',p)$ is non-decreasing as $p$ increases. □

**Lemma 10.4.** *Let* $\alpha(q,q',p)$ *be the exponent defined in (8.11) for* $0 \le p \le q \le q'$. *Then we have* $\alpha(q,q',p) \in (0,1/2]$ *and the three following assertions hold:*

*(i)* *For any* $q \ge 1$, $\alpha(q,q+1,q) = 1/2 - d < 1/2$.
*(ii)* *If* $d \le 1/4$, *then for all* $1 \le q \le q'$ *and* $0 \le p \le \min(q,q'-1)$ *such that* $q' \neq q+1$, *we have* $\alpha(q,q',p) = 1/2$.
*(iii)* *If* $d > 1/4$, *then for all* $q \ge 2$, $\alpha(q,q,q-1) = 1 - 2d < 1/2$.

*Proof.* Assertion (i) follows by applying (8.11), since $\delta_+(0) = 1/2$ and $\delta_+(1) = d$.

Suppose that $d \le 1/4$. If $p = 0$, $\alpha(q,q',p) = 1/2$ for any $q,q'$ by (8.11). Let now $p \ge 1$. Let $(q,q')$ be such that $1 \le q \le q'$, $0 \le p \le \min(q,q'-1)$ and $q' \neq q+1$. Then either $q = q' \ge p+1$ (first case) or $q' \ge q+2$ and $q \ge p$ (second case). Then by Lemma 10.3(i), we have in the first case

$$\alpha(q,q,p) \ge \alpha(p+1,p+1,p) = \min(1-2d, 1/2) = 1/2,$$

since $d \le 1/4$. In the second case, Lemma 10.3 (i) implies

$$\alpha(q,q',p) \ge \alpha(p,p+2,p) = \min(1/2 - \delta_+(2), 1/2) = 1/2,$$

since $d \le 1/4$ implies $\delta_+(2) = 0$. This proves Assertion (ii).

To obtain Assertion (iii), we remark that

$$\alpha(q,q,q-1) = \min(1-2d, 1/2) = 1 - 2d < 1/2,$$

since $d > 1/4$. □

**Lemma 10.5.** *Consider a sequence* $\{q_\ell,\ \ell \in \mathcal{L}\}$ *with* $\mathcal{L}$ *a set of consecutive integers starting at* 0. *Let* $\nu_c(d)$ *be as in Definition 4.1 for all* $d \in (1/2(1 - 1/q_0), 1/2)$, *so that (2.10) holds. Then the following assertions hold:*

*(i)* *If* $q_0 = 1$, $\nu_c(d)$ *is non-increasing as* $d$ *increases.*
*(ii)* *If* $q_0 \ge 2$, $\nu_c(d)$ *is non-decreasing as* $d$ *increases.*

*Proof.* We first consider the case $q_0 \geq 2$. In this case, either $I_0 = \emptyset$ and $\nu_c(d) = \infty$, or $I_0 \neq \emptyset$ and $\nu_c$ is a continuous function taking values

$$\nu_c(d) = \begin{cases} 1 + \frac{4(\delta(q_0) - \delta(q_{\ell_0}))}{1-2d} = 1 + 2(q_{\ell_0} - q_0) & \text{if } \delta(q_{\ell_0}) > 0, \\ 1 + \frac{4\delta(q_0)}{1-2d} = 1 - 2q_0 + 2/(1-2d) & \text{otherwise.} \end{cases}$$

Hence we obtain (ii).

We now consider the case $q_0 = 1$. In this case, with the convention $a/0 = \infty$ for $a > 0$ the following formula can be applied in all cases:

$$\nu_c(d) = \min\left(\frac{1 - 2\delta_+(q_1 - 1)}{\delta_+(2)}, \frac{2d + 1/2 - 2\delta_+(q_{\ell_r}) - \delta(r+1)}{\delta(r+1)} : r \in \mathcal{R}\right)$$

This comes from the fact that if $d \leq 1/4$, we have $\delta_+(2) = 0$ and $\mathcal{R} \subset \{0\}$ with equality if and only if $I_0 \neq \emptyset$. Let us denote

$$\tilde{\mathcal{R}} = \{1 + q_{\ell+1} - q_\ell : \ell \in \mathcal{L}\},$$

so that $\mathcal{R} = \tilde{\mathcal{R}} \cap \{r : \delta(r+1) > 0\}$. Since $2d + 1/2 - 2\delta_+(q_{\ell_r}) - \delta(r+1) = 2(d - \delta_+(q_{\ell_r})) + (1/2 - \delta(r+1)) > 0$ we get using the same convention as above that

$$\nu_c(d) = \min\left(\frac{1 - 2\delta_+(q_1 - 1)}{\delta_+(2)}, \frac{2d + 1/2 - 2\delta_+(q_{\ell_r}) - \delta(r+1)}{\delta_+(r+1)} : r \in \tilde{\mathcal{R}}\right),$$

where now the set $\tilde{\mathcal{R}}$ does not depend on $d$. To prove (i), we thus only need to show the following two assertions (setting $q = q_1 - 1$ and then $p = r + 1$ and $q = q_{\ell_r}$).

(a) For any given positive integer $q$, $(1 - 2\delta_+(q))/\delta_+(2)$ is non-increasing as $d$ increases,
(b) For any given positive integers $p$ and $q$, $\mu(d) := (2d + 1/2 - 2\delta_+(q) - \delta(p))/\delta_+(p)$ is non-increasing as $d$ increases.

Assertion (a) follows from the fact that $\delta(q)$ is increasing with $d$ for any given $q \geq 1$. Finally, we need to prove Assertion (b). Take some integers $p, q \geq 1$ and denote $\mu(d)$ as in (b). If $\delta_+(p) = 0$, which is equivalent to $d \leq 1/2(1 - 1/p)$, $\mu(d) = \infty$. Now $\mu(d)$ is continuous over $d > 1/2(1 - 1/p)$ and takes value

$$\mu(d) = \min\left(\frac{1/2 + 2d}{dp + (p-1)/2}, \frac{1/2 + (q-1)(1-2d)}{dp + (p-1)/2}\right).$$

Since the two arguments in the min are decreasing functions of $d$ over $d > 1/2(1 - 1/p)$, we conclude that (b) holds. The proof of the lemma is achieved. □

## Appendix A: Integral representations

It is convenient to use an integral representation in the spectral domain to represent the random processes (see for example [19, 21]). The stationary Gaussian process $\{X_k, k \in \mathbb{Z}\}$ with spectral density (1.3) can be written as

$$X_\ell = \int_{-\pi}^{\pi} \mathrm{e}^{\mathrm{i}\lambda\ell} f^{1/2}(\lambda) \mathrm{d}\widehat{W}(\lambda) = \int_{-\pi}^{\pi} \frac{\mathrm{e}^{\mathrm{i}\lambda\ell} f^{*1/2}(\lambda)}{|1-\mathrm{e}^{-\mathrm{i}\lambda}|^d} \mathrm{d}\widehat{W}(\lambda), \quad \ell \in \mathbb{N}. \tag{A.1}$$

This is a special case of

$$\widehat{I}(g) = \int_{\mathbb{R}} g(x) \mathrm{d}\widehat{W}(x), \tag{A.2}$$

where $\widehat{W}(\cdot)$ is a complex–valued Gaussian random measure satisfying, for any Borel sets $A$ and $B$ in $\mathbb{R}$, $\mathbb{E}(\widehat{W}(A)) = 0$, $\mathbb{E}(\widehat{W}(A)\overline{\widehat{W}(B)}) = |A \cap B|$ and

$$\widehat{W}(A) = \overline{\widehat{W}(-A)}.$$

The integral (A.2) is defined for any function $g \in L^2(\mathbb{R})$ and one has the isometry

$$\mathbb{E}(|\widehat{I}(g)|^2) = \int_{\mathbb{R}} |g(x)|^2 \mathrm{d}x.$$

The integral $\widehat{I}(g)$, moreover, is real–valued if

$$g(x) = \overline{g(-x)}.$$

We shall also consider multiple Itô–Wiener integrals

$$\widehat{I}_q(g) = \int_{\mathbb{R}^q}'' g(\lambda_1, \ldots, \lambda_q) \mathrm{d}\widehat{W}(\lambda_1) \cdots \mathrm{d}\widehat{W}(\lambda_q)$$

where the double prime indicates that one does not integrate on hyperdiagonals $\lambda_i = \pm\lambda_j, i \neq j$. The integrals $\widehat{I}_q(g)$ are handy because we will be able to expand our non–linear functions $G(X_k)$ introduced in Section 1 in multiple integrals of this type.

These multiples integrals are defined for $g \in \overline{L^2}(\mathbb{R}^q, \mathbb{C})$, the space of complex valued functions defined on $\mathbb{R}^q$ satisfying

$$g(-x_1, \ldots, -x_q) = \overline{g(x_1, \ldots, x_q)} \text{ for } (x_1, \ldots, x_q) \in \mathbb{R}^q, \tag{A.3}$$

$$\|g\|_{L^2}^2 := \int_{\mathbb{R}^q} |g(x_1, \ldots, x_q)|^2 \,\mathrm{d}x_1 \cdots \mathrm{d}x_q < \infty. \tag{A.4}$$

Hermite polynomials are related to multiple integrals as follows: if $X = \int_{\mathbb{R}} g(x) \mathrm{d}\widehat{W}(x)$ with $\mathbb{E}(X^2) = \int_{\mathbb{R}} |g(x)|^2 \mathrm{d}x = 1$ and $g(x) = \overline{g(-x)}$ so that $X$ has unit variance and is real–valued, then

$$H_q(X) = \widehat{I}_q(g^{\otimes q}) = \int_{\mathbb{R}^q}'' g(x_1) \cdots g(x_q) \mathrm{d}\widehat{W}(x_1) \cdots \mathrm{d}\widehat{W}(x_q). \tag{A.5}$$

## Appendix B: The wavelet filters

The sequence $\{Y_t\}_{t\in\mathbb{Z}}$ can be formally expressed as

$$Y_t = \Delta^{-K} G(X_t), \quad t \in \mathbb{Z}.$$

The study of the asymptotic behavior of the scalogram of $\{Y_t\}_{t\in\mathbb{Z}}$ at different scales involve multidimensional wavelets coefficients of $\{G(X_t)\}_{t\in\mathbb{Z}}$ and of $\{Y_t\}_{t\in\mathbb{Z}}$. To obtain them, one applies a multidimensional linear filter $\mathbf{h}_j(\tau), \tau \in \mathbb{Z} = (h_{j,\ell}(\tau))$, at each scale index $j \geq 0$. We shall characterize below the multidimensional filters $\mathbf{h}_j(\tau)$ by their discrete Fourier transform:

$$\widehat{\mathbf{h}}_j(\lambda) = \sum_{\tau\in\mathbb{Z}} \mathbf{h}_j(\tau) \mathrm{e}^{-\mathrm{i}\lambda\tau},\ \lambda \in [-\pi,\pi], \quad \mathbf{h}_j(\tau) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \widehat{\mathbf{h}}_j(\lambda)\mathrm{e}^{\mathrm{i}\lambda\tau}\mathrm{d}\lambda, \tau \in \mathbb{Z}. \tag{B.1}$$

The resulting wavelet coefficients $\mathbf{W}_{j,k}$, where $j$ is the scale index and $k$ the location are defined as

$$\mathbf{W}_{j,k} = \sum_{t\in\mathbb{Z}} \mathbf{h}_j(\gamma_j k - t) Y_t = \sum_{t\in\mathbb{Z}} \mathbf{h}_j(\gamma_j k - t)\Delta^{-K} G(X_t),\ j \geq 0, k \in \mathbb{Z}, \tag{B.2}$$

where $\gamma_j \uparrow \infty$ as $j \uparrow \infty$ is a sequence of non–negative scale factors applied at scale index $j$, for example $\gamma_j = 2^j$. We do not assume that the wavelet coefficients are orthogonal nor that they are generated by a multiresolution analysis. Our assumption on the filters $\mathbf{h}_j = (h_{j,\ell})$ are as follows:

(W-1) Finite support: For each $\ell$ and $j$, $\{h_{j,\ell}(\tau)\}_{\tau\in\mathbb{Z}}$ has finite support. Further there exists some $A > 0$ such that for any $j$ and any $\ell$ one has

$$\mathrm{supp}(h_{j,\ell}) \subset \gamma_j[-A, A]. \tag{B.3}$$

(W-2) Uniform smoothness: There exists $M \geq K$, $\alpha > 1$ and $C > 0$ such that for all $j \geq 0$ and $\lambda \in [-\pi, \pi]$,

$$|\widehat{\mathbf{h}}_j(\lambda)| \leq \frac{C\gamma_j^{1/2}|\gamma_j\lambda|^M}{(1+\gamma_j|\lambda|)^{\alpha+M}}. \tag{B.4}$$

By $2\pi$-periodicity of $\widehat{h}_j$ this inequality can be extended to $\lambda \in \mathbb{R}$ as

$$|\widehat{\mathbf{h}}_j(\lambda)| \leq C\frac{\gamma_j^{1/2}|\gamma_j\{\lambda\}|^M}{(1+\gamma_j|\{\lambda\}|)^{\alpha+M}}. \tag{B.5}$$

where $\{\lambda\}$ denotes the element of $(-\pi, \pi]$ such that $\lambda - \{\lambda\} \in 2\pi\mathbb{Z}$.

(W-3) Asymptotic behavior: There exists a sequence of phase functions $\Phi_j : \mathbb{R} \to (-\pi, \pi]$ and some non identically zero function $\widehat{\mathbf{h}}_\infty$ such that

$$\lim_{j\to+\infty} (\gamma_j^{-1/2}\widehat{\mathbf{h}}_j(\gamma_j^{-1}\lambda)) = \widehat{\mathbf{h}}_\infty(\lambda), \tag{B.6}$$

locally uniformly on $\lambda \in \mathbb{R}$.

In (W-3) *locally uniformly* means that for all compact $K \subset \mathbb{R}$,

$$\sup_{\lambda \in K} \left| \gamma_j^{-1/2} \widehat{\mathbf{h}}_j(\gamma_j^{-1}\lambda) \mathrm{e}^{\mathrm{i}\Phi_j(\lambda)} - \widehat{\mathbf{h}}_\infty(\lambda) \right| \to 0.$$

It implies in particular that $\widehat{\mathbf{h}}_\infty$ is continuous over $\mathbb{R}$.

A more convenient way to express the wavelet coefficients $\mathbf{W}_{j,k}$ than in (B.2) is to incorporate the linear filter $\Delta^{-K}$ into the filter $\mathbf{h}_j$ and denote the resulting filter $\mathbf{h}_j^{(K)}$. Then

$$\mathbf{W}_{j,k} = \sum_{t \in \mathbb{Z}} \mathbf{h}_j^{(K)}(\gamma_j k - t) G(X_t), \tag{B.7}$$

where

$$\widehat{\mathbf{h}}_j^{(K)}(\lambda) = (1 - \mathrm{e}^{-\mathrm{i}\lambda})^{-K} \widehat{\mathbf{h}}_j(\lambda) \tag{B.8}$$

is the discrete Fourier transform of $\mathbf{h}_j^{(K)}$, see [10] for more details.

## Appendix C: The multiscale wavelet inference setting

We state here two theorems that are used in Section 6 to derive statistical properties of the estimator of the memory parameter $d_0$. This parameter is obtained from univariate multiscale wavelet filters $g_j$. Since, Theorem 3.2 applies to multivariate filters $\mathbf{h}_j$ which define the multivariate scalogram $\mathbf{S}_{n,j}$, we explain in this appendix the connection between these two perspectives.

We first give some details about the definition of the estimator of the memory parameter. We use dyadic scales here, as in the standard wavelet analysis described in [20], where the univariate wavelet coefficients are defined as

$$W_{j,k} = \sum_{t \in \mathbb{Z}} g_j(2^j k - t) Y_t, \tag{C.1}$$

which corresponds to (2.14) with $\gamma_j = 2^j$ and with $(g_j)$ denoting a sequence of filters that satisfies (W-1)–(W-3) with $m = 1$. In the case of a multiresolution analysis, $g_j$ can be deduced from the associated mirror filters.

The number $n_j$ of wavelet coefficients available at scale $j$, is related both to the number $N$ of observations $Y_1, \ldots, Y_N$ of the time series $Y$ and to the length $T$ of the support of the wavelet $\psi$. More precisely, one has

$$n_j = [2^{-j}(N - T + 1) - T + 1] = 2^{-j}N + 0(1), \tag{C.2}$$

where $[x]$ denotes the integer part of $x$ for any real $x$. Details about the above facts can be found in [20, 24].

The univariate scalogram is an empirical measure of the distribution of "energy of the signal" along scales, based on the $N$ observations $Y_1, \ldots, Y_N$. It is defined as

$$\widehat{\sigma}_j^2 = \frac{1}{n_j} \sum_{k=0}^{n_j - 1} W_{j,k}^2, \quad j \geq 0, \tag{C.3}$$

and is identical to $S_{n_j,j}$ defined in (2.15). The wavelet spectrum is defined as

$$\sigma_j^2 = \mathbb{E}[\widehat{\sigma}_j^2] = \mathbb{E}[W_{j,k}^2] \quad \text{for all } k, \tag{C.4}$$

where the last equality holds for $M \geq K$ since in this case $\{W_{j,k}, k \in \mathbb{Z}\}$ is weakly stationary.

To define our wavelet estimator of the memory parameter $d_0$, we are given some positive weights $w_0, \dots, w_p$ such that

$$\sum_{i=0}^{p} w_i = 0 \text{ and } \sum_{i=0}^{p} i w_i = \frac{1}{2\log(2)}.$$

We then set

$$\hat{d}_0 = \sum_{i=0}^{p} w_i \log(\widehat{\sigma}_{j+i}). \tag{C.5}$$

To derive statistical properties of this estimator, we apply Theorem 3.2 using a sequence of multivariate filters $(\mathbf{h}_j)_{j\geq 0}$ related to the family of univariate filters $g_j$ in a way indicated below.

We first give an example and consider the case $p = 1$. To investigate the asymptotic properties of $\hat{d}_0$, we then have to study the *joint behavior* of $W_{j-u,k}$ for $u = 0, 1$. Recall that $j - 1$ is a finer scale than $j$. Following the framework of [24], we consider the multivariate coefficients $\mathbf{W}_{j,k} = (W_{j,k}, W_{j-1,2k}, W_{j-1,2k+1})$, since, in addition to the wavelet coefficients $W_{j,k}$ at scale $j$, there are twice as many wavelet coefficients at scale $j - 1$, the additional coefficients being $W_{j-1,2k}$, $W_{j-1,2k+1}$. These coefficients can be viewed in this case as the output of a three-dimensional filter $\mathbf{h}_j$ defined as $\mathbf{h}_j(\tau) = (g_j(\tau), g_{j-1}(\tau), g_{j-1}(\tau + 2^{j-1}))$. These three entries correspond to $(u, v)$ below equal to $(0, 0)$, $(1, 0)$ and $(1, 1)$, respectively, in the general case below.

In the general case, each $\mathbf{h}_j$ is defined as follows. For all, $j \geq 0$, $u \in \{0, \dots, j\}$ and $v \in \{0, \dots, 2^u - 1\}$, let $\ell = 2^u + v$ and define a filter $h_{\ell,j}$ by

$$h_{\ell,j}(t) = g_{j-u}(t + 2^{j-u}v), \quad t \in \mathbb{Z}. \tag{C.6}$$

Applying this definition and (C.1) with $\gamma_j = 2^j$, we get

$$W_{j-u,2^u k+v} = \sum_{t\in\mathbb{Z}} h_{\ell,j}(2^j k - t) Y_t.$$

These coefficients are stored in a vector $\mathbf{W}_{j,k} = [W_{\ell,j,k}]_\ell$, say of length $m = 2^p - 1$,

$$W_{\ell,j,k} = W_{j-u,2^u k+v}, \quad \ell = 2^u + v = 1, 2, \dots, m, \tag{C.7}$$

which corresponds to the multivariate wavelet coefficient (2.17) with $\mathbf{h}_j(t)$ having components $h_{\ell,j}(t)$, $\ell = 1, 2, \dots, m$ defined by (C.6). This way of proceeding allows us to express the vector $[\widehat{\sigma}_{j-u}^2 - \sigma_{j-u}^2]_{u=0,\dots,p-1}$ as a linear function of the vector $\overline{\mathbf{S}}_{n_j,j}$ defined by (2.16), up to a negligible term. We can then deduce, as in Section 6, the asymptotic behavior of $\hat{d}_0$ of the multivariate scalogram $\overline{\mathbf{S}}_{n_j,j}$ using (C.5).

We now indicate the asymptotic behavior of the univariate multiscale scalogram in the case $G = H_{q_0}$ since it will be needed in Section 6. We state the results separately for $q_0 = 1$ and for $q_0 \geq 2$.

We first consider the case $q_0 = 1$:

**Theorem C.1.** *Suppose $G = H_{q_0}$ with $q_0 = 1$ and that Assumptions A(i), (ii) in Section 2 hold. Set $\gamma_j = 2^j$ and let $\{(g_j)_{j\geq 0},\, g_\infty\}$ be a sequence of univariate filters satisfying (W-1)–(W-3) with $m = 1$ and $M \geq d + K$. Then, as $j \to \infty$,*

$$\sigma_j^2 \sim f^*(0)\, L_1(\widehat{g}_\infty)\, 2^{2j(d+K)}, \tag{C.8}$$

*where $L_1$ has been defined in (3.4). Let now $j = j(N)$ be an increasing sequence such that $j \to \infty$ and $N2^{-j} \to \infty$. Define $n_j$, $\widehat{\sigma}_j^2$ and $\sigma_j^2$ as in (C.2), (C.3) and (C.4), respectively. Then, as $N \to \infty$,*

$$\left\{ n_j^{1/2} \left( \frac{\widehat{\sigma}_{j-u}^2}{\sigma_{j-u}^2} - 1 \right) \right\}_{u \geq 0} \xrightarrow{\text{fidi}} \left\{ Q_u^{(d)} \right\}_{u \geq 0}, \tag{C.9}$$

*where $Q^{(d)}$ denotes a centered Gaussian process with covariance function*

$$\mathrm{Cov}(Q_u^{(d)}, Q_{u'}^{(d)}) = \frac{4\pi\, 2^{2(d+K)|u'-u| - \max(u,u')}}{L_1(\widehat{g}_\infty)^2} \int_{-\pi}^{\pi} |D_{\infty,u-u'}(\lambda)|^2 \mathrm{d}\lambda, \tag{C.10}$$

*with for all $m \in \mathbb{Z}$ and $\lambda \in (-\pi, \pi)$,*

$$D_{\infty,m}(\lambda) = \sum_{\ell \in \mathbb{Z}} |\lambda + 2\pi\ell|^{-2(d+K)} \mathbf{e}_m(\lambda + 2\pi\ell) \overline{\widehat{g}_\infty(\lambda + 2\pi\ell)} \widehat{g}_\infty(2^{-m}(\lambda + 2\pi\ell)),$$

*and*

$$\mathbf{e}_m(\xi) = 2^{-m/2}[\mathrm{e}^{-i2^{-m}v\xi},\, v = 0, \dots, 2^m - 1]^T.$$

*Proof.* We first observe that the proof of formula (4.5) in Theorem 4.1 of [11] remains valid in the case $q_0 = 1$. This yields (C.8).

We now prove the convergence (C.9). To do so we adapt the corresponding proof of Theorem 4.1 of [11] done for $q_0 \geq 2$. From [11] (see equality (9.5)), we have

$$\widehat{\sigma}_{j-u}^2 - \sigma_{j-u}^2 = \frac{n_j}{n_{j-u}} \sum_{v=0}^{2^u - 1} \overline{S}_{n_j,j}(2^u + v) + O_P(\sigma_{j-u}^2 / n_{j-u}), \quad u = 0, \dots, p-1,$$

where we denoted the entries of the multivariate scalogram $\overline{\mathbf{S}}_{n_j,j}$ in (2.16) as $[\overline{S}_{n_j,j}(\ell)]_{\ell=1,\dots,m}$. In addition, we also proved in Section 9 of [11] that the multivariate filters $\mathbf{h}_j(t)$ involved in the definition of the multivariate wavelet coefficients, defined by (C.6), satisfy the assumptions of Theorem 3.2 of [11]. We can then apply Theorem 3.2(a) of [11] which provides the asymptotic behavior of the multivariate scalogram $\overline{S}_{n_j,j}$. Using the equality (9.6) of [11] relating $\widehat{h}_{\ell,\infty}$ and $\widehat{g}_\infty$ as,

$$\widehat{h}_{\ell,\infty}(\lambda) = 2^{-u/2}\widehat{g}_\infty(2^{-u}\lambda)\mathrm{e}^{\mathrm{i}2^{-u}v\lambda},$$

we then deduce that as $j \to \infty$,

$$\left\{n_j^{1/2}2^{-2(j-u)(d+K)}\overline{\mathbf{S}}_{n_j,j}(2^u+v)\right\}_{u,v} \xrightarrow{(\mathcal{L})} \mathcal{N}(0,\widetilde{\Gamma}),$$

where (we denote $\lambda_p = \lambda + 2p\pi$),

$$\begin{aligned}\widetilde{\Gamma}_{(u,v),(u',v')} &= 2^{2(u+u')(d+K)}\Gamma_{2^u+v,2^{u'}+v} \\ &= 4\pi(f^*(0))^2\, 2^{2(u+u')(d+K-\frac{1}{2})} \\ &\quad\times \int_{-\pi}^{\pi}\left|\sum_{p\in\mathbb{Z}}|\lambda_p|^{-2(K+d)}\widehat{g}_\infty(2^{-u}\lambda_p)\overline{\widehat{g}_\infty}(2^{-u'}\lambda_p)\mathrm{e}^{\mathrm{i}(2^{-u}v-2^{-u'}v')\lambda_p}\right|^2 \mathrm{d}\lambda,\end{aligned}$$

and $(u,v)$ (resp $(u',v')$) take values $u = 0,\dots,p-1$ (resp $u' = 0,\dots,p-1$) and $v = 0,\dots,2^u-1$ (resp $v' = 0,\dots,2^{u'}-1$). We showed in [11], Relation (9.4), that as $j\to\infty$, $n_j/n_{j-u} \sim 2^{-u}$. Using also (C.8), which implies that $\sigma_{j-u}^2 \sim f^*(0)\,L_1(\widehat{g}_\infty)\;2^{2(j-u)(d+K)}$ as $j\to\infty$, and following the proof of Theorem 4.1 of [11], we get

$$\left\{n_j^{1/2}\frac{1}{\sigma_{j-u}^2}\frac{n_j}{n_{j-u}}\sum_{v=0}^{2^u-1}\overline{\mathbf{S}}_{n_j,j}(2^u+v)\right\}_u \xrightarrow{(\mathcal{L})} \mathcal{N}(0,\overline{\Gamma}),$$

with

$$\overline{\Gamma}_{u,u'} = \frac{2^{-u-u'}}{(f^*(0))^2L_1(\widehat{g}_\infty)^2}\sum_{v=0}^{2^u-1}\sum_{v'=0}^{2^{u'}-1}\widetilde{\Gamma}_{(u,v),(u',v')} \tag{C.11}$$

$$\begin{aligned}&= \frac{2^{2(u+u')(d+K-\frac{1}{2})}}{(f^*(0))^2L_1(\widehat{g}_\infty)^2}\sum_{v=0}^{2^u-1}\sum_{v'=0}^{2^{u'}-1}\Gamma_{2^u+v,2^{u'}+v} \\ &= \frac{4\pi 2^{2(u+u')(d+K-1)}}{L_1(\widehat{g}_\infty)^2} \\ &\quad\times\int_{-\pi}^{\pi}\sum_{v=0}^{2^u-1}\sum_{v'=0}^{2^{u'}-1}\left|\sum_{p\in\mathbb{Z}}\frac{\widehat{g}_\infty(2^{-u}\lambda_p)\overline{\widehat{g}_\infty(2^{-u'}\lambda_p)}\mathrm{e}^{\mathrm{i}(2^{-u}v-2^{-u'}v')\lambda_p}}{|\lambda_p|^{2(K+d)}}\right|^2\mathrm{d}\lambda,\end{aligned} \tag{C.12}$$

and where $u,u' = 0,\dots,p-1$. Thereafter, we follow the same lines that in the proof of [24, Theorem 2]. Assume for example that $u' \geq u$. We have to estimate

$$\sum_{v=0}^{2^u-1}\sum_{v'=0}^{2^{u'}-1}\left|\sum_{p\in\mathbb{Z}}|\lambda_p|^{-2(K+d)}\widehat{g}_\infty(2^{-u}\lambda_p)\overline{\widehat{g}_\infty(2^{-u'}\lambda_p)}\mathrm{e}^{\mathrm{i}(2^{-u}v-2^{-u'}v')\lambda_p}\right|^2,$$

which reads $\sum_{v'=0}^{2^{u'}-1} G_{u,u',v'}(\lambda)$ with

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

where $g_{u,u'}(\xi) = |2^u\xi|^{-2(K+d)}\widehat{g}_\infty(\xi)\overline{\widehat{g}_\infty(2^{u-u'}\xi)}$. We now observe that $G_{u,u',v'}$ is a $2\pi$-periodic function and write $p = 2^u q + r$ with $r \in \{0,\dots,2^u-1\}$. Hence ($\lambda_p = \lambda_r + 2^u q \times 2\pi$ and $\mathrm{e}^{2\mathrm{i}\pi v}$, if $v$ is integer),

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

with

$$h_{u,u',v'}(\xi) = \sum_{q\in\mathbb{Z}} \mathrm{e}^{-\mathrm{i}2^{u-u'}v'(2\pi q+\xi)} g_{u,u'}(\xi + 2\pi q).$$

Hence

$$G_{u,u',v'}(\lambda) = \sum_{v=0}^{2^u-1}\sum_{r=0}^{2^u-1}\sum_{r'=0}^{2^u-1} \mathrm{e}^{\mathrm{i}2^{-u}v2\pi(r-r')} h_{u,u',v'}(2^{-u}\lambda_r)\overline{h_{u,u',v'}(2^{-u}\lambda_{r'})}.$$

Observe that if $r \neq r'$

$$\sum_{v=0}^{2^u-1} \mathrm{e}^{\mathrm{i}2^{-u}v2\pi(r-r')} = 0,$$

whereas in the case $r = r'$ this sum equals $2^u$. Hence

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

As in the proof of [24, Theorem 2], we apply Lemma 1 of [25] with $g = |h_{u,u',v'}|^2$, $\gamma = 2^u$ and get

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

We then deduce that

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

Using (C.11), the definition of $G_{u,u',v'}$ and the last display, we deduce that

$$\overline{\Gamma}_{u,u'} = \frac{4\pi 2^{2(u+u')(d+K-1)}}{L_1(\widehat{g}_\infty)^2}\left(2^{2u}\int_{-\pi}^{\pi}\sum_{v'=0}^{2^{u'}-1}|h_{u,u',v'}(\lambda)|^2\right)$$

$$= \frac{4\pi 2^{2(u+u')(d+K-1)}}{L_1(\widehat{g}_\infty)^2}\times 2^{2u}2^{-4u(d+K)}$$

$$\times\sum_{v'=0}^{2^{u'}-1}\int_{-\pi}^{\pi}\left|\sum_{q\in\mathbb{Z}}|\lambda_q|^{-2(d+K)}\mathrm{e}^{-\mathrm{i}2^{u-u'}v'\lambda_q}\widehat{g}_\infty(\lambda_q)\overline{\widehat{g}_\infty(2^{-(u-u')}\lambda_q)}\right|^2\mathrm{d}\lambda$$

$$= \frac{4\pi 2^{2(u'-u)(d+K)-2u'}}{L_1(\widehat{g}_\infty)^2}$$

$$\times\sum_{v'=0}^{2^{u'}-1}\int_{-\pi}^{\pi}\left|\sum_{q\in\mathbb{Z}}|\lambda_q|^{-2(d+K)}\mathrm{e}^{-\mathrm{i}2^{u-u'}v'\lambda_q}\widehat{g}_\infty(\lambda_q)\overline{\widehat{g}_\infty(2^{-(u-u')}\lambda_q)}\right|^2\mathrm{d}\lambda.$$

For $v'\in\{0,\dots,2^{u'}-1\}$, we write $v'=v+k2^{u'-u}$ with $v\in\{0,\dots,2^{u'-u}-1\}$ and $k\in\{0,\dots,2^u-1\}$ and transform the sum in $v'$ into a sum over $v$ and $k$. We obtain

$$\overline{\Gamma}_{u,u'} = \frac{4\pi 2^{2(u'-u)(d+K)-2u'}}{L_1(\widehat{g}_\infty)^2}\sum_{v=0}^{2^{u'-u}-1}\sum_{k=0}^{2^u-1}\int_{-\pi}^{\pi}\Bigg|\sum_{q\in\mathbb{Z}}|\lambda_q|^{-2(d+K)}$$

$$\times\,\mathrm{e}^{-\mathrm{i}2^{u-u'}(v+k2^{u'-u})\lambda_q}\widehat{g}_\infty(\lambda_q)\overline{\widehat{g}_\infty(2^{-(u-u')}\lambda_q)}\Bigg|^2\mathrm{d}\lambda.$$

Since $\mathrm{e}^{-\mathrm{i}2^{u-u'}v'\lambda_q}=\mathrm{e}^{-\mathrm{i}2^{u-u'}v\lambda_q}\mathrm{e}^{-\mathrm{i}k\lambda}$ and $\sum_{k=0}^{2^u-1}|\mathrm{e}^{-\mathrm{i}k\lambda}|^2=2^u$, one has

$$\overline{\Gamma}_{u,u'} = \frac{4\pi 2^{2(u'-u)(d+K)-u'}2^{u-u'}}{L_1(\widehat{g}_\infty)^2}$$

$$\times\sum_{v=0}^{2^{u'-u}-1}\int_{-\pi}^{\pi}\left|\sum_{q\in\mathbb{Z}}|\lambda_q|^{-2(d+K)}\mathrm{e}^{-\mathrm{i}2^{u-u'}v\lambda_q}\widehat{g}_\infty(\lambda_q)\overline{\widehat{g}_\infty(2^{-(u-u')}\lambda_q)}\right|^2\mathrm{d}\lambda.$$

Define now for any $m\in\mathbb{Z}$, the vector

$$\mathbf{e}_m(\xi)=2^{-m/2}[\mathrm{e}^{\mathrm{i}2^{-m}v\xi},v=0,\dots,2^m-1]^T.$$

We then recover (C.10) which concludes the proof. □

The case $q_0\geq 2$ has been considered in [11, Theorem 4.1]. We recall it here.

**Theorem C.2.** *Suppose $G=H_{q_0}$, $q_0\geq 2$ and that Assumptions A(i), (ii) hold with $q_0\geq 2$. Set $\gamma_j=2^j$ and let $\{(g_j)_{j\geq 0},\,g_\infty\}$ be a sequence of univariate filters satisfying (W-1)–(W-3) with $m=1$ and $M\geq\delta(q_0)+K$. Then, as $j\to\infty$,*

$$\sigma_j^2\sim q_0!\,(f^*(0))^{q_0}\,L_{q_0}(\widehat{g}_\infty)\,2^{2j(\delta(q_0)+K)},\tag{C.13}$$

*where $L_p$ has been defined in (3.4) for any $p \geq 1$. Let now $j = j(N)$ be an increasing sequence such that $j \to \infty$ and $N2^{-j} \to \infty$. Define $n_j$, $\widehat{\sigma}_j^2$ and $\sigma_j^2$ as in (C.2), (C.3) and (C.4), respectively. Then, as $N \to \infty$,*

$$\left\{ n_j^{1-2d} \left( \frac{\widehat{\sigma}_{j-u}^2}{\sigma_{j-u}^2} - 1 \right) \right\}_{u \geq 0} \xrightarrow{\text{fidi}} \left\{ 2^{(2d-1)u} \, \frac{L_{q_0-1}(\widehat{g}_\infty)}{q_0! \, L_{q_0}(\widehat{g}_\infty)} \, Z_d(1) \right\}_{u \geq 0} . \tag{C.14}$$


## Acknowledgements

Marianne Clausel's research was partially supported by the PEPS project *AGREE* and LabEx *PERSYVAL-Lab* (ANR-11-LABX-0025-01) funded by the French program Investissement d'avenir. François Roueff's research was partially supported by the ANR project *MATAIM* NT09 441552. Murad S. Taqqu was supported in part by the NSF grants DMS–1007616 and DMS-1309009 at Boston University.


## References


[1] Abry, P., Helgason, H., and Pipiras, V. (2011). Wavelet-based analysis of non-Gaussian long-range dependent processes and estimation of the Hurst parameter. *Lithuanian Mathematical Journal*, 51(3):287–302. MR2832225

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

[3] Abry, P., Veitch, D., and Flandrin, P. (1998). Long-range dependence: Revisiting aggregation with wavelets. *J. Time Ser. Anal.*, 19(3):253–266. MR1626734

[4] Bardet, J., Bibi, H., and Jouini, A. (2008). Adaptive wavelet based estimator of the memory parameter for stationary gaussian processes. *Bernoulli*, 14:691–724. MR2537808

[5] Bardet, J.-M. (2000). Testing for the presence of self-similarity of Gaussian time series having stationary increments. *Journal of Time Series Analysis*, 21:497–515. MR1794484

[6] Bardet, J.-M. (2002). Statistical study of the wavelet analysis of fractional Brownian motion. *IEEE Trans. Inform. Theory*, 48(4):991–999. MR1908463

[7] Bardet, J.-M., Lang, G., Moulines, E., and Soulier, P. (2000). Wavelet estimator of long-range dependent processes. *Stat. Inference Stoch. Process.*, 3(1–2):85–99. 19th "Rencontres Franco-Belges de Statisticiens" (Marseille, 1998). MR1819288

[8] Bardet, J.-M. and Tudor, C. A. (2010). A wavelet analysis of the Rosenblatt process: Chaos expansion and estimation of the self-similarity parameter. *Stochastic Process. Appl.*, 120(12):2331–2362. MR2728168

[9] Clausel, M., Roueff, F., Taqqu, M. S., and Tudor, C. (2012). Large scale behavior of wavelet coefficients of non-linear subordinated processes

with long memory. *Applied and Computational Harmonic Analysis*, 32:223–241. MR2880280

[10] Clausel, M., Roueff, F., Taqqu, M. S., and Tudor, C. (2013). High order chaotic limits of wavelet scalograms under long–range dependence. *ALEA Lat. Am. J. Probab. Math. Stat.*, 10(2):979–1011. MR3151747

[11] Clausel, M., Roueff, F., Taqqu, M. S., and Tudor, C. (2014). Wavelet estimation of the long memory parameter for Hermite polynomial of Gaussian processes. *ESAIM: Probability and Statistics*, 18:42–76. MR3143733

[12] Faÿ, G., Moulines, E., Roueff, F., and Taqqu, M. (2009). Estimators of long-memory: Fourier versus wavelets. *J. of Econometrics*, 151(2):159–177. MR2559823

[13] Flandrin, P. (1989a). On the spectrum of fractional Brownian motions. *IEEE Transactions on Information Theory*, IT-35(1):197–199. MR0995341

[14] Flandrin, P. (1989b). Some aspects of nonstationary signal processing with emphasis on time-frequency and time-scale methods. In Combes, J., Grossman, A., and Tchamitchian, P., editors, *Wavelets*, pages 68–98. Springer-Verlag. MR1010899

[15] Flandrin, P. (1991). Fractional Brownian motion and wavelets. In Farge, M., Hung, J., and Vassilicos, J., editors, *Fractals and Fourier Transforms-New Developments and New Applications*. Oxford University Press.

[16] Flandrin, P. (1992). Wavelet analysis and synthesis of fractional Brownian motion. *IEEE Trans. Inform. Theory*, 38(2, part 2):910–917. MR1162229

[17] Flandrin, P. (1999). *Time-Frequency/Time-Scale Analysis*. Academic Press, 1st edition. MR1681043

[18] Fox, R. and Taqqu, M. S. (1986). Large-sample properties of parameter estimates for strongly dependent stationary Gaussian time series. *Ann. Statist.*, 14(2):517–532. MR0840512

[19] Major, P. (1981). *Multiple Wiener-Itô Integrals*, volume 849 of *Lecture Notes in Mathematics*. Springer, Berlin. MR0611334

[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(2):155–187. MR2345656

[21] Nualart, D. (2006). *The Malliavin Calculus and Related Topics*. Springer. MR2200233

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

[23] Robinson, P. M. (1995b). Log-periodogram regression of time series with long range dependence. *The Annals of Statistics*, 23:1048–1072. MR1345214

[24] Roueff, F. and Taqqu, M. S. (2009a). Asymptotic normality of wavelet estimators of the memory parameter for linear processes. *J. Time Ser. Anal.*, 30(5):534–558. MR2560417

[25] Roueff, F. and Taqqu, M. S. (2009b). Central limit theorems for arrays of decimated linear processes. *Stoch. Proc. App.*, 119(9):3006–3041. MR2554037

[26] Taqqu, M. S. (1975). Weak convergence to fractional Brownian motion and to the Rosenblatt process. *Z. Wahrsch. verw. Gebiete*, 31:287–302. MR0400329

[27] Taqqu, M. S. (1979). Central limit theorems and other limit theorems for functionals of Gaussian processes. *Z. Wahrsch. verw. Gebiete*, 70:191–212.

[28] Veillette, M. S. and Taqqu, M. S. (2013). Properties and numerical evaluation of the Rosenblatt distribution. *Bernoulli*, 19(3):982–1005. MR3079303

[29] Veitch, D. and Abry, P. (1999). A wavelet-based joint estimator of the parameters of long-range dependence. *IEEE Trans. Inform. Theory*, 45(3):878–897. MR1682517

[30] Wornell, G. W. and Oppenheim, A. V. (1992). Estimation of fractal signals from noisy measurements using wavelets. *IEEE Trans. Signal Process.*, 40(3):611 – 623.

[31] Yaglom, A. M. (1958). Correlation theory of processes with random stationary $n$th increments. *Amer. Math. Soc. Transl. (2)*, 8:87–141. MR0091554