



**HAL**  
open science

# Wavelet Packets of Nonstationary Random Processes: Contributing Factors for Stationarity and Decorrelation

Abdourrahmane Atto, Yannick Berthoumieu

► **To cite this version:**

Abdourrahmane Atto, Yannick Berthoumieu. Wavelet Packets of Nonstationary Random Processes: Contributing Factors for Stationarity and Decorrelation. *IEEE Transactions on Information Theory*, 2012, 58 (1), pp. 317-330. 10.1109/TIT.2011.2167496 . hal-00548105v2

**HAL Id: hal-00548105**

**<https://hal.science/hal-00548105v2>**

Submitted on 1 Sep 2011

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Wavelet Packets of Nonstationary Random Processes: Contributing Factors for Stationarity and Decorrelation

Abdourrahmane M. ATTO<sup>1</sup>, *Member, IEEE* and Yannick BERTHOUMIEU<sup>2</sup>, *Member, IEEE*

**Abstract**—The paper addresses the analysis and interpretation of second order random processes by using the wavelet packet transform. It is shown that statistical properties of the wavelet packet coefficients are specific to the filtering sequences characterizing wavelet packet paths. These statistical properties also depend on the wavelet order and the form of the cumulants of the input random process. The analysis performed points out the wavelet packet paths for which stationarization, decorrelation and higher order dependency reduction are effective among the coefficients associated with these paths. This analysis also highlights the presence of singular wavelet packet paths: the paths such that stationarization does not occur and those for which dependency reduction is not expected through successive decompositions. The focus of the paper is on understanding the role played by the parameters that govern stationarization and dependency reduction in the wavelet packet domain. This is addressed with respect to semi-analytical cumulant expansions for modeling different types of nonstationarity and correlation structures. The characterization obtained eases the interpretation of random signals and time series with respect to the statistical properties of their coefficients on the different wavelet packet paths.

**keywords:** Wavelet Transforms ; Nonstationary Random Processes ; Cumulant analysis.

## I. INTRODUCTION

Information processing from signals and time series is substantially simpler when the data follows from independent and identically distributed, *iid*, random variables. As *iid* assumption is irrelevant for most practical applications, the challenging issue of representing the data in some transform domain, in order to meet or approach the above *iid* statistical property arises naturally. In this respect, a suitable transform is required to have stationarization and decorrelating properties and, more generally, to reduce higher order dependencies between the random variables describing the time evolution of the random process under consideration.

Among the transforms that approximately achieve this goal, wavelet decompositions are highly effective because wavelets operate unconditionally with respect to the input

process and tend to achieve the desirable stationarization and decorrelating properties for a large class of stochastic processes, see [1], [2], [3], [4], among others. This class contains stationary random processes (see for instance [4], [5], [6], [7], [8]) and some nonstationary random processes such as cyclostationary random processes and processes with stationary increments (the literature on the topic is abundant and instructive, see [1], [2], [3], [9], [10], [11], [12], [13], [14], [15], [16], [17], [18], [19], [20], [21]).

The statistical properties of the wavelet coefficients of second order random processes are specific to the wavelet transform used. For instance, the statistical properties of the *continuous wavelet transform* have been addressed in [13] and [16]. In these references, a characterization is provided in order for stationarity to hold true for the coefficients of some second order nonstationary random processes. However, it is shown in [1] that the above characterization does not apply when considering the *discrete wavelet transform*, due to the multiscale decomposition scheme involved in the latter.

Furthermore, when considering the *discrete wavelet packet transform*, it is shown in [8] that for stationary random processes, the decorrelating property involved by this transform is not a trivial consequence of the characterization derived for the discrete wavelet transform in [5], [6], [7]: the wavelet order has no impact in the latter case whereas this order plays an important role in the derivation of wavelet packet statistical properties. Furthermore, it is shown in [3] that the statistical properties of the wavelet packet transform of a fractional Brownian motion are path-dependent and the generalization from the case of the discrete wavelet transform (see [9], [10], [11], [12], [17]) requires large wavelet orders as well.

A significant drawback of the discrete wavelet transform in terms of decorrelating properties is the residual correlation occurring among the coefficients located at first levels of this transform. For strongly correlated processes, this residual correlation can be significant, impacting information processing, whereas very few issues exist since these coefficients have to be processed as well: otherwise, information is lost. One solution could consist in designing the wavelet filters adaptively with respect to the input random process and the specification of some desirable statistical properties. See for instance [22] where the wavelet filters may change from one scale to another depending on the input random process, in order for the wavelet transform to

<sup>1</sup> LISTIC - Université de Savoie, Polytech Annecy-Chambéry, B.P. 80439, 74944 Annecy-le-Vieux Cedex, France, [Abdourrahmane.Atto@univ-savoie.fr](mailto:Abdourrahmane.Atto@univ-savoie.fr)

<sup>2</sup> IMS - Université de Bordeaux, IPB, ENSEIRB-MATMECA, 351 cours de la libération, 33405 Talence Cedex, France, [Yannick.Berthoumieu@ims-bordeaux.fr](mailto:Yannick.Berthoumieu@ims-bordeaux.fr)

Copyright (c) 2011 IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending a request to [pubs-permissions@ieee.org](mailto:pubs-permissions@ieee.org).

meet some specifications. For certain stochastic processes with compact time/spatial supports, sometimes referred to as reciprocal processes, an alternative consists in modeling their statistical dependencies by using pyramidally coarse to fine organized trees, [23], [24], [25], [26]. Such tree-like decomposition schemes can be exploited to develop filtering algorithms (that essentially operate smoothing) and the corresponding filters can somehow be associated with wavelet filters. Both issues given above constrain the wavelet decomposition to be conditional with respect to the input random process.

In many situations involving information processing from a huge number of observations, a fixed representation is of much interest. In this respect, the present work relates to non-adaptive frameworks, in continuation of [1], [2], [3]. In order to obtain a general framework which can make stationarization and decorrelation effective for a large class of random processes, we consider the discrete wavelet packet transform, in continuation to [3], [4]. The insight in using wavelet packets is the fact of splitting again, the subbands that have not reached a given stationarization or decorrelation level.

The consequence of the above splitting scheme is a description of wavelet packet paths through a double indexed sequence, where the indices are the decomposition level  $j$  and the frequency index  $n$ . This double indexed sequence have the particularity that the frequency index  $n = n(j)$  depends on  $j$ , which makes the use of dominated convergence theorems prohibited (see [8] where it is shown that the Lebesgue dominated convergence theorem does not apply for deriving the limit of the autocorrelation functions of the wavelet packet coefficients per path). This is why direct inference from the case of the discrete wavelet transform [1] to the case tackled in this paper (wavelet packet transform) does not apply directly.

More specifically, the analysis presented hereafter is performed with respect to the key parameters that govern stationarization and dependency reduction in wavelet packet coefficients: the number of vanishing moments of the wavelet function and the wavelet order. As highlighted below, these two parameters have different roles, but they can be aggregated in a capital parameter when considering standard families of wavelet functions.

The results given in the paper describe the impact of these parameters on the statistical properties of the wavelet packet coefficients for a large class second order random processes. A random process pertaining to this class is such that its cumulant of order  $N$  can be expanded in 3 terms: a *projective term* (with dimension  $N-1$ ), a *stationary term* (which lies along the dimension  $N-1$ ) and an  *$N$ -variate polynomial term*. Furthermore, the paper also points out the singular wavelet paths: the paths associated with nodes (subbands) such that no stationarization can occur and those for which dependency reduction is not expected to hold true.

The paper is organized as follows. In Section II, we present notation and some preliminary assumptions used throughout the paper. Section III deals with the conditions

under which stationarization occurs for wavelet packet coefficients. Section IV provides asymptotic results on the decorrelation and the higher order dependency reduction that can be reached for the class of random processes introduced in Section III. Section V discusses the contribution of such an analysis to characterize generalized fractionally integrated random processes from the identification of singular wavelet packet paths. Section VI concludes the work and mentions some prospects.

## II. PRELIMINARY NOTATION AND ASSUMPTIONS

### A. Wavelet paths and subbands

Throughout the paper,  $H_0$  denotes a *scaling filter* and  $W_0$  is the *scaling function* associated with this filter. Similarly,  $H_1$  denotes the *wavelet filter* associated with  $H_0$  and  $W_1$  is the *wavelet function* associated with  $H_1$ , [27], [28]. The functional subspace of  $L^2(\mathbb{R})$  generated from the translated versions of  $W_0$  is denoted by  $\mathbf{W}_{0,0}$  (input space).

The wavelet packet decomposition assumes splitting  $\mathbf{W}_{0,0}$  into orthogonal subspaces  $\mathbf{W}_{1,0}$  and  $\mathbf{W}_{1,1}$  and continuing the splitting recursively on every *subband*  $\mathbf{W}_{j,n}$  for obtaining  $\mathbf{W}_{j+1,2n}$  and  $\mathbf{W}_{j+1,2n+1}$ . The splitting involves paraunitary wavelet filters  $(H_0, H_1)$  at every decomposition level (see [28] for details).

Let  $(H_{\epsilon_1}, H_{\epsilon_2}, \dots, H_{\epsilon_j})$  be the sequence of wavelet filters successively applied for decomposing  $\mathbf{W}_{0,0}$ , with  $\epsilon_\ell \in \{0, 1\}$  for every  $\ell \in \{1, 2, \dots, j\}$ . This sequence is associated with a unique wavelet packet subspace  $\mathbf{W}_{j,n}$  (*subband*), where

$$n = n(j) = \sum_{\ell=1}^j \epsilon_\ell 2^{j-\ell}. \quad (1)$$

A full wavelet packet path  $\mathcal{P}$ , starting from the root node  $\mathbf{W}_{0,0}$ , is specified from the infinite sequence  $(H_{\epsilon_\ell})_{\ell \in \mathbb{N}}$  of filters to be used in the recursive splitting scheme. This is equivalent to associating path  $\mathcal{P}$  with the infinite binary sequence  $\{\epsilon_\ell\}_{\ell \in \mathbb{N}}$  indexing the sequence of filters given above. At the decomposition level  $j$ , in path  $\mathcal{P}$ , the frequency index is  $n = n_{\mathcal{P}}(j)$  given by Eq. (1) where the corresponding subsequence  $\{\epsilon_\ell\}_{\ell=1,2,\dots,j}$  is composed of the  $j$  first term of the infinite binary sequence associated with  $\mathcal{P}$ . For convenience, we will use either the notation  $\mathcal{P} = \{\epsilon_j\}_{j \in \mathbb{N}}$  or the notation  $\mathcal{P} = (\mathbf{W}_{0,0}, \{\mathbf{W}_{j,n_{\mathcal{P}}(j)}\}_{j \in \mathbb{N}})$  involving the subbands associated with the filtering sequence  $(H_{\epsilon_\ell})_{\ell \in \mathbb{N}}$ , see [4] for more details on wavelet packet path characterization.

The *approximation path*, hereafter denoted by  $\mathcal{P}_0$ , is the wavelet packet path such that:  $n = n_{\mathcal{P}_0}(j) = 0, \forall j \in \mathbb{N}$ . All other paths are said to be *detail paths*.

Assume that orthonormal wavelet transforms are concerned. Then  $\mathbf{W}_{j,n}$  is generated from the sequence of wavelet functions  $\{W_{j,n,k} : k \in \mathbb{Z}\}$ , where  $W_{j,n,k}(\cdot) = W_{j,n}(\cdot - 2^j k)$ , with  $W_{j,n}$  satisfying, in the Fourier domain<sup>1</sup>:

$$\mathcal{F} W_{j,n}(\omega) = 2^{j/2} \left[ \prod_{\ell=1}^j H_{\epsilon_\ell}(2^{\ell-1} \omega) \right] \mathcal{F} W_0(\omega), \quad (2)$$

<sup>1</sup>Fourier transform:  $\mathcal{F}f(\omega) = \int_{\mathbb{R}} f(t) e^{-i\omega t} dt$  if  $f \in L^1(\mathbb{R})$ .

the above equality holding in  $L^2(\mathbb{R})$  sense. Note that depending on the binary sequence  $\{\varepsilon_\ell\}_{\ell=1,2,\dots,j}$  associated with  $n$ , then the decomposition involved is

- a *discrete wavelet transform* when  $n \in \{0,1\}$  at every decomposition level,
- a *full discrete wavelet packet transform* when  $n \in \{0,1,\dots,2^\ell-1\}$  for every decomposition level  $\ell$ ,  $1 \leq \ell \leq j$ .

We consider the wavelet packet framework in the following.

### B. Vanishing moments

A wavelet function  $W_1$  is said to have  $r$  vanishing moments if

$$\int_{\mathbb{R}} t^m W_1(t) dt = 0, \quad \text{for every } m = 0, 1, \dots, r-1. \quad (3)$$

Let  $\mathcal{M}_{j,n,k}(m)$  denote the  $(m+1)$ -th *moment* of the function  $W_{j,n,k}$ ,  $m \in \{0, 1, 2, \dots\}$ :

$$\mathcal{M}_{j,n,k}(m) = \int_{\mathbb{R}} t^m W_{j,n,k}(t) dt. \quad (4)$$

Under the condition given by Eq. (3) and if we assume that the paraunitary filters  $(H_0, H_1)$  have finite impulse responses, then for  $n \neq 0$ , we have:

$$\mathcal{M}_{j,n,k}(m) = 0, \quad \text{for every } m = 0, 1, \dots, r-1. \quad (5)$$

It follows that  $W_{j,n,k}$  has at least  $r$  vanishing moments.

Note that by construction, every wavelet function  $W_1$  is with at least one vanishing moment whereas the contrary holds for the scaling function  $W_0$  (see for instance [27]). Thus, we have  $\mathcal{M}_{j,0,k}(0) \neq 0$  whereas  $\mathcal{M}_{j,n,k}(0) = 0$  for every  $n \in \{1, 2, \dots, 2^j - 1\}$ .

### C. Preliminary assumptions

Let  $X$  be a second order real random process, continuous in quadratic mean. Let  $R(t, s) = \mathbb{E}[X(t)X(s)]$  be the autocorrelation function of  $X$ . Assume that

$$\iint_{\mathbb{R}^2} R(t, s) W_{j,n,k}(t) W_{j,n,k}(s) dt ds < \infty. \quad (6)$$

Then, the projection of  $X$  on subband  $\mathbf{W}_{j,n}$  yields coefficients that define a discrete second order real random process  $c_{j,n}$ , with:

$$c_{j,n}[k] = \int_{\mathbb{R}} X(t) W_{j,n,k}(t) dt, \quad k \in \mathbb{Z}, \quad (7)$$

The statistical properties of  $c_{j,n}$  depend on the analytical form of  $R(t, s)$ . In what follows, we assume that this autocorrelation function admits the following expansion:

#### Condition (ACF)

The *AutoCorrelation Function*, *ACF*, can be written in the form:

$$R(t, s) = \underbrace{F(t) + F(s)}_{\text{Projective terms}} + \underbrace{\mathcal{S}(t-s)}_{\text{Stationary term}} + \underbrace{\sum_{1 \leq p, q \leq M} \alpha_{p,q} t^p s^q}_{\text{Bivariate-polynomial}} \quad (8)$$

with

$$F(t) W_{j,n,k}(t) \in L^1(\mathbb{R})$$

and

$$t^p s^q W_{j,n,k}(t) W_{j,n,\ell}(s) \in L^1(\mathbb{R}^2)$$

for every  $1 \leq p, q \leq M$ , where  $\mathcal{S}$  is an even function.

Due to the symmetry of the autocorrelation function, we can order the terms of the bivariate-polynomial with respect to one of the variables  $t, s$ . Therefore, we call  $M$ , the *degree of the bivariate-polynomial* involved in Eq. (8). It is worth noticing that in Condition (ACF) above, no restriction is imposed on  $M$ . In practice, when the degree  $M$  is fixed (for modeling purpose, for instance), then only bivariate monomial terms with high degrees are excluded from the bivariate-polynomial of Condition (ACF). Now, when  $M$  tends to infinity, the expansion Eq. (8) involves Taylor-like expansions and analytic power series when both  $F$  and  $\mathcal{S}$  are polynomial functions and monomial coefficients are chosen adequately. From these considerations, we have that Condition (ACF) is reasonable for approximating the autocorrelation functions of a wide class of random signals and time series.

From now on, under Eq. (8), the Fourier transform  $\mathcal{F}\mathcal{S}$  of  $\mathcal{S}$ , when it exists, will refer as the *spectrum* of  $X$ . The justification of this terminology follows as a consequence of the properties of the wavelet packet coefficients of  $X$ : an adequate choice of the wavelet function leads to stationary wavelet coefficients for the decomposition of  $X$  and makes possible the definition of a power spectral density of the form given by

$$\gamma_{j,n}(\omega) = \frac{1}{2\pi} \mathcal{F}\mathcal{S}(\omega) |\mathcal{F}W_{j,n}(\omega)|^2$$

for the subband  $\mathbf{W}_{j,n}$  wavelet packet coefficients, see Theorem 1 for details.

A specific class of random processes considered in this work is characterized by spectra  $\mathcal{F}\mathcal{S}$  with the following form:

$$\mathcal{F}\mathcal{S}(\omega) = \Lambda(\omega) \times \Psi(\omega), \quad (9)$$

where  $\omega \in [-\pi, \pi]$  and  $\Psi$  is a “fractional  $K$ -factor” function with the form given by

$$\Psi(\omega) = \prod_{k=1}^K \frac{\sigma^2}{\{2|\cos\omega - \psi_k|\}^{2\delta_k}}. \quad (10)$$

This form allows for a substantial generalization of the definition of a wide class of fractionally integrated random processes, as illustrated below.

Note that the spectrum  $\mathcal{F}\mathcal{S}$  of such a random process may be unbounded in neighborhoods of many frequency points, denoting the presence of possibly several long memory parameters governing the behavior of this random process. Depending on the form of function  $\Lambda$ , we will use the following terminology from the literature on stochastic processes and time series:

#### 1) $K$ -factor FI random processes

If  $\Lambda(\omega) = 1$  in Eq. (9), then  $\mathcal{F}\mathcal{S}$  is the spectrum of a  $K$ -factor Fractionally Integrated (FI) random process. The case of the standard FI random process corresponds to  $K = 1$  and  $\psi_1 = 1$ .

## 2) $K$ -factor FEXP random processes

Let

$$\Lambda(\omega) = \prod_{\ell=0}^L e^{\eta_{\ell} v_{\ell}(\omega)}, \quad (11)$$

where  $(v_{\ell})_{\ell=0,1,\dots,L}$  are piecewise continuous functions satisfying  $v_0 \equiv 1$ ,  $v_{\ell}(\omega) = v_{\ell}(-\omega)$ , and the matrix  $(v_{\ell}(2\pi m/N))_{1 \leq m \leq m^*, 0 \leq \ell \leq L}$  is non-singular for any natural number  $N$ . Then  $\mathcal{F}\mathcal{S}$  is the spectrum of a  $K$ -factor Fractional EXPonential (FEXP) random process. The case of the standard FEXP random process [29] follows by letting  $K = 1$  and  $\psi_1 = 1$ .

## 3) $K$ -factor GARMA random processes

The spectrum  $\mathcal{F}\mathcal{S}$  of a  $K$ -factor GARMA (Gegenbauer AutoRegressive Moving Average) random process is of the form Eq. (9) [30], [31], [32], with  $\Psi$  given by Eq. (10) and

$$\Lambda(\omega) = \frac{\Theta(e^{-i\omega})}{\Phi(e^{-i\omega})} = \frac{\left|1 - \sum_{p=1}^P \theta_p e^{-ip\omega}\right|^2}{\left|1 - \sum_{q=1}^Q \phi_q e^{-iq\omega}\right|^2}. \quad (12)$$

Parameters  $(\phi_{\ell})_{\ell}$  model the contribution of autoregressive terms and parameters  $(\theta_{\ell})_{\ell}$  correspond to moving average contributions.

Note that when  $K = 1$  and  $\psi_1 = 1$ , then  $\mathcal{F}\mathcal{S}$  is known from the literature as the spectrum of an ARFIMA (AutoRegressive Fractionally Integrated Moving Average) random process. Further details concerning the properties of ARFIMA random processes can be found in [33], [34], [35], [36], [37], [38], [39], [40], among others.

Before detailing the characterization of the wavelet packet coefficients of random processes having autocorrelation functions with the form Eq. (8), let us provide some examples of such random processes.

### Example 1 (WSS random processes)

For a Wide Sense Stationary (WSS) random process  $X(t)$ , we have  $R_X(t, s) = R_X(t-s, 0) \equiv R_X(t-s) = \mathcal{S}(t-s)$  (thus,  $F = 0$  and the autocorrelation admits no bivariate polynomial term). As a consequence, the terminology of ‘‘spectrum’’ used for  $\mathcal{F}\mathcal{S}$  is the ‘‘natural’’ one since  $X$  is stationary.

A stationary  $K$ -factor GARMA (Gegenbauer AutoRegressive Moving Average) model [30], [31], [32] satisfies in the time domain

$$\Phi(B) \prod_{k=1}^K (1 - 2\psi_k B + B^2)^{\delta_k} X(t) = \Theta(B) Z(t), \quad (13)$$

where  $Z(t)$  is a zero-mean white noise with variance  $\sigma^2$ , functions  $\Theta, \Phi$  are defined by  $\Theta(B) = I - \sum_{p=1}^P \theta_p B^p$ ,  $\Phi(B) = I - \sum_{q=1}^Q \phi_q B^q$ ,  $B$  is the *backshift operator*:  $BX(t) = X(t-1)$  and  $I$  represents the identity operator.

The following conditions are required in order to ensure wide sense stationarity for a random process  $X$  satisfying Eq. (13): first, parameters  $(\psi_k)_{k=1,2,\dots,K}$  are assumed to be distinct. In addition, one among the following conditions must holds true:

- i)  $\delta_k < 1/2$  and  $|\psi_k| < 1$ , for every  $k \in \{1, 2, \dots, K\}$ ,
- ii)  $\delta_k < 1/4$  and  $|\psi_k| = 1$ , for every  $k \in \{1, 2, \dots, K\}$ .

Finally, the zeros of functions  $\Theta(x)$  and  $\Phi(x)$  need to be distinct and must lie outside the unit circle. ■

### Example 2 (Trend and stationary autocorrelation)

Consider a random process  $Z$  with mean  $\mu_Z(t)$  and stationary autocorrelation function  $R_Z(t, s) \equiv R_Z(t-s)$ . Define a zero-mean random process  $X(t)$  from

$$X(t) = Z(t) - \mu_Z(t).$$

Then,  $X$  has autocorrelation function

$$R_X(t, s) = R_Z(t-s) - \mu_Z(t)\mu_Z(s).$$

Assume that  $\mu$  is a polynomial function,  $\mu_Z(t) = \sum_{k=0}^{M^*} a_k t^k$ , that does not reduce to a constant (otherwise,  $X$  is wide-sense stationary). Then an expression of the form Eq. (8) follows for the autocorrelation function of  $X$  by letting  $\mathcal{S}(t) = R_Z(t)$ ,  $F(t) = a_0^2/2 + a_0 \times \sum_{k=1}^{M^*} a_k t^k$  and the bivariate polynomial is  $\sum_{k=1}^{M^*} \sum_{\ell=1}^{M^*} a_k a_{\ell} t^k s^{\ell}$ . ■

### Example 3 (Polynomial moments)

The autocorrelation function of a random process with finite order polynomial moments satisfies the form Eq. (8). Consider for instance the polynomial random modulation defined by

$$X(t) = \sum_{k=0}^{M^{\#}} X_k t^k,$$

where  $(X_k)_{k=0,1,\dots,M^{\#}}$  is a sequence of zero-mean uncorrelated random variables. Let  $\sigma_k^2 = \mathbb{E}[X_k^2]$  for every  $k \in \{0, 1, \dots, M^{\#}\}$ . Then, random process  $X$  have autocorrelation function

$$R_X(t, s) = \sum_{k=0}^{M^{\#}} \sigma_k^2 (ts)^k,$$

which can be written in the form Eq. (8) by letting  $F(t) = \sigma_0^2/2$  and  $\mathcal{S} = 0$ . ■

### Example 4 (fBm and Linear transformations of fBms)

The nonstationary random processes given below have no bivariate polynomial terms in their autocorrelation expansion. In this respect, we only specify the close form of the projective and stationary terms.

#### 1) fBm random process

For a zero-mean fractional Brownian motion (fBm), we have  $F(t) = \mathcal{S}(t) = |t|^{2H}$  where  $H \geq 0$  is the Hurst parameter.

#### 2) fBmMA random process

Consider now a random process  $X$  resulting from the linear filtering of an fBm random process  $Z$  by using a filter with finite impulse response. We assume for the sake of simplicity the following first order moving average model (process  $X$  is an fBmMA random process):

$$X(t) = Z(t) + \alpha Z(t - T_0), \quad (14)$$

where  $Z(t)$  is a zero-mean fBm with Hurst parameter  $H$ . Then we have

$$\mathcal{S}(t) = (1 + \alpha^2)|t|^{2H} + \alpha(|t + T_0|^{2H} + |t - T_0|^{2H})$$

and

$$F(t) = (1 + \alpha)(|t|^{2H} + \alpha|t - T_0|^{2H}).$$

■

Example 5 below concern some random processes for which an expansion of the form Eq. (8), when it exists, does not necessarily simplify the derivation of the properties of wavelet coefficient autocorrelation functions.

#### Example 5 (Separable autocorrelation)

Many stochastic processes admit separable autocorrelation function:  $R(t, s) = v(t)v(s)$ . For instance, this is the case for the random cosine modulation  $X(t) = Z \cos \omega_0 t$ , where  $Z$  is a random variable: this random process have autocorrelation function

$$R_X(t, s) = \mathbb{E}[Z^2] \cos \omega_0 t \cos \omega_0 s.$$

Assume that function  $v$  has no finite order polynomial expansion<sup>2</sup> (example of the random modulation above). Then, deriving an expansion of the form Eq. (8) is not straightforward<sup>3</sup> and does not necessarily ease the wavelet coefficient analysis<sup>4</sup>. Nevertheless, the separability of the autocorrelation function with respect to variables  $t, s$  is useful for establishing the properties of the wavelet coefficient autocorrelations which, themselves, appear to be separable (see Example 5 [Continued]).

■

#### Remark 1

Some specific random processes have very intricate autocorrelations that are not separable and cannot be expanded in the form Eq. (8) straightforwardly. For instance, the autocorrelation function of a zero-mean multifractional Brownian motion (mfBm) is [41]:

$$R(t, s) = C(H(t), H(s)) (t^{H(t)+H(s)} + s^{H(t)+H(s)} - |t - s|^{H(t)+H(s)})$$

with

$$C(x, y) = \frac{\sqrt{\Gamma(2x+1)\Gamma(2y+1)\sin\pi x \sin\pi y}}{2\Gamma(x+y+1)\sin(\pi(x+y+1)/2)}.$$

This form involves several compositions of non trivial functions and does not follow Eq. (8), unless further simplifying assumptions are used concerning the behaviour of the time-dependent Hurst parameter. The class of mfBm random processes given above requires a specific case study for the derivation of the statistical properties of its wavelet packet coefficients.

■

<sup>2</sup>One can note that if  $v$  is a polynomial, expansion of the form Eq. (8) follows from monomial and binomial terms of the product  $v(t)v(s)$ , as it can be seen in Example 2.

<sup>3</sup>When  $v$  is an entire analytic function, expansion of the form Eq. (8) can be obtained from the power series of  $v$ .

<sup>4</sup>Remark 2 below emphasizes that a bivariate polynomial expansion with infinite order does not make it possible to simplify the wavelet coefficient autocorrelation function.

We now consider the higher order cumulants of the random process  $X$ . Let  $N \in \mathbb{N}$  and denote by

$$\text{cum}(t_1, t_2, \dots, t_N) = \text{cum}\{X(t_1), X(t_2), \dots, X(t_N)\},$$

the *cumulant* of order  $N$  of the random process  $X$ . We recall that the *cumulant generating function* can be seen as the logarithm of the *moment generating function* and the relationship between cumulants and moments is:

$$\begin{aligned} \text{cum}(t_1, t_2, \dots, t_N) &= \sum_{k=1}^N (-1)^{k-1} (k-1)! \times \mathbb{E} \left[ \prod_{\ell \in \mathfrak{s}_1} X(t_\ell) \right] \times \\ &\quad \mathbb{E} \left[ \prod_{\ell \in \mathfrak{s}_2} X(t_\ell) \right] \times \dots \times \mathbb{E} \left[ \prod_{\ell \in \mathfrak{s}_k} X(t_\ell) \right], \end{aligned} \quad (15)$$

where, for a given  $k$ ,  $1 \leq k \leq N$ , the summation extends over all partitions  $(\mathfrak{s}_1, \mathfrak{s}_2, \dots, \mathfrak{s}_k)$  of the set  $\{1, 2, \dots, N\}$ . The above cumulant is hereafter assumed to belong to  $L^1(\mathbb{R}^N) \cup L^2(\mathbb{R}^N)$ . For  $N \geq 2$ , we will also consider the following assumptions:

#### Condition (C1)

The cumulant  $\text{cum}(t_0, t_1, t_2, \dots, t_N)$  can be written in the form

$$\begin{aligned} \text{cum}(t_0, t_1, t_2, \dots, t_N) &= F^N(t_0, t_1, \dots, t_N) \\ &\quad + \mathcal{S}^N(t_1 - t_0, t_2 - t_0, \dots, t_N - t_0) \\ &\quad + \sum_{1 \leq q_0, q_1, \dots, q_N \leq M_N} \alpha_{q_0, q_1, \dots, q_N} t_0^{q_0} t_1^{q_1} \dots t_N^{q_N} \end{aligned} \quad (16)$$

where function  $F^N$ , the *projective term*, is of the form:

$$F^N(t_0, t_1, \dots, t_N) = \sum_{\ell=0}^N F_\ell(t_0, t_1, \dots, t_{\ell-1}, t_{\ell+1}, \dots, t_N)$$

and function  $\mathcal{S}^N$  (*stationary term*) is such that

$$\mathcal{S}^N(t_{k_1} - t_{k_0}, \dots, t_{k_N} - t_{k_0}) = \mathcal{S}^N(t_1 - t_0, \dots, t_N - t_0)$$

for any permutation  $\{k_0, k_1, \dots, k_N\}$  of  $\{0, 1, \dots, N\}$ .

#### Condition (C2)

$$F_\ell(t_1, t_2, \dots, t_N) \times W_{j,n,\ell_1}(t_1) \times W_{j,n,\ell_2}(t_2) \times \dots \times W_{j,n,\ell_N}(t_N) \in L^1(\mathbb{R}^N)$$

where  $F_\ell$  is any of the function involved in the sum defining projective term  $F^N$  and

$$t_0^{q_0} \dots t_N^{q_N} W_{j,n,\ell_0}(t_0) \dots W_{j,n,\ell_N}(t_N) \in L^1(\mathbb{R}^{N+1}).$$

### III. STATIONARIZATION

Subband wavelet and wavelet packet coefficients of stationary random processes are stationary [4], [5], [6], [7], [8]. For the class of nonstationary random processes satisfying Condition (ACF), this section presents theoretical results establishing their wide sense stationary, provided that the wavelet used has at least  $r \geq M + 1$  vanishing moments, where  $M$  is the degree of the bivariate-polynomial involved in Eq. (8). Furthermore, assume that: 1) conditions (C1), (C2) are satisfied and, 2) the sequence  $(M_N)_{N \geq 2}$

of multivariate-polynomial degrees involved in Eq. (16) is bounded, with  $M_\infty = \sup\{M_N : N \geq 2\}$ . Then, strict sense stationarity of the subband coefficients of  $X$  follows from wavelets having  $r \geq \max\{M, M_\infty\} + 1$  vanishing moments. These results are formalized in Theorems 1 and 2.

In the rest of the section, an upper index  $r$  (notation  $c_{j,n}^r$ ,  $R_{j,n}^r$ ,  $\text{cum}_{j,n}^r$  and  $\mathcal{M}_{j,n,k}^r$ ) will be used to specify that wavelet subbands are generated from a wavelet function  $\mathbf{W}_1$  having  $r$  vanishing moments. We will also assume that the wavelet filters are with finite impulse response. From this assumption and by using the notation of Section II-B, we have:  $\mathcal{M}_{j,n,k}^r(p) = 0$  for every  $n \neq 0$  and  $p = 0, 1, \dots, r-1$ .

#### A. Wide-sense stationarity

Let  $R_{j,n}^r$  be the autocorrelation function of the discrete random process  $c_{j,n}^r$  representing the wavelet packet coefficients of  $X$  on subband  $\mathbf{W}_{j,n}^r$ :

$$R_{j,n}^r[k, \ell] = \iint_{\mathbb{R}^2} R(t, s) W_{j,n,k}^r(t) W_{j,n,\ell}^r(s) dt ds. \quad (17)$$

Assume that condition **(ACF)** holds true. Then  $F(t) W_{j,n,k}^r(t) \in L^1(\mathbb{R})$ ,  $t^p s^q W_{j,n,k}^r(t) W_{j,n,\ell}^r(s) \in L^1(\mathbb{R}^2)$  and thus, we have:

$$\begin{aligned} R_{j,n}^r[k, \ell] &= \mathcal{M}_{j,n,\ell}^r(0) \times \left( \int_{\mathbb{R}} F(t) W_{j,n,k}^r(t) dt \right) \\ &+ \mathcal{M}_{j,n,k}^r(0) \times \left( \int_{\mathbb{R}} F(t) W_{j,n,\ell}^r(t) dt \right) \\ &+ \sum_{1 \leq p, q \leq M} \alpha_{p,q} \mathcal{M}_{j,n,k}^r(p) \times \mathcal{M}_{j,n,\ell}^r(q) \\ &+ \iint_{\mathbb{R} \times \mathbb{R}} \mathcal{S}(t-s) W_{j,n,k}^r(t) W_{j,n,\ell}^r(s) dt ds. \end{aligned}$$

Taking into account Eq. (5), it follows that  $R_{j,n}^r$  only depends on  $\mathcal{S}$  when  $n \neq 0$  and  $r \geq M+1$ :

$$R_{j,n}^r[k, \ell] = \iint_{\mathbb{R} \times \mathbb{R}} \mathcal{S}(t-s) W_{j,n,k}^r(t) W_{j,n,\ell}^r(s) dt ds. \quad (18)$$

If we assume further that:

$$\iint_{\mathbb{R}^2} \mathcal{S}(t-s) W_{j,n,k}^r(t) W_{j,n,\ell}^r(s) dt ds < \infty$$

and that  $\mathcal{S}$  has a Fourier transform (in  $L^1(\mathbb{R})$  or  $L^2(\mathbb{R})$  sense), then we have, using the same notation and with the above assumptions:

#### Theorem 1

*Under Condition (ACF), and if we assume that  $X$  is with zero-mean or have a polynomial mean of order  $M$ , then the discrete random sequence  $c_{j,n}^r, n \neq 0$ , is wide sense stationary for  $r \geq M+1$ :  $R_{j,n}^r[k, \ell] \equiv R_{j,n}^r[k-\ell]$ , with*

$$R_{j,n}^r[m] = \frac{1}{2\pi} \int_{\mathbb{R}} \mathcal{F}\mathcal{S}(\omega) \left| \mathcal{F}W_{j,n}^r(\omega) \right|^2 e^{i2^j m \omega} d\omega. \quad (19)$$

When the wavelet function used has enough vanishing moments, it follows from Theorem 1 that: for random processes with autocorrelation function given by Eq. (8), projective and bivariate-polynomial terms do not impact the

autocorrelation functions of the wavelet packet coefficients. From Theorem 1, function  $\mathcal{S}$  plays an important role in the wavelet coefficient autocorrelation function. As mentioned throughout the paper,  $\mathcal{S}$  is the contribution of a stationary term in the autocorrelation function of a - not necessarily stationary - random process. Hence, from Theorem 1, one can associate a *spectrum*  $\mathcal{F}\mathcal{S}$  to a nonstationary random processes that have wavelet autocorrelation restricted to the contribution of  $\mathcal{S}$ .

#### Remark 2

Consider the decomposition of a random process  $X$  with autocorrelation function  $R_X$  given by Eq. (8) and let  $\mathcal{P}$  be a wavelet packet path, with  $\mathcal{P} \neq \mathcal{P}_0$ .

If  $R_X$  admit no bivariate-polynomial term, then wavelet filters have the same stationarization effect. Indeed, only the first moment is required for annihilating the projective terms and all wavelets functions  $W_{j,n,k}$ ,  $n \neq 0$ , have their first moments that vanish:  $\mathcal{M}_{j,n,k}^r(0) = 0$  for every  $n \neq 0$  and every  $r \geq 1$ , by construction. In contrast, when a bivariate-polynomial term is present in  $R_X$ , then only wavelets with  $r \geq M+1$  vanishing moments have the desirable stationarization property. Note that some standard families of wavelet functions are synthesized so that their number of vanishing moments can be arbitrarily large (Example of Daubechies and spline wavelets). However, for these families, order  $r = \infty$  is not attainable in practice due to the fact that it requires for the wavelet filters to have infinite impulse responses.

The following discusses the consequences of Theorem 1 on the decompositions of 1) the random processes obtained by detrending a random process with stationary autocorrelation and 2) the fBm and fBmMA random processes.

#### Example 2 [Continued] (Trend and stationary autocorrelation)

If  $X$  is the random process defined in Example 2, with  $\mu_Z(t) = \sum_{k=0}^{M^*} a_k t^k$ , then stationarity follows for wavelet packet coefficients of  $X$  when the wavelet function has  $r \geq M^* + 1$  vanishing moments. ■

#### Example 4 [Continued] (fBm and Linear transformations of fBms)

##### 1) fBm random process

If  $X$  is a fractional Brownian motion with Hurst parameter  $H$ ,  $0 < H < 1$ , then only one vanishing moment is required for stationarity to holds true for its the wavelet coefficients. By applying the Fourier transform (from the distributional sense) to  $\mathcal{S}(t) = |t|^{2H}$ , we have

$$\mathcal{F}\mathcal{S}(\omega) = \frac{\sigma^2 D(H)}{|\omega|^{2H+1}}, \quad (20)$$

where  $D(H) = \Gamma(2H+1) \sin(\pi H)$  and  $\Gamma$  is the standard Gamma function. Note that the result above can be proven without referring to the theory of distributions, as performed in [11] (wavelet framework) and [3] (wavelet packet framework). Note also that the result above holds true for the detail wavelet packet subbands  $(\mathbf{W}_{j,n}^r)_{j \geq 1, n \neq 0}$ : the wavelet functions generating these

subbands have at least one vanishing moment, by construction. In contrast, the approximation subbands  $(W_{j,0}^r)_{j \geq 1}$  are characterized by scaling functions having no vanishing moment and the coefficients associated with these subbands remain nonstationary.

## 2) fBmMA random process

With the same remarks as above, the detail wavelet subbands of the fBmMA random process defined in Eq. (14) are stationary for  $r \geq 1$  and have autocorrelation function given by Eq (19), where

$$\mathcal{F} \mathcal{S}(\omega) = \frac{\sigma^2 D(H) (1 + \alpha^2 + 2\alpha \cos T_0 \omega)}{|\omega|^{2H+1}}. \quad (21)$$

■

The following example deals with the particular case of a separable autocorrelation.

### Example 5 [Continued] (Separable autocorrelation)

Let  $X$  be a random process having autocorrelation function  $R(t, s) = v(t)v(s)$ . Assume that  $v \in L^1(\mathbb{R}) \cup L^2(\mathbb{R})$  and that model Eq. (8) with a finite order  $M$  fails for  $R(t, s)$ .

Then the autocorrelation function of the sequence  $c_{j,n}^r$  associated with the decomposition of  $X$  have the following form:

$$R_{j,n}^r[k, \ell] = U_{j,n}^r[k] U_{j,n}^r[\ell]$$

with

$$U_{j,n}^r[k] = \frac{1}{2\pi} \int_{\mathbb{R}} \mathcal{F} v(\omega) \mathcal{F} W_{j,n}^r(-\omega) e^{i2^j k \omega} d\omega.$$

Consequently, we can conclude that  $c_{j,n}^r$  is nonstationary, in general.

Consider for instance the random modulation  $X(t)$  defined in Example 5. Assume that  $\mathbb{E}[Z^2] = 1$ . Then we have

$$U_{j,n}^r[k] = \frac{1}{2} \left( e^{i2^j \omega_0 k} \mathcal{F} W_{j,n}^r(-\omega_0) + e^{-i2^j \omega_0 k} \mathcal{F} W_{j,n}^r(\omega_0) \right)$$

and we derive  $U_{j,n}^r[k] = \mathcal{F} W_{j,n}^r(\omega_0) \cos 2^j \omega_0 k$  for even wavelet functions. Thereby, autocorrelation function of  $c_{j,n}^r$  is  $R_{j,n}^r[k, \ell] = \left( \mathcal{F} W_{j,n}^r(\omega_0) \right)^2 \cos 2^j \omega_0 k \cos 2^j \omega_0 \ell$ . This autocorrelation function have, up to a scaling factor, the same form that of  $X$ . ■

### B. Higher order stationarity

Let  $N$  be a natural number. The cumulant of order  $N+1$  of the random process  $c_{j,n}^r$  is given by

$$\begin{aligned} \text{cum}_{j,n}^r[k, \ell_1, \ell_2, \dots, \ell_N] &= \text{cum}\{c_{j,n}^r[k] c_{j,n}^r[\ell_1] c_{j,n}^r[\ell_2] \dots c_{j,n}^r[\ell_N]\} \\ &= \int_{\mathbb{R}^{N+1}} dt ds_1 ds_2 \dots ds_N \text{cum}(t, s_1, s_2, \dots, s_N) W_{j,n,k}^r(t) \\ &\quad W_{j,n,\ell_1}^r(s_1) W_{j,n,\ell_2}^r(s_2) \dots W_{j,n,\ell_N}^r(s_N). \end{aligned}$$

When  $N = 1$ , then  $\text{cum}_{j,n}^r[k, \ell] = R_{j,n}^r[k, \ell]$  and wavelet packets are wide sense stationary under condition (ACF) and further assumptions used in Section III-A. These assumptions are supposed to hold true in this section and we also assume that conditions (C1) and (C2) hold true. Then, if we proceed as in Section III-A and if we take into

account the null-moment condition given by Eq. (5), we obtain

$$\begin{aligned} \text{cum}_{j,n}^r[k, \ell_1, \ell_2, \dots, \ell_N] &= \int_{\mathbb{R}^{N+1}} dt ds_1 ds_2 \dots ds_N \mathcal{S}^N(s_1, s_2, \dots, s_N) W_{j,n,k}^r(t) \\ &\quad W_{j,n,\ell_1}^r(t + s_1) W_{j,n,\ell_2}^r(t + s_2) \dots W_{j,n,\ell_N}^r(t + s_N) \end{aligned}$$

for  $r \geq M_N + 1$ , that is:  $\text{cum}_{j,n}^r[k, \ell_1, \ell_2, \dots, \ell_N]$  depends only on the stationary term  $\mathcal{S}^N$ .

Furthermore, let us assume that:

$$\begin{aligned} \mathcal{S}^N(s_1, s_2, \dots, s_N) W_{j,n,k}^r(t) W_{j,n,\ell_1}^r(t + s_1) \\ \times W_{j,n,\ell_2}^r(t + s_2) \dots W_{j,n,\ell_N}^r(t + s_N) \in L^1(\mathbb{R}^{N+1}) \end{aligned}$$

and that the Fourier transform of  $\mathcal{S}^N$  exists. Then,  $\text{cum}_{j,n}^r[k, \ell_1, \ell_2, \dots, \ell_N]$  can be written in the following form given in Theorem 2.

### Theorem 2

*With the same assumptions used in Theorem 1, conditions (C1), (C2) and the assumptions used above for guaranteeing the existence of the different integrals involved in the expansion of the wavelet packet cumulants, we have:  $c_{j,n}^r$ ,  $n \neq 0$ , is strictly stationary for  $r \geq \max\{M, M_\infty\} + 1$ . For  $N \geq 1$ ,  $\text{cum}_{j,n}^r[k, \ell_1, \ell_2, \dots, \ell_N] \equiv \text{cum}_{j,n}^r[\ell_1 - k, \ell_2 - k, \dots, \ell_N - k]$ , with*

$$\begin{aligned} \text{cum}_{j,n}^r[k_1, k_2, \dots, k_N] &= \frac{1}{(2\pi)^N} \int_{\mathbb{R}^N} d\omega_1 d\omega_2 \dots d\omega_N \\ &\quad e^{-iM^j(k_1\omega_1 + k_2\omega_2 + \dots + k_N\omega_N)} \\ &\quad \mathcal{F} \mathcal{S}^N(-\omega_1, -\omega_2, \dots, -\omega_N) \\ &\quad \mathcal{F} W_{j,n}^r(-\omega_1 - \omega_2 - \dots - \omega_N) \\ &\quad \mathcal{F} W_{j,n}^r(\omega_1) \mathcal{F} W_{j,n}^r(\omega_2) \dots \mathcal{F} W_{j,n}^r(\omega_N). \end{aligned}$$

### Remark 3

In presence of nonstationary terms in the cumulants of  $X$ , the stationarization properties obtained in Theorems 1 and 2 are not likely to hold true for the approximation path  $\mathcal{P}_0$ . Indeed, as mentioned in Section II-B, any subband  $W_{j,n=0}^r$  for  $j \geq 1$  is generated from scaling functions  $\{W_{j,0,k}^r : k \in \mathbb{Z}\}$  having no vanishing moment. Thus the contribution of the projective and the multivariate polynomial terms do not annihilate: this implies that higher order stationarity does not occur for the approximation coefficients, in general.

## IV. DEPENDENCIES

This section presents some results concerning the capability of wavelet packets for decorrelating the coefficients of stochastic processes satisfying assumption (ACF). It provides additional results concerning higher order dependency reduction induced by wavelets on stochastic processes satisfying assumptions (ACF), (C1) and (C2).

From Theorem 1, the autocorrelation function of the wavelet packet coefficients can be written in the form of Eq. (19), under the assumption that the wavelet function  $W_1$  has  $r \geq M + 1$  vanishing moments. Distributing  $c_{j,n}^r$  as a sequence of decorrelated coefficients involves finding parameters that make  $R_{j,n}^r[m]$  vanishes for every  $m \in \mathbb{Z} \setminus \{0\}$ . Since no restrictions are imposed on  $\mathcal{F}\mathcal{S}$ , apart those required for integrability, then the parameters that govern the behavior of  $R_{j,n}^r$  are the shape and the support of  $\mathcal{F}W_{j,n}^r$ .

One can probably design a wavelet function, depending on the close form of  $\mathcal{F}\mathcal{S}$  so as to yield vanishing  $R_{j,n}^r[m]$ , for  $m \neq 0$ . In such a scenario, the wavelet function is computed adaptively with respect to the input random process spectrum, yielding a Karhunen-Loève-like expansion. The first limitation of this approach is that an adaptive consideration can be restrictive when a large class of stochastic processes is concerned. In addition, the above consideration is also limited because the spectrum  $\mathcal{F}\mathcal{S}$  is usually unknown.

In order to seek for wavelet decorrelating capability, unconditionally with respect to the input process, the remaining parameter is the size of the support of  $\mathcal{F}W_{j,n}^r$ . Indeed, by drastically reducing this size, we can expect to reduce the amplitude  $R_{j,n}^r[m]$ . Note that support reduction is the trick used to construct the Shannon wavelets: by dividing the support of  $\mathcal{F}W_{j,n}^r$  per 2 when  $j$  increases, spectrum  $\mathcal{F}\mathcal{S}$  is analyzed on a very tight frequency interval when  $j$  is large. In this respect, these wavelets provide us with a framework for analyzing wavelet decorrelating properties.

Theorem 3 below formalizes the above heuristic considerations. In this theorem, as well as in the rest of the paper, we need to describe  $\mathcal{F}W_{j,n}^r$  with an additional parameter that relates to the size of the support of  $\mathcal{F}W_{j,n}^r$  or, equivalently, describes how close  $\mathcal{F}W_{j,n}^r$  is, with respect to the corresponding Shannon wavelet function. For the sake of generality, this parameter needs to be different with the number  $r$  of vanishing moments of the wavelet function  $W_0$  since the support size of a function is not necessarily connected with the number of vanishing moments of this function. However, for the standard family of wavelet filters, the support size is linked to the number of wavelet vanishing moments [27] so that we can save notation. Section IV-A below provides this connection.

#### A. Wavelet order: the connection between the wavelet support size and the number of wavelet vanishing moments

Consider a filter with impulse response  $h_0 = (h_0[\ell])_{\ell \in \mathbb{Z}}$ . Let us define, up to a factor  $1/\sqrt{2}$ , the Fourier transform of  $h_0$  by:

$$H_0(\omega) = \frac{1}{\sqrt{2}} \sum_{\ell \in \mathbb{Z}} h_0[\ell] e^{-i\ell\omega}.$$

The non-negative integer  $s$  such that  $H_0$  admits the polynomial factorization [42]

$$H_0(\omega) = \left( \frac{1 + e^{-i\omega}}{2} \right)^s Q(e^{i\omega}), \quad (22)$$

is called the *filter order* and is also called *s-regularity* in [43]. The factorization given by Eq. (22) assumes that  $Q$  has no poles or zeros at  $\omega = \pi$ . From this factorization, it follows that order  $s$  describes the flatness of  $H_0$  at  $\omega = 0$  and  $\omega = \pi$ . It is worth stressing that  $H_0(\pi) = 0$  so that increasing the flatness of  $H_0$  at  $\omega = \pi$  involves reducing the support size of  $H_0$ .

Assume now that  $H_0$  is a *scaling filter* [27], [28], [43]. Let  $H_1$  be the *wavelet filter* associated with  $H_0$ :  $(H_0, H_1)$  is a couple of paraunitary filters. Then the same remark as above holds true for  $H_1$  due to this paraunitary condition: by inverting the role played by  $\omega = 0$  and  $\omega = \pi$ , it follows that increasing the flatness of  $H_0$  at  $\omega = \pi$  involves increasing the flatness of  $H_1$  at  $\omega = 0$  and, consequently, this implies reducing the support size of  $H_1$ . Furthermore, the wavelet function  $\mathcal{F}W_1$  inherits the above properties of  $H_1$  by taking into account that  $\mathcal{F}W_1(\omega) = H_1(\frac{\omega}{2})\mathcal{F}W_0(\frac{\omega}{2})$ . Moreover, parameter  $s$  is exactly the number of vanishing moments (parameter  $r$  in the previous sections) of the wavelet function  $W_1$ .

From the above analysis, we decide to recover the upper index  $r$  in the notation of wavelet functions (see Section III). However, this index will be written with the following conventional notation: “[ $r$ ]”, meaning that wavelet functions have  $r$  vanishing moments and are generated from, or generate paraunitary filters with order  $r$  (factorization given by Eq. (22)). These wavelets will be said with order  $r$ .

If we consider standard families of wavelet filters such as Daubechies, Symmlet or Battle-Lemarié spline filters, then the filters corresponding to  $r = 1$  and  $r = +\infty$  are respectively the Haar and the Shannon filters. These filters play an important role in describing the filter families mentioned above: for a given order  $1 < r < +\infty$ , the shapes of a couple of paraunitary filters  $(H_c^{[r]})_{c \in \{0,1\}}$  can be inferred by their closeness to the shapes of  $(H_c^{[1]})_{c \in \{0,1\}}$ , Haar paraunitary filters, or  $(H_c^S)_{c \in \{0,1\}}$ , Shannon paraunitary filters. This assumption is reasonable because when the order  $r$  increases from 1 to  $+\infty$ , then the sequence  $(H_c^{[r]})_{[r] \geq 1}$  converges almost everywhere to  $H_c^S$  and the shapes of  $(H_c^{[r]})_r$  vary smoothly between the shape of  $H_c^{[1]}$  and that of  $H_c^S$  on both frequency intervals  $]-\pi/2, \pi/2[$  and  $]-\pi, -\pi/2 \cup ]\pi/2, \pi[$  (see [3] for details).

#### B. Correlation structure

This section provides the decorrelating properties of wavelet packets for the coefficients of stochastic processes satisfying Condition (ACF). These properties are consequences of some asymptotic results, depending on the wavelet decomposition level and the wavelet order. Stating these asymptotic results require a whole wavelet packet path specification. This specification is addressed by using the notation of Section II-A.

Let  $[z]$  denotes the largest integer less than or equal to the real  $z$  and  $G$  be the permutation recursively defined by  $G(2\ell + \epsilon) = 3G(\ell) + \epsilon - 2 \lfloor \frac{G(\ell) + \epsilon}{2} \rfloor$ . With the same notation and assumptions proposed in Section III-A, we have:

**Theorem 3**

Let  $\mathcal{P} = \{\epsilon_j\}_{j \in \mathbb{N}} = \left( \mathbf{W}_{0,0}^{[r]}, \left\{ \mathbf{W}_{j,n_{\mathcal{P}}(j)}^{[r]} \right\}_{j \in \mathbb{N}} \right)$  be a path in the wavelet packet tree. Assume that  $\mathcal{P} \neq \mathcal{P}_0$  where  $\mathcal{P}_0$  is the approximation path.

Assume  $\mathcal{F}\mathcal{S}$  is continuous at the frequency  $\omega_{\mathcal{P}}$  defined by

$$\omega_{\mathcal{P}} = \lim_{j \rightarrow +\infty} \frac{G(n_{\mathcal{P}}(j))\pi}{2^j}. \quad (23)$$

Then the autocorrelation  $R_{j,n_{\mathcal{P}}(j)}^{[r]}$  of  $c_{j,n_{\mathcal{P}}(j)}^{[r]}$  uniformly satisfies:

$$\lim_{j \rightarrow +\infty} \left( \lim_{[r] \rightarrow +\infty} R_{j,n_{\mathcal{P}}(j)}^{[r]}[k] \right) = \mathcal{F}\mathcal{S}(\omega_{\mathcal{P}})\delta[k]. \quad (24)$$

*Proof:* Since the contributions of the nonstationary terms induced by the projective and the bivariate-polynomial terms annihilate when  $\mathcal{P} \neq \mathcal{P}_0$ , then the result follows by mimicking the proofs of [8, Proposition 1, Theorem 1]. ■

Let us consider the approximation path  $\mathcal{P}_0$ , the unique path not being concerned by Theorem 3. From Eq. (23),  $\mathcal{P}_0$  is a path associated with the limit frequency  $\omega = 0$ . More precisely, we have:

**Proposition 1**

$$\omega_{\mathcal{P}} = 0 \iff \mathcal{P} = \mathcal{P}_0.$$

*Proof:* Proposition 1 follows by noting that  $\mathcal{P}_0$  is the unique path such that the sequence  $(n_{\mathcal{P}}(j))_{j \in \mathbb{N}}$  of frequency indices associated with path  $\mathcal{P}$  can be upper bounded by a constant independent with  $j$ . ■

**Remark 4**

From Proposition 1, the approximation path is the unique path associated with frequency  $\omega = 0$ . By considering Theorem 3, it follows that asymptotic decorrelation cannot be expected in the neighborhood of the null frequency, even when  $\omega = 0$  is a continuity point of the spectrum  $\mathcal{F}\mathcal{S}$ : in addition with the contribution of  $\mathcal{F}\mathcal{S}$ , one must take into consideration, the terms issued from projective and bivariate polynomial terms.

The decorrelating properties stated in Theorem 3 are effective for detail wavelet packet paths associated with a spectrum  $\mathcal{F}\mathcal{S}$  which is regular once restricted on the support of  $\mathcal{F}W_{j,n}^r$ . In the following, we focus on the implications of Theorem 3 with respect to singularities in  $\mathcal{F}\mathcal{S}$ , respectively for  $K$ -factor FI - FEXP - GARMA, fBm and fBmMA type spectra, whether the input random process is stationary or not. From now on, a path associated with a frequency  $\nu$  will be denoted by  $\mathcal{P}_{\nu}$ .

**Example 6 (Singular paths)**

- 1) [Singular paths for FI, FEXP, fBm, fBmMA type spectra]

In the following contexts, the unique singular path is  $\mathcal{P}_0$  (function  $\mathcal{F}\mathcal{S}$  has a unique singular point, the pole  $\omega = 0$ ), when  $X$  is:

- a standard FI random process ( $K = 1$ ,  $\psi_1 = 1$  and  $\Lambda(\omega) = 1$  in Eq. (12)) that does not reduce to a white Gaussian process (thus,  $\delta \neq 0$ ),
- a standard FEXP random process ( $K = 1$ ,  $\psi_1 = 1$  and  $\Lambda(\omega)$  is given by Eq. (11)), with  $\delta \neq 0$ ,
- an fBm (see Eq. (20)) or an fBmMA random process (see Eq. (21)), with  $0 < H < 1$ .

Thus,  $c_{j,n_{\mathcal{P}}(j)}^{[r]}$  tends to decorrelate with  $j, r$  in every path  $\mathcal{P} \neq \mathcal{P}_0$ .

**2) [Singular paths for ARFIMA type spectrum]**

Assume that the above standard FI process is constrained to follow an ARMA representation, that is,  $X$  is associated with a rational function  $\Lambda(\omega)$  in Eq. (12) and as such, defines an ARFIMA random process. Assume that  $\Lambda(\omega)$  has  $Q^*$  distinct and non-null poles  $\omega_{\ell}$ ,  $\ell = 1, 2, \dots, Q^*$ .

Then,  $\mathcal{F}\mathcal{S}$  has poles  $\{0, \omega_1, \omega_2, \dots, \omega_{Q^*}\}$  and asymptotic decorrelation follows for all wavelet packet paths, except for the paths associated with the above poles.

**3) [Singular paths for  $K$ -factor FI, FEXP type spectrum]**

Assume that  $X$  is either a  $K$ -factor FI or a  $K$ -factor FEXP process. Then, it follows from Eq. (10) that the poles of  $\mathcal{F}\mathcal{S}$  are the Gegenbauer frequencies

$$\omega_{G_i} = \cos^{-1} \psi_k, \quad k = 1, 2, \dots, K.$$

Thus, the decorrelating properties of the wavelet packet coefficients concern paths  $\mathcal{P} \notin \{\mathcal{P}_0, \mathcal{P}_{\omega_{G_k}} : k = 1, 2, \dots, K\}$ .

**4) [Singular paths for  $K$ -factor GARMA type spectrum]**

For a  $K$ -factor GARMA random process, in addition with the pole  $\omega = 0$ , spectrum  $\mathcal{F}\mathcal{S}$  admits  $Q^*$  poles issued from function  $\Lambda$  and  $K$  poles resulting from function  $\Psi$ . When the number  $K^* = Q^* + K + 1$  of poles of  $\mathcal{F}\mathcal{S}$  is large, then, decorrelating the wavelet packets at small decomposition levels is unfeasible. Indeed, first, decorrelation is unreachable in at least the  $K^*$  paths. Second, due to the sharpness of the spectrum near its poles, the decorrelation requires very large decomposition levels in any path associated with a frequency that lie in the neighborhood of a spectrum pole. ■

**C. High order dependencies**

Similarly as in Section IV-B, the filter order plays an important role for high order dependency reduction through wavelet packet paths. When this order is maximal, that is, when the Shannon filters are concerned, we have:

$$\begin{aligned} & \text{cum}_{j,n}^S[k_1, k_2, \dots, k_N] \\ &= \frac{2^{j(N+1)/2}}{(2\pi)^N} \int_{\Delta_{j,G(n)}^N} d\omega_1 d\omega_2 \dots d\omega_N \\ & \quad e^{-i2^j(k_1\omega_1 + k_2\omega_2 + \dots + k_N\omega_N)} \\ & \quad \mathcal{F}\mathcal{S}^N(-\omega_1, -\omega_2, \dots, -\omega_N) \\ & \quad \mathbb{1}_{\Delta_{j,G(n)}}(\omega_1 + \omega_2 + \dots + \omega_N). \end{aligned} \quad (25)$$

where  $\Delta_{j,G(n)}^N = \underbrace{\Delta_{j,G(n)} \times \Delta_{j,G(n)} \times \dots \times \Delta_{j,G(n)}}_{N \text{ times}}$ .

The above cumulant involves computing the integrand in Eq. (25) on the narrow hypercube  $\Delta_{j,G(n)}^N$ .

Let  $\mathcal{P} = \left( \mathbf{W}_{0,0}^{[r]}, \left\{ \mathbf{W}_{j,n_{\mathcal{P}}(j)}^{[r]} \right\}_{j \in \mathbb{N}} \right)$  be a wavelet packet path,  $\mathcal{P} \neq \mathcal{P}_0$ . Let  $\omega_{\mathcal{P}}^N = \underbrace{(\omega_{\mathcal{P}}, \omega_{\mathcal{P}}, \dots, \omega_{\mathcal{P}})}_{N \text{ times}}$  and assume that there exists a neighbourhood  $V_{\omega_{\mathcal{P}}^N}$  of  $\omega_{\mathcal{P}}^N$  where  $\mathcal{S}^N$  is bounded. Then, in path  $\mathcal{P}$ ,  $\text{cum}_{j,n}^S[k_1, k_2, \dots, k_N]$ ,  $n = n_{\mathcal{P}}(j)$ , vanishes when  $j$  tends to  $\infty$ .

More precisely, let  $j_0$  such that  $\Delta_{j_0, G(n_{\mathcal{P}}(j_0))}^N \subset V_{\omega_{\mathcal{P}}^N}$ . Such a natural number exists since the diameter of  $\Delta_{j,G(n)}$  tends to 0 as  $j$  tends to infinity. Then, for any  $j \geq j_0$  and any natural number  $N \geq 2$ , we have:

$$\left| \text{cum}_{j,n}^S[k_1, k_2, \dots, k_N] \right| \leq \frac{\left\| \mathcal{F} \mathcal{S}_{V_{\omega_{\mathcal{P}}^N}}^N \right\|_{\infty}}{2^{j(N-1)/2}},$$

where  $\|f\|_{V,\infty}$  denotes the  $L^\infty$  norm of the function  $f$  restricted on support  $V$ . This gives the decay order of the the cumulant of order  $N \geq 2$  of  $X$ .

Finally, since the sequences of filters considered in this paper (see Eq. (22)) converge to the Shannon filters when their order  $r$  increases. Then, it is reasonable to expect that higher order filters will yield cumulant decay with a factor that is close to the decay induced by the Shannon filters.

#### D. Statistical dependencies: the role of the wavelet support sizes

From the theoretical results given above, it follows that statistical dependency reductions are strongly linked to the width of the support and the shape of  $\mathcal{F}W_{j,n}$ . By considering Eq. (2), then the dependency reduction also relates to the support size of the wavelet filters ( $H_0, H_1$ ), smaller support sizes being expected to yield stronger dependency reductions. This section provides details on the effective size of the support of the wavelet packet *equivalent filter*  $\mathbf{H}_{j,n}$  used for computing  $c_{j,n}$  from the input random process  $X$ . The illustrations concern Daubechies wavelets [28].

From Eq. (2), the equivalent filter applied to obtain the subband  $(j, n)$  wavelet coefficients is  $2^{j/2} \mathbf{H}_{j,n}(\omega)$ , with:

$$\mathbf{H}_{j,n}(\omega) = \left[ \prod_{\ell=1}^j H_{\epsilon_\ell}(2^{\ell-1} \omega) \right]. \quad (26)$$

When  $j = 1$ , the filter involved in Eq. (26) is either the scaling filter  $H_0$  (low-pass) or the wavelet filter  $H_1$  (high-pass). When  $j > 1$ , then  $\mathbf{H}_{j,n}$  is obtained from a combination of low-pass (when  $\epsilon_\ell = 0$ ) and high-pass (when  $\epsilon_\ell = 1$ ) filterings.

Figure 1 provides the graphs of  $\mathbf{H}_{j,n}^{[r]}(\omega)$  corresponding to some Daubechies filters, for a decomposition level  $j = 5$  and frequency indices  $n = 0, 20$ . We recall that the Daubechies scaling filter is obtained from Eq. (22), where  $Q$  is the Daubechies polynomial (see [28]). We recall also that the

Daubechies filters with order  $r = 1$  correspond to Haar filters defined for  $\epsilon \in \{0, 1\}$  by

$$H_\epsilon^{[1]}(\omega) = \frac{1}{2} \left( 1 + (1 - 2\epsilon)e^{-i\omega} \right) \quad (27)$$

and Daubechies filters with order  $r = \infty$  are the Shannon filters satisfying

$$H_\epsilon^S(\omega) = \sum_{\ell \in \mathbb{Z}} \mathbb{1}_{\Delta_\epsilon}(\omega - 2\pi\ell), \quad (28)$$

where  $\Delta_0 = [-\pi/2, \pi/2]$  and  $\Delta_1 = [-\pi, -\pi/2] \cup [\pi/2, \pi]$ .

As it can be seen in Figure 1, the Daubechies filters with  $[r] \geq 7$  have tight supports and the analysis performed in the paper suggests that dependency reductions can be reasonably attained in non-singular paths by using these filters: when analyzed on a small window size, a polyspectrum  $\mathcal{F} \mathcal{S}^N$  with regular shape can be seen as approximately constant on this support.

#### V. IDENTIFICATION OF A SINGULAR WAVELET PACKET PATH AND STOCHASTIC MODELING

From theorems 1, 2, 3 and the results of Section IV-C, it follows that under the null-moment condition, then detail wavelet packet coefficients become stationary in the wide sense (resp. strict sense) for processes satisfying conditions **(ACF)** (resp. **(C1)**, **(C2)**). In addition, correlation and higher order dependency reduction follow, provided that the decomposition level and the filter order are large enough. The above properties depend on the shape of the input process spectrum: decorrelation is attained in any detail path of the wavelet packet tree associated with a frequency  $\omega_{\mathcal{P}}$  that is a continuity point of the spectrum  $\mathcal{F} \mathcal{S}$ . Higher order cumulant decay occurs in any path associated with a frequency  $\omega_{\mathcal{P}}$  such that the polyspectrum  $\mathcal{F} \mathcal{S}^N$ ,  $N \geq 2$ , is bounded in a neighborhood of  $\omega_{\mathcal{P}}$ .

In practice, when modeling of random signals or time series is concerned, then it is convenient to first identify singularities of the spectrum. Note that when dealing with discrete time stochastic processes, the main singularities of interest are those associated with the poles of  $\mathcal{F} \mathcal{S}$ , and thus, associated with paths for which decorrelation is not expected to follow.

In order to identify the wavelet packet path  $\mathcal{P}^* = \{\epsilon_j^*\}_{j \in \mathbb{N}}$  associated with a given pole  $\omega^*$ ,  $0 \leq \omega^* < \pi$ , we need to find the sequence of frequency indices  $(n_{\mathcal{P}^*(j)})_{j \in \mathbb{N}}$  such that

$$n_{\mathcal{P}^*(j)} = \sum_{\ell=1}^j \epsilon_\ell^* 2^{j-\ell} \text{ and satisfying}$$

$$\lim_{j \rightarrow +\infty} \frac{G(n_{\mathcal{P}^*(j)})\pi}{2^j} = \omega^*.$$

From the partitioning of the interval  $[0, \pi[$  that realize the supports of the Shannon wavelet functions  $\mathcal{F}W_{j,n}$  when  $n = 0, 1, \dots, 2^j - 1$ , the index associated with frequency  $\omega^*$  is the non-negative integer:

$$n_{\mathcal{P}^*(j)} = G^{-1} \left( \left\lfloor \frac{2^j \omega^*}{\pi} \right\rfloor \right),$$

where  $G^{-1}$  is the inverse of the permutation  $G$ .

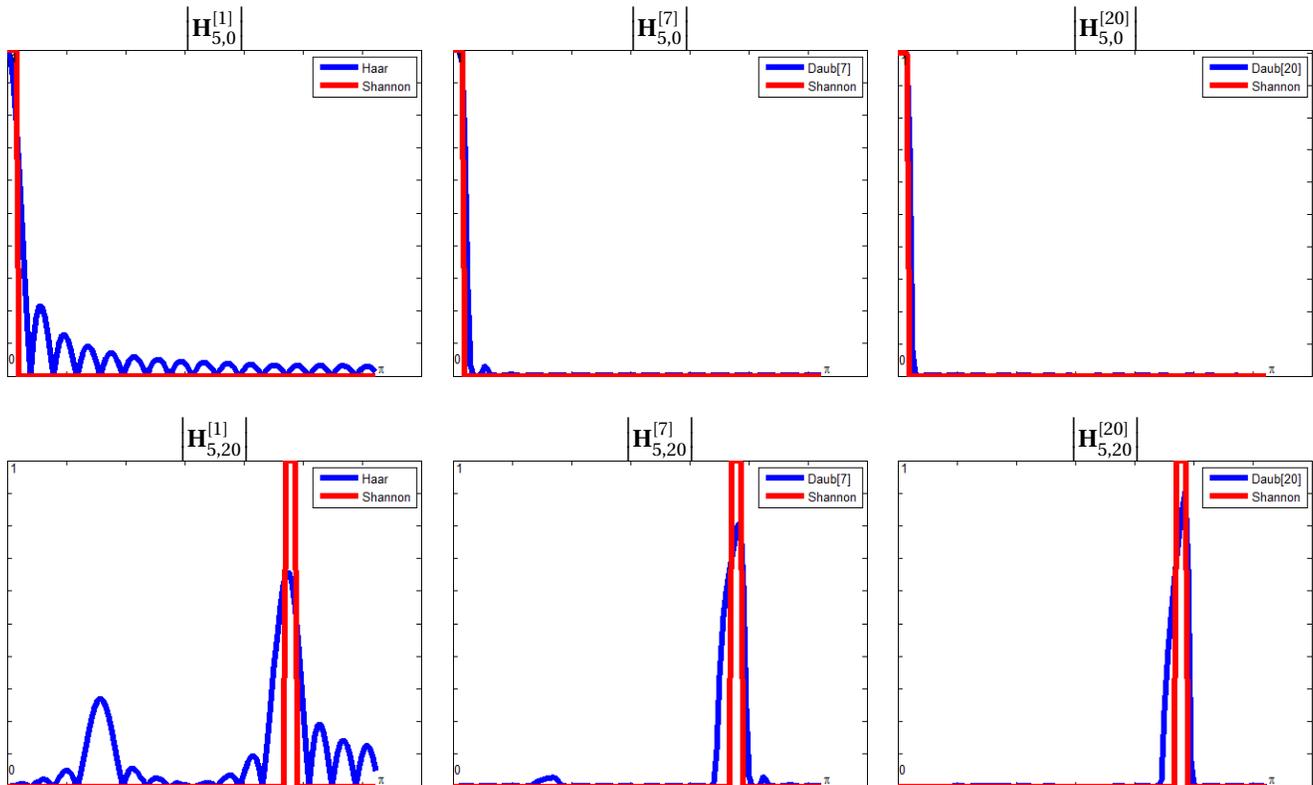


Fig. 1. Equivalent wavelet packet filters  $|\mathbf{H}_{5,0}^{[r]}(\omega)|$  after 5 decompositions. Filters are given (in blue) for orders  $r = 1, 7, 20$ , where  $r = 1$  corresponds to the Haar filters. Equivalent Shannon filters ( $r = \infty$ ) are represented (in red) as references corresponding to the “ideal” cases. The filtering sequences used for obtaining equivalent filters are those involved in the firsts 5 terms associated with paths  $\mathcal{P}_0$  and  $\mathcal{P}_{3\pi/4}$ . These sequences are  $(h_0^{[r]}, h_0^{[r]}, h_0^{[r]}, h_0^{[r]}, h_0^{[r]})$  for path  $\mathcal{P}_0$  and  $(h_1^{[r]}, h_0^{[r]}, h_1^{[r]}, h_0^{[r]}, h_0^{[r]})$  for path  $\mathcal{P}_{3\pi/4}$ .

The following example determines the path  $\mathcal{P}_{3\pi/4}$  associated with frequency  $\omega = 3\pi/4$ . For  $j \geq 2$ , path  $\mathcal{P}_{3\pi/4}^j$  is characterized by frequency indices

$$G(n_{\frac{3\pi}{4}}(j)) = \left[ 3 \times 2^{j-2} \right] = 2^{j-1} + 2^{j-2}.$$

Thus, by using the inverse permutation  $G^{-1}$  (see [3]), this path is characterized by the frequency indices  $n_{\frac{3\pi}{4}}(j) = 2^{j-1} + 2^{j-3}$  for  $j \geq 3$ .

In practical applications where we want to identify a fractional model (FI, ARFIMA, GARMA, etc) from the observation of a stochastic process, the presence of narrow peaks within a spectrum inform us with the presence of poles in the wavelet packet spectrum of the process under consideration. At this stage, the relevance of the spectrum estimation is crucial since it determines the selection of the model. In particular, in situations where some poles are very close to each other, the estimation method is required to discriminate the corresponding spectrum peaks.

The following experimental results show that *wavelet packet spectrum*, [3], can be particularly adapted to such situations. It is worth noticing that the role of these experiments concerns the relevancy of a spectrum estimated from data and do not concern “parameter estimation from a given spectrum estimate”. For more details on parameter estimators (in wavelet and the Fourier domain), the reader can refer to [44], [45], [46], see also [47], [48] for the robust

estimation of the autocorrelation function in presence of a long memory parameter.

Consider a random process with ARFIMA type obtained by letting  $K = 1$ ,  $\psi_1 = 1$ ,  $\sigma = 1$ ,  $\delta = 1/2$ ,  $\theta_1 = 1$ ,  $\phi_1 = e^{i\omega_1} + e^{i\omega_2}$ ,  $\phi_2 = e^{2i\omega_2} (1 - e^{-i\omega_2} (e^{i\omega_1} + e^{i\omega_2}))$ , in Eq (13), with  $\omega_1 = 3\pi/4 - \Delta\omega$  and  $\omega_2 = 3\pi/4 + \Delta\omega$ . This random process is nonstationary since its parameters do not satisfy conditions given in Example 1. However, a spectrum  $\mathcal{F}\mathcal{S}$  of the form given by Eq. (12) is associated with this random process, in the sense of [49]. According to the parameters given above, the function  $\Lambda$  involved in this spectrum admits 2 poles located at frequencies  $\omega_1$  and  $\omega_2$ .

When  $\Delta\omega$  is small, the poles  $\omega_1$  and  $\omega_2$  of  $\mathcal{F}\mathcal{S}$  are close to each other and lie on both sides of the “central” frequency  $3\pi/4$ . To analyze the sensitivity of the wavelet packet spectrum with respect to the closeness of poles in the spectrum, we let  $\Delta\omega \in \{\pi/48, \pi/72\}$ . We then synthesize the corresponding ARFIMA processes and compute their spectra by using the wavelet packet method described in [3].

Figure 2 presents the wavelet packet spectra obtained, as well as spectra computed by using Welch’s method (averaged and modified periodogram computed on the basis of the discrete Fourier transform). As it can be seen in this figure, the two poles are well detected by using the wavelet packet method (narrow peaks in the neighborhoods of the poles). In contrast, the peaks tend to overlap when

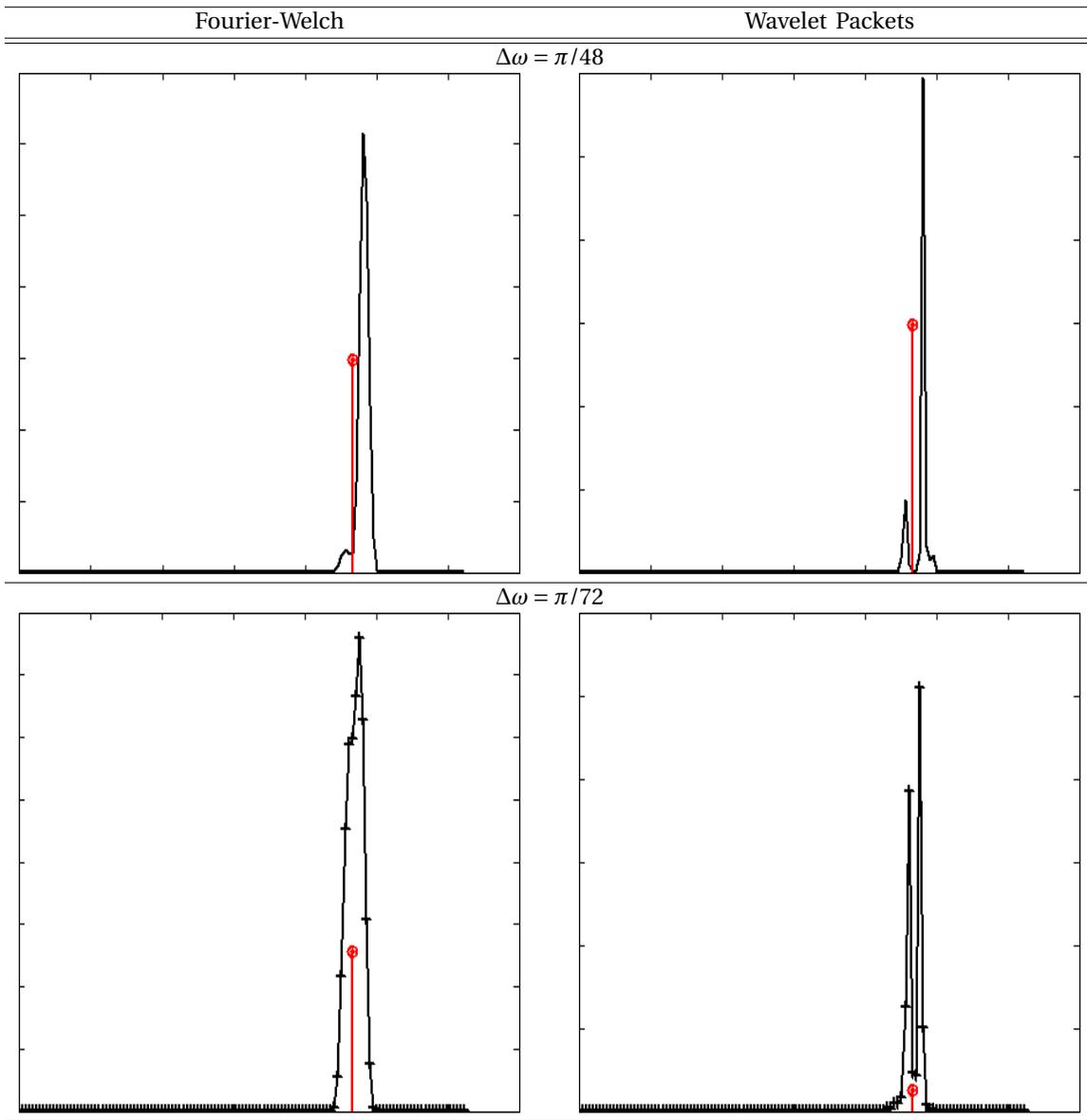


Fig. 2. “Fourier-Welch” and “Wavelet packet” spectra of ARFIMA processes, where any ARFIMA function  $B$  used admits two poles at frequencies  $\omega = 3\pi/4 \pm \Delta\omega$ . The graphs are given for  $\Delta\omega = \pi/48$  and  $\Delta\omega = \pi/72$ . The “central” frequency  $3\pi/4$  is marked by a vertical (red) line.

using Welch’s method for small values of  $\Delta\omega$ .

When considering the wavelet packet method, the positive half support of  $\mathcal{F}W_{j,n}^{[r]}$  has width  $\pi/2^j$  if we assume that ideal Shannon wavelet function is used ( $r = +\infty$ ). In this respect, we need a wavelet packet decomposition with  $J \geq 5$  (resp.  $J \geq 6$ ) levels in order to expect for differencing significantly spectra peaks that are located at a distance of  $2\Delta\omega$  from each other, when  $\Delta\omega = \pi/48$  (resp.  $\Delta\omega = \pi/72$ ).

Note that the spectra given in Figure 2 have been obtained with  $J = 6$  decomposition levels and the Daubechies wavelet with order  $r = 7$ . It follows from these spectra that order  $r = 7$  suffices for the discrimination of the two peaks. Thus, wavelet packet spectrum is efficient for characterizing spectra of generalized fractional random processes. The main concluding remark for this section is the following: for the wavelet packet spectrum estimation, one needs to choose a sufficiently large  $J$  and order  $r$  to avoid confusing

some spectra peaks which are close to each other.

From the above experimental results, it follows that wavelet packets are relevant for the analysis of second order random processes and well adapted for the characterization of generalized fractional random processes.

## VI. CONCLUSION

The paper has investigated the statistical properties of the wavelet packet transform of second order random processes. The analysis performed has highlighted the important role played by the wavelet vanishing-moment condition in the stationarization property of the wavelet packet transform. The decorrelating properties are strongly linked on 1) the support width and 2) the shape of the Fourier transform of the wavelet packet functions.

Theoretical results stating the statistical dependency reduction in the wavelet packet domain hold true for a

wide class of nonstationarity random processes. This class is characterized by some cumulant expansions having as nonstationary contributions, projective and bivariate polynomial terms given in the paper. These results are path-dependent: the coefficients associated with any random process  $X$  pertaining to the class-of-interest become stationary and tend to decorrelate in a path  $\mathcal{P} \neq \mathcal{P}_0$  when the decomposition level and the filter order increases, provided that the limit frequency  $\omega_{\mathcal{P}}$  is not a singular point of the wavelet packet spectrum associated with  $X$ .

When the wavelet packet coefficients have become stationary the impact of non-summability of the autocorrelation function of some subband wavelet packet coefficients has been analyzed. The analysis has highlighted that non-summability of the autocorrelation function associated with some subbands results in singular wavelet packet paths: these paths are associated with frequencies such that the spectrum of the input random process is unbounded in the neighborhood of these specific frequencies. Stochastic modeling for processes yielding singular wavelet packet paths can thus be addressed by associating generalized fractional processes having poles that are located at the above specific frequencies.

By exhibiting singular wavelet packet paths and by providing their interpretation with respect to the form of the polyspectra of the input random process, the paper has made a contribution in the analysis, modeling and synthesis of generalized fractional processes. For these random processes, the paper have also emphasized the relevancy of the wavelet packet based spectrum estimation, in comparison with Fourier based spectrum.

#### ACKNOWLEDGEMENT

The authors are very grateful to the anonymous reviewers and the associate editor for their insightful and useful comments.

#### REFERENCES

- [1] R. Averkamp and C. Houdré, "A note on the discrete wavelet transform of second-order processes," *Information Theory, IEEE Transactions on*, vol. 46, no. 4, pp. 1673 – 1676, Jul 2000.
- [2] P. F. Craigmile and D. B. Percival, "Asymptotic decorrelation of between-scale wavelet coefficients," *IEEE Transactions on Information Theory*, vol. 51, no. 3, pp. 1039 – 1048, Mar. 2005.
- [3] A. M. Atto, D. Pastor, and G. Mercier, "Wavelet packets of fractional brownian motion: Asymptotic analysis and spectrum estimation," *IEEE Transactions on Information Theory*, vol. 56, no. 9, Sep. 2010.
- [4] A. M. Atto and D. Pastor, "Central limit theorems for wavelet packet decompositions of stationary random processes," *IEEE Transactions on Signal Processing*, vol. 58, no. 2, pp. 896 – 901, Feb. 2010.
- [5] C. Houdré, "Wavelets, probability, and statistics: Some bridges," *In Wavelets: Mathematics and applications*, CRC Press, pp. 365 – 398, 1994.
- [6] J. Zhang and G. Walter, "A wavelet-based KL-like expansion for wide-sense stationary random processes," *IEEE Transactions on Signal Processing*, vol. 42, no. 7, pp. 1737 – 1745, July 1994.
- [7] D. Leporini and J.-C. Pesquet, "High-order wavelet packets and cumulant field analysis," *IEEE Transactions on Information Theory*, vol. 45, no. 3, pp. 863 – 877, Apr. 1999.
- [8] A. M. Atto, D. Pastor, and A. Isar, "On the statistical decorrelation of the wavelet packet coefficients of a band-limited wide-sense stationary random process," *Signal Processing*, vol. 87, no. 10, pp. 2320 – 2335, Oct. 2007.
- [9] P. Flandrin, "Wavelet analysis and synthesis of fractional brownian motion," *IEEE Transactions on Information Theory*, vol. 38, no. 2, pp. 910 – 917, Mar. 1992.
- [10] A. H. Tewfik and M. Kim, "Correlation structure of the discrete wavelet coefficients of fractional brownian motion," *IEEE Transactions on Information Theory*, vol. 38, no. 2, pp. 904 – 909, Mar. 1992.
- [11] E. Masry, "The wavelet transform of stochastic processes with stationary increments and its application to fractional brownian motion," *IEEE Transactions on Information Theory*, vol. 39, no. 1, pp. 260 – 264, Jan. 1993.
- [12] R. W. Dijkerman and R. R. Mazumdar, "On the correlation structure of the wavelet coefficients of fractional brownian motion," *IEEE Transactions on Information Theory*, vol. 40, no. 5, pp. 1609 – 1612, Sep. 1994.
- [13] S. Cambanis and C. Houdré, "On the continuous wavelet transform of second-order random processes," *IEEE Transactions on Information Theory*, vol. 41, no. 3, pp. 628 – 642, May 1995.
- [14] E. J. McCoy and A. T. Walden, "Wavelet analysis and synthesis of stationary long-memory processes," *Journal of Computational and Graphical Statistics*, vol. 5, pp. 26 – 56, 1996.
- [15] M. Vannucci and F. Corradi, "A review of wavelet in biomedical applications," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 61, no. 4, pp. 971 – 986, 1999.
- [16] R. Averkamp and C. Houdré, "Some distributional properties of the continuous wavelet transform of random processes," *IEEE Transactions on Information Theory*, vol. 44, no. 3, pp. 1111 – 1124, May 1998.
- [17] T. Kato and E. Masry, "On the spectral density of the wavelet transform of fractional brownian motion," *Journal of Time Series Analysis*, vol. 20, no. 50, pp. 559 – 563, 1999.
- [18] M. J. Jensen, "An alternative maximum likelihood estimator of long-memory processes using compactly supported wavelets," *Journal of Economic Dynamics and Control*, vol. 24, no. 3, pp. 361 – 387, Mar. 2000.
- [19] P. Abry, P. Flandrin, M. S. Taqqu, and D. Veitch, "Self-similarity and long-range dependence through the wavelet lens," *In Theory and applications of long-Range Dependence*, Birkhäuser Boston, pp. 527 – 556, 2002.
- [20] S. Touati and J.-C. Pesquet, "Some results on the wavelet packet decomposition of nonstationary processes," *EURASIP Journal on Applied Signal Processing*, vol. 2002, no. 11, pp. 1289 – 1295, Nov. 2002.
- [21] T.-H. Li and H.-S. Oh, "Wavelet spectrum and its characterization property for random processes," *IEEE Transactions on Information Theory*, vol. 48, no. 11, pp. 2922 – 2937, Nov 2002.
- [22] G. Didier and V. Pipiras, "Adaptive wavelet decompositions of stationary time series," *Journal of Time Series Analysis*, vol. 31, no. 3, pp. 182 – 209, 2010.
- [23] A. S. Willsky, "Multiresolution markov models for signal and image processing," *Proceedings of the IEEE*, vol. 90, no. 8, pp. 1396 – 1458, Aug. 2002.
- [24] K. Daoudi, A. B. Frakt, and A. S. Willsky, "Multiscale autoregressive models and wavelets," *IEEE Transactions on Information Theory*, vol. 45, no. 3, pp. 828 – 845, Apr. 1999.
- [25] R. W. Dijkerman, R. R. Mazumdar, and A. Bagchi, "Reciprocal processes on a tree-modeling and estimation issues," *IEEE Transactions on Automatic Control*, vol. 40, no. 2, pp. 330 – 335, Feb. 1995.
- [26] R. W. Dijkerman and R. R. Mazumdar, "Wavelet representations of stochastic processes and multiresolution stochastic models," *IEEE Transactions on Signal Processing*, vol. 42, no. 7, pp. 1640 – 1652, Jul. 1994.
- [27] S. Mallat, *A wavelet tour of signal processing, second edition*. Academic Press, 1999.
- [28] I. Daubechies, *Ten lectures on wavelets*. SIAM, Philadelphia, PA, 1992.
- [29] J. Beran, "Fitting long-memory models by generalized linear regression," *Biometrika*, vol. 80, no. 4, pp. 817 – 822, 1993.
- [30] H. L. Gray, N.-F. Zhang, and W. A. Woodward, "On generalized fractional processes," *Journal of Time Series Analysis*, vol. 10, no. 3, pp. 233 – 257, May 1989.
- [31] W. A. Woodward, Q. C. Cheng, and H. L. Gray, "A k-factor gamma long-memory model," *Journal of Time Series Analysis*, vol. 19, no. 4, pp. 485 – 504, Jul. 1998.
- [32] R. Ramachandran and P. Beaumont, "Robust estimation of gamma model parameters with an application to cointegration among interest rates of industrialized countries," *Computational Economics*, vol. 17, no. 2-3, pp. 179 – 201, 2001.

- [33] C. W. J. Granger and R. Joyeux, "An introduction to long-memory time series models and fractional differencing," *Journal of Time Series Analysis*, vol. 1, no. 1, pp. 15 – 29, Jan. 1980.
- [34] J. R. M. Hosking, "Fractional differencing," *Biometrika*, vol. 68, no. 1, pp. 165 – 176, 1981.
- [35] F. Sowell, "Maximum likelihood estimation of stationary univariate fractionally integrated time series models," *Journal of Econometrics*, vol. 53, no. 1-3, pp. 165 – 176, 1992.
- [36] —, "Modeling long-run behavior with the fractional arima model," *Journal of Monetary Economics*, vol. 29, no. 2, pp. 277 – 302, Apr. 1992.
- [37] C.-F. Chung, "A note on calculating the autocovariances of the fractionally integrated arma models," *Economics Letters*, vol. 45, no. 3, pp. 0165 – 1765, 1994.
- [38] J. Beran, *Statistics for Long-Memory Processes*. Chapman and Hall/CRC, 1994.
- [39] —, "Maximum likelihood estimation of the differencing parameter for invertible short and long memory autoregressive integrated moving average models," *Journal of the Royal Statistical Society. Series B. Methodological*, vol. 57, no. 4, pp. 659 – 672, 1995.
- [40] P. Brockwell and R. Davis, *Introduction to Time Series and Forecasting*. Springer, 2nd edition, 2002.
- [41] A. Ayache, S. Cohen, and J. L. Vehel, "The covariance structure of multifractional brownian motion, with application to long range dependence," *IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP*, vol. 6, pp. 3810 – 3813, June 2000.
- [42] C. S. Burrus, R. A. Gopinath, and H. Guo, *Introduction to Wavelets and Wavelet Transforms: A Primer*. Prentice Hall, 1998.
- [43] —, *Introduction to Wavelets and Wavelet Transforms: A Primer*. Prentice Hall, 1998.
- [44] O. Kouamo, C. Lévy-Leduc, and E. Moulines, "Central limit theorem for the robust log-regression wavelet estimation of the memory parameter in the gaussian semi-parametric context," *Preprint, [Available Online]*, 2011.
- [45] G. Faÿ, E. Moulines, F. Roueff, and M. S. Taqqu, "Estimators of long-memory: Fourier versus wavelets," *Journal of Econometrics*, vol. 151, no. 2, pp. 159 – 177, 2009.
- [46] D. Veitch and P. Abry, "A wavelet-based joint estimator of the parameters of long-range dependence," *IEEE Transactions on Information Theory*, vol. 45, no. 3, pp. 878 – 897, apr 1999.
- [47] C. Lévy-Leduc, H. Boistard, E. Moulines, M. S. Taqqu, and V. A. Reisen, "Robust estimation of the scale and of the autocovariance function of gaussian short and long-range dependent processes," *Journal of Time Series Analysis*, vol. 32, no. 2, pp. 135 – 156, 2011.
- [48] —, "Asymptotic properties of u-processes under long-range dependence," *Annals of Statistics*, vol. 39, no. 3, pp. 1399 – 1426, 2011.
- [49] A. M. Yaglom, "Correlation theory of processes with random stationary  $n$ th increments," *American Mathematical Society Translations (Series 2)*, vol. 8, pp. 87 – 141, 1958.