

# Simulation of Gegenbauer processes using wavelet packets

Collet J.J. \*      Fadili J.M. †

## Abstract

In this paper, we propose to study the synthesis of Gegenbauer processes using the wavelet packets transform. In order to simulate 1-factor Gegenbauer process, we introduce an original algorithm, inspired by the one proposed by Coifman and Wickerhauser [CW92], to adaptively search for the best-ortho-basis in the wavelet packet library where the covariance matrix of the transformed process is nearly diagonal. Our method clearly outperforms the one recently proposed by [Whi01], is very fast, does not depend on the wavelet choice, and is not very sensitive to the length of the time series. From these first results we propose an algorithm to build bases to simulate  $k$ -factor Gegenbauer processes. Given the simplicity of programming and running, we feel the general practitioner will be attracted to our simulator. Finally we evaluate the approximation due to the fact that we consider the wavelet packet coefficients as uncorrelated. An empirical study is carried out which supports our results.

**Keywords:** Gegenbauer process, Wavelet packet transform, Best-basis, Autocovariance

## 1 Introduction

The simulation of long memory processes is an issue of a paramount importance in many statistical problems. In the time domain, there exist different methods devoted to this task (see [Ber94] for a non exhaustive review of them). Alternative efficient approaches, which operate in the frequency domain, were also proposed (see [DH87] and [Ber94]). More recently, owing to their scale-invariance property, wavelets have since widely adopted as a natural tool for analyzing and synthesizing  $1/f$  long-memory processes. They were demonstrated to provide almost Karhunen-Loève expansion of such processes [Wor96].

The simulation of fractional differenced Gaussian noise (fdGn) using discrete wavelet transform (DWT) has been studied by [MW96]. This kind of processes is characterized by an unbounded power spectral density (PSD) at 0. The proposed method lies on the fact that the DWT approximately decorrelates long memory processes (see [DT93, TK92, Wor96, Jen99, PW00]). The orthonormal wavelet decomposition "only" ensures approximate decorrelation. These approximations have been studied in [Wor90, Fla92, Dij94, TK92, Wor96, Jen99, Jen00] for a variety of  $1/f$  long memory processes.

The DWT is only adapted to processes whose PSD is unbounded at the origin. Gegenbauer processes (sometimes also called seasonal persistent processes) are also long memory processes and are characterized by an unbounded PSD. The main difference with the fdGn processes is that the singularities of the PSD of the Gegenbauer processes can be located at one or many frequencies in the Nyquist domain, not necessary at the origin. Therefore, a natural tool to analyze such processes appears to be the *wavelet packet transform*, which is a generalization of the wavelet transform. The wavelets packets adaptively divides the frequency axis in separate intervals of various sizes. They segment unconditionally the frequency axis and are uniformly translated in time. Moreover, a discrete time series of size  $N$  is decomposed in more than  $2^{N/2}$  wavelet packet bases. Among these bases, one is a very good candidate to whiten the series and then almost diagonalizes the covariance of the seasonal process.

---

\*School of Economics and Finance, Queensland University of Technology, GP0 Box 2434, Brisbane QLD 4001, Australia, j.collet@qut.edu.au

†GREYC UMR CNRS 6072, ENSICAEN 6, Bd du Maréchal Juin, 14050 Caen Cedex, France, Jalal.Fadili@greyc.ensicaen.fr

Recently, Mallat, Zhang and Papanicolaou [MZP98], and, following their work, Donoho, Mallat and von Sachs [DMvS98], studied the idea of estimating the covariance of locally stationary processes by approximating the covariance of the process by a covariance which is almost diagonal in a specially constructed basis (cosine packets for their locally stationary processes) using an adaptation Coifman-Wickerhauser best ortho-basis algorithm. Our work here can be seen as the spectral dual of theirs, since we are interested in studying the covariance of seasonal processes in the wavelet packet domain.

To the best of our knowledge, the simulation of the Gegenbauer process using the wavelet packet transform has been first studied by Whitcher [Whi01]. The DWPT creates a redundant collection of wavelet coefficients at each level of the transform organized in a binary tree structure. Different methods exist to determine the best candidate orthonormal basis. Whitcher used a method which depends on both the location of the singularity and the wavelet used in the DWPT. Once the basis found, it remains to apply the DWPT using the same approximation as in [MW96].

This method consists first in considering the square gain function of the wavelet filter associated with each wavelet packet coefficient sufficiently small at the Gegenbauer frequency. Then a pruning of this family is done to obtain the ortho-basis. The main advantage of this method is its simplicity and its rapidity. However several points are still questionable and must be eluded. First the notion "sufficiently small" implies the introduction of a threshold which seems to depend both on the used wavelet and the length of the simulated series. None indication is given to choose this threshold which remains ad hoc. Furthermore, it is not clear why the basis should depend on the wavelet. Lastly and more annoying, this method inherently leads to an overpartitioning of the spectra which depends on the wavelet and the threshold considered (see Figure 4 in [Whi01]). Hence, it could be much more interesting to build, for each Gegenbauer process, an unique valid basis for all wavelets.

In this article, we propose an alternative way to determine the appropriate basis for the simulation of 1-factor Gegenbauer process and then the simulation of  $k$ -factor Gegenbauer process. Indeed while it is natural that this basis depends on the location of the singularity, it is not clear why this basis should depend on the wavelet used. We propose to use an algorithm inspired by the algorithm of [CW92] to find this basis. The main characteristic of this algorithm is that it returns basis depending only on the location of the Gegenbauer frequency unlike the method of construction proposed by Whitcher which provides basis depending both on the location of the singularity and the length of the wavelet. To point out the role of the wavelet used, we will study the decorrelation properties of the wavelet packet coefficients of a Gegenbauer process when it is expressed in this basis. In particular, the influence of the wavelet regularity and the location of the singularity on the decorrelation decay speed will be established.

The organization of this paper is as follows. After some preliminaries and notations related to the wavelet packets theory (Section 2.1) and to the Gegenbauer process (Section 2.2), we will define the algorithm of Coifman and Wickerhauser and the cost function we propose (Section 3.1 and 3.2). We develop an algorithm to build appropriate basis to simulate 1-factor Gegenbauer process (Section 3.3). This method will be used to determine basis for the simulation of  $k$ -factor processes (Section 3.4). For this two constructions we illustrate our results with some examples. In an other section we will evaluate the approximation due to the fact that we consider the wavelet packet coefficient as uncorrelated (Section 4). We illustrate our results with some simulations.

## 2 Preliminaries

### 2.1 The wavelet packet transform

Wavelet packets were introduced by Coifman, Meyer and Wickerhauser [CMW92], by generalizing the link between multi-resolution approximations and wavelets. Let the sequence of functions defined recursively as follows:

$$\psi_{j+1}^{2p}(t) = \sum_{n=-\infty}^{\infty} h(n)\psi_j^p(t - 2^j n) \quad (1)$$

$$\psi_{j+1}^{2p+1}(t) = \sum_{n=-\infty}^{\infty} g(n)\psi_j^p(t - 2^j n) \quad (2)$$

for  $j \in \mathbb{N}$  and  $p = 0, \dots, 2^j - 1$ , where  $h$  and  $g$  are the conjugate pair of quadrature mirror filters (QMF). At the first scale, the functions  $\psi_0$  and  $\psi_1$  can be respectively identified with the father and the mother wavelets  $\phi$  and  $\psi_1$  with the classical properties (among others):

$$\int \phi(t) = 1, \int \psi(t) = 0 \quad (3)$$

The collection of translated, dilated and normalized functions  $\psi_{j,n}^p \stackrel{\text{def}}{=} 2^{-j/2} \psi_p(2^{-j} - n)$  makes up what we call the (multi-scale) wavelet packets associated to the QMFs  $h$  and  $g$ .  $j \in \mathbb{N}$  is the scale index,  $p = 0, \dots, 2^j - 1$  can be identified with a frequency index and  $k$  is the position index. It has been proved (see e.g. Wickerhauser (1994)) that if  $\{\psi_j^{p,n}\}_{n \in \mathbb{Z}}$  is an orthonormal basis of a space  $\mathbf{V}_j$ , then the family  $\{\psi_j^{2p,n}, \psi_j^{2p+1,n}\}_{n \in \mathbb{Z}}$  is also an orthonormal basis of  $\mathbf{V}_j$ .

The recursive splitting of vector spaces is represented in a binary tree. To each node  $(j, p)$ , with  $j \in \mathbb{N}$  and  $p = 0, \dots, 2^j - 1$ , we associate a space  $\mathbf{V}_j^p$  with the orthonormal basis  $\{\psi_j^p(t - 2^j n)\}_{n \in \mathbb{Z}}$ . As the splitting relations creates two orthogonal basis, it is obvious that  $\mathbf{V}_j^p = \mathbf{V}_{j+1}^{2p} \oplus \mathbf{V}_{j+1}^{2p+1}$ .

We call admissible tree any binary tree where each node has either 0 or 2 children. If  $\{j_i, p_i\}_{i \in I}$  are the leaves of an admissible tree, then, we can verify that the spaces  $\{\mathbf{V}_{j_i}^{p_i}\}_{i \in I}$  are mutually orthogonal and add-up to:

$$\mathbf{L}^2(\mathbb{R}) \ni \mathbf{V}_0 = \bigoplus_{i=1}^I \mathbf{V}_{j_i}^{p_i}. \quad (4)$$

The wavelet packet representation is overcomplete. That is, there are many subsets of wavelet packets which constitute orthonormal bases for the original space  $\mathbf{V}_0$  (typically more than  $2^{2^{J-1}}$  for a binary tree of depth  $J$ ). While they form large library, these bases can be easily organized in a binary tree and efficiently searched for extreme points of certain cost functions, see [CW92] for details. Such a search algorithm and associated cost function are at the heart of this paper.

In the following we call the collection  $(\mathcal{B}_{j_i}^{p_i})_{i=1, \dots, I}$  the basis of  $\mathbf{L}^2(\mathbb{R})$  and we denote it  $\mathcal{B}$  and the tree for which the collection  $\{j_i, p_i\}_{i \in I}$  are the leaves, the associated tree.

Given a basis  $\mathcal{B}$  and its associated tree  $\mathcal{T}$  it is possible to decompose any function  $f$  of  $\mathbf{L}^2(\mathbb{R})$  in the basis  $\mathcal{B}$ . At each node  $(j_i, p_i)$  of the tree  $\mathcal{T}$ , the wavelet packet coefficients  $W_{j_i}^{p_i}(n)$  of  $f$  in the subspace  $\mathbf{V}_{j_i}^{p_i}$  at position  $n$  are given by the inner product:

$$W_{j_i}^{p_i}(n) = \int \psi_{j_i}^{p_i}(t - 2^{j_i} n) f(t) dt. \quad (5)$$

For a discrete signal of  $N$  equally-spaced samples, the wavelet packet transform is calculated using a fast filter bank algorithm that requires  $O(N \log N)$  operations. The interested reader may refer to the books of Mallat [Mal98] and Wickerhauser [Wic94] for more details about the wavelet packet transform.

## 2.2 Gegenbauer process

The  $k$ -factor Gegenbauer process is a  $1/f$ -type process introduced by Gray *et al.* [GZW89]. The spectral density  $f$  of a such process  $(X_t)_t$  is given by for all  $|\lambda| \leq 1/2$

$$f(\lambda) = \frac{\sigma_\varepsilon^2}{2\pi} \prod_{i=1}^k (2|\cos 2\pi\lambda - \cos 2\pi\nu_i|)^{-2d_i}, \quad (6)$$

where  $k$  is a finite integer and  $0 < d_i < 1/2$  if  $0 < |\nu_i| < 1/2$  and  $0 < d_i < 1/4$  if  $|\nu_i| = 0$  for  $i = 1, \dots, k$ . The parameter  $d_i$  and  $\nu_i$  are respectively called the memory parameter and the Gegenbauer frequency. The  $k$ -factor Gegenbauer process is a generalization of the fractionally differenced Gaussian white noise process (see [Hos81] and [GJ80]) in the sense that the spectral density is unbounded at  $k$  different frequencies not necessary located in 0.

The Gegenbauer process  $(X_t)_t$  is related to a white noise process  $(\varepsilon_t)_t$  with mean 0 and variance  $\sigma_\varepsilon^2$  through the relationship

$$\prod_{i=1}^k (I - 2\nu_i B + B^2)^{d_i} X_t = \varepsilon_t, \quad (7)$$

where  $BX_t = X_{t-1}$  and  $\nu_i = \cos 2\pi\nu_i$ .

The main characteristic of the Gegenbauer processes in the time domain is the slow decays of autocovariance function. In the case of a 1-factor Gegenbauer process, Gray *et al.* [GZW89] and then Chung [Chu96] exhibit the asymptotic behavior of the autocovariance function,

$$\rho(h) \sim h^{2d-1} \cos(2\pi\nu h) \quad \text{as } h \rightarrow \infty \quad (8)$$

The next section is devoted to the construction of the best basis diagonalizing the covariance of a  $N$ -sample realization of a Gegenbauer process with the convention  $N = 2^J$ .

### 3 Best-basis construction algorithm

#### 3.1 Approximate Diagonalization in an Best-Ortho-basis

Let  $X_t$  be a stationary Gegenbauer process and  $\Gamma$  its covariance matrix. Let  $\gamma_{i,j}[\mathcal{B}]$  the entries of  $\Gamma[\mathcal{B}] = \mathcal{B}^T \Gamma \mathcal{B}$ ; the covariance matrix of the coordinates  $W$  of  $X_t$  in the ortho-basis  $\mathcal{B}$ . We define diagonalization as an optimization of the functional:

$$\max_{\mathcal{B}} \mathcal{E}(\mathcal{B}) = \max_{\mathcal{B}} \sum_i e(\gamma_{ii}[\mathcal{B}])$$

where  $e$  is a convex function. The optimization formulation of diagonalization is not widely used, presumably because it generally doesn't help in computing diagonalizations. Optimization of an arbitrary objective  $\mathcal{E}$  over finite libraries of orthogonal bases - the cosine packets library and the wavelet packets library - is not a problem with good algorithmic solutions. Wickerhauser [Wic91] suggested applying these libraries in problems related to covariance estimation. He proposed the notion of selecting a "best basis" for representing a covariance by optimization of the "entropy functional"  $\mathcal{E}(\mathcal{B})$  over all bases in a restricted library. Authors in [MZP98], developed a proposal which uses the specific choice  $e_2(\gamma) = \gamma^2$ . In the Wickerhauser formulation, one is optimizing over a finite library and there will not generally be a basis in this library which exactly diagonalizes  $\Gamma$ . Then the different convex functions  $e(\gamma)$  may end up picking different bases. The quadratic cost function  $e_2$  has a special interpretation in this context: it leads to a basis which best diagonalizes  $\Gamma$  in a least-squares sense [DMvS98].

This cost function  $e_2$  is closely related to the Hilbert-Schmidt (HS) norm of the diagonalization error. Indeed, define the operator  $D_{\mathcal{B}}\Gamma = \mathcal{B} \text{diag}(\Gamma[\mathcal{B}]) \mathcal{B}^T$ . In other words, this is the operator on matrices formed by rotating into basis  $\mathcal{B}$ , then discarding off-diagonal elements, then rotating back into the natural basis. A fundamental result proved in [DMvS98] is that the basis optimizing  $\mathcal{E}_2 = \sum_i \gamma_{ii}^2[\mathcal{B}]$  gives a best approximate diagonalization and the optimal value  $\mathcal{E}_2$  measures the error of diagonalization, that is:

$$\arg \min \|\Gamma - D_{\mathcal{B}}\Gamma\|_{HS}^2 = \arg \min \left\{ \|\Gamma\|_{HS}^2 - \sum_i \gamma_{ii}^2[\mathcal{B}] \right\} = \arg \max \sum_i \gamma_{ii}^2[\mathcal{B}]$$

#### 3.2 Proposed Algorithm

Even if the the approach developed in [MZP98, DMvS98] wa specialized to the case of  $e_2$ , it is not really tied to the specific entropy measure; other "additive" convex measures can be accomodated such as the  $l_1$  norm. Thus, we here consider the sum of the wavelet packet coefficients variances as the additive cost function  $\mathcal{E}_2 = \sum_i \gamma_{ii}[\mathcal{B}] = \sum_{i \in I} \mathbb{V}[W_{j_i}^{p_i}]$  (i.e.  $e_1(\gamma) = |\gamma|$ ). The best basis  $\mathcal{B}_j^p$  of  $\mathbf{V}_j^p$  is the basis that minimizes the wavelet packet coefficients variances, among all the bases of  $\mathbf{V}_j^p$  that can be constructed from the tree. The construction of best bases can be accomplished efficiently using the recursive Coifman-Wickerhauser algorithm (CW algorithm) defined by

$$\mathcal{B}_j^p = \begin{cases} \mathcal{B}_{j+1}^{2p} \cup \mathcal{B}_{j+1}^{2p+1} & \text{if } \mathbb{V}[W_{j+1}^{2p}] + \mathbb{V}[W_{j+1}^{2p+1}] < \mathbb{V}[W_j^p], \\ \mathcal{B}_j^p & \text{if } \mathbb{V}[W_{j+1}^{2p}] + \mathbb{V}[W_{j+1}^{2p+1}] \geq \mathbb{V}[W_j^p]. \end{cases} \quad (9)$$

The chosen criterion lies on the comparison between variance of wavelet packet coefficients. It needs the calculation of these variances. In the following, we will modify this criterion for two main reasons.

On the one hand to calculate these variances we need to choose a wavelet. This choice implies that the basis we construct in the following will depend on this wavelet. Proceeding this way, we cannot expect to find an unique basis, independently of the choice of the wavelet.

On the other hand, to avoid the a priori choice of a wavelet, we may consider the case of a wavelet as a perfect band-pass filter in the same vein as in [MW96]. Unfortunately, in this case, one must compute the variances of each wavelet packet coefficient, i.e. the calculation of the integral of the spectral density of the Gegenbauer process over dyadic intervals. Because of the expression of the spectral density of the Gegenbauer processes, it is not possible to obtain an analytic expression of this integral and the numerical evaluation of this integral will be of a prohibitive computational burden.

We then define for the wavelet packet coefficients  $W_{j+1}^{2p}$  and  $W_{j+1}^{2p+1}$ , a new type of wavelet variance denoted by  $\overline{\mathbb{V}[W_{j+1}^{2p}]}$  and  $\overline{\mathbb{V}[W_{j+1}^{2p+1}]}$ , as follows:

$$\overline{\mathbb{V}[W_{j+1}^{2p}]} = \begin{cases} 0 & \text{if } \mathbb{V}[W_{j+1}^{2p}] \leq A_0 \mathbb{V}[W_{j+1}^{2p+1}] \\ \mathbb{V}[W_{j+1}^{2p}] & \text{otherwise.} \end{cases} \quad (10)$$

$$\overline{\mathbb{V}[W_{j+1}^{2p+1}]} = \begin{cases} 0 & \text{if } \mathbb{V}[W_{j+1}^{2p+1}] \leq A_0 \mathbb{V}[W_{j+1}^{2p}] \\ \mathbb{V}[W_{j+1}^{2p+1}] & \text{otherwise.} \end{cases} \quad (11)$$

where  $A_0$  is a fixed constant that will be determined later.

In the following, when we write that  $\overline{\mathbb{V}[W_{j+1}^{2p}]} = 0$  or  $\overline{\mathbb{V}[W_{j+1}^{2p+1}]} = 0$ , it will mean respectively that  $\mathbb{V}[W_{j+1}^{2p}] \leq A_0 \mathbb{V}[W_{j+1}^{2p+1}]$  or  $\mathbb{V}[W_{j+1}^{2p+1}] \leq A_0 \mathbb{V}[W_{j+1}^{2p}]$ . In these cases we will also use respectively the notations,

$$\mathbb{V}[W_{j+1}^{2p}] \ll \mathbb{V}[W_{j+1}^{2p+1}] \quad \text{and} \quad \mathbb{V}[W_{j+1}^{2p+1}] \ll \mathbb{V}[W_{j+1}^{2p}]$$

Using the criterion defined above, algorithm (9) becomes,

$$\mathcal{B}_j^p = \begin{cases} \mathcal{B}_{j+1}^{2p} \cup \mathcal{B}_{j+1}^{2p+1} & \text{if } \overline{\mathbb{V}[W_{j+1}^{2p}]} = 0 \text{ or } \overline{\mathbb{V}[W_{j+1}^{2p+1}]} = 0, \\ \mathcal{B}_j^p & \text{Otherwise.} \end{cases} \quad (12)$$

In the following, we use this algorithm to build the best-ortho-basis for a Gegenbauer process.

### 3.3 The 1-factor case

It is natural to build the best basis according to the shape of the spectral density of our process. More precisely, the basis is a function of the location of the singularities. It means that in the case of 1-factor Gegenbauer process, the basis depends directly on the value of the Gegenbauer frequency. Using the notations defined in the previous section, the recursive construction is summarized in the following proposition,

**Proposition 3.1** *If  $(X_t)_t$  is a stationary 1-factor Gegenbauer process, with parameters  $(d, \nu, \sigma)$  then, at node  $(j, p)$ , if the frequency  $\nu$  is in the interval  $I_j^p = [\frac{p}{2^j}, \frac{p+1}{2^j}[$ , then*

$$\overline{\mathbb{V}[W_{j+1}^{2p}]} = 0 \quad \text{or} \quad \overline{\mathbb{V}[W_{j+1}^{2p+1}]} = 0,$$

and consequently for algorithm (12),

$$\mathcal{B}_j^p = \mathcal{B}_{j+1}^{2p} \cup \mathcal{B}_{j+1}^{2p+1}.$$

Furthermore, if the frequency  $\nu$  is in both the closure of the intervals  $I_{j+1}^{2p}$  and  $I_{j+1}^{2p+1}$ , then,

$$\overline{\mathbb{V}[W_{j+2}^{4p+1}]} = 0 \quad \text{and} \quad \overline{\mathbb{V}[W_{j+2}^{4p+2}]} = 0,$$

and consequently for algorithm (12),

$$\mathcal{B}_j^p = \mathcal{B}_{j+2}^{4p} \cup \mathcal{B}_{j+2}^{4p+1} \cup \mathcal{B}_{j+2}^{4p+2} \cup \mathcal{B}_{j+2}^{4p+3}.$$

**Proof:** See Appendix A.

This proposition is composed of two main parts. The first one is related to the construction of bases in the general case. The second part is a slight adaptation in the particular case where the Gegenbauer frequency is a power of  $1/2$ .

To obtain the best-ortho-basis of a 1-factor Gegenbauer process, we propose an algorithm built according to Proposition 3.1 and the algorithm defined in (12):

---

**Algorithm 1** 1-factor Best-Basis Search Algorithm

---

**Require:** A Gegenbauer frequency  $\nu$  and sample size  $N = 2^J$ ,

**Initialization**

```

1: for  $j = 0, \dots, J$  and  $p = 0, \dots, 2^j - 1$  do
2:    $Tree(j, p) = 0$ .
3: end for

```

**Main Loop**

```

4: for  $j = 1, \dots, J$  do
5:   for  $p = 0, 2, \dots, 2^j - 2$  do
6:     if  $\nu \in [p/2^j, (p+1)/2^j]$  then
7:        $Tree(j, p+1) = 1$ 
8:     end if
9:     if  $\nu \in [(p+1)/2^j, (p+2)/2^j]$  then
10:       $Tree(j, p) = 1$ 
11:    end if
12:   end for
13: end for

```

**Pruning**

```

14: for  $j = 1, \dots, J$  do
15:   for  $p = 0, 2, \dots, 2^j - 2$  do
16:     if  $\nu \in [p/2^j, (p+1)/2^j]$  and  $\max_{r=1, \dots, J-j-1; s=0, \dots, 2^r-1} Tree(j+r, 2^r p + s) > 0$  then
17:        $Tree(j, p) = 0$ 
18:     end if
19:   end for
20: end for

```

---

This algorithm is decomposed into two main loops. The first one builds a family where the best-ortho-basis is included. The second loop is a pruning of the family to obtain the best-ortho-basis. This second loop corresponds to the second part of Proposition 3.1.

The algorithm we propose is very fast because it does not require the calculation of variances of wavelet packet coefficients. To illustrate the rapidity of our algorithm we provide in the next figure some computation time to build bases using our method and the Whitcher's method. In this example, we are only interested in the time needed to build the basis. These bases are built to simulate Gegenbauer process with a singularity located at  $1/12$  and length equal to  $2^J$ , with  $J = 5, \dots, 13$ . The black line corresponds to the computation time of the algorithm we propose. The signs plus, times, dots and squares correspond to the computation time to build basis using the Whitcher's method in the case of respectively Daubechies wavelet (10), Symmlet (10), Coiflet (5) and Battle-Lemarie wavelet (5). In every case the computation time increases with the length of the process. However this time increases always much faster for the Whitcher method than for the method we develop (the ratio of computation times is 10 to 300 times larger for Whitcher's method for series of length 32 to 8192). Typically, for a 8192-sample series, it takes 100 ms to our algorithm to find the best basis while Whitcher's algorithm requires 30 s.

**Examples of bases**

We give here two examples of construction of bases. Figure 2, represents the basis built to simulate a Gegenbauer process with Gegenbauer frequency  $\nu = 1/12$  and memory parameter  $0 < d < 1/2$ . It was

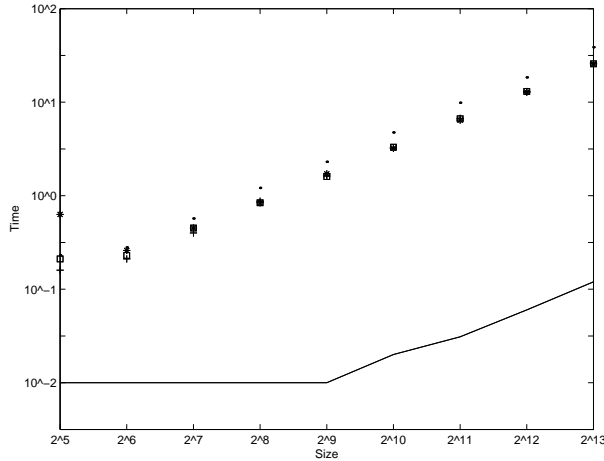


Figure 1: Computation time (in seconds) to build basis.

built using the first part of the previous algorithm.

Figure 3 represent the basis built in the case of a Gegenbauer process with  $\nu = 1/4$  and  $0 < d < 1/2$ . This second case corresponds to the second case of Proposition 3.1. One may remark that unlike the first case where the tree has at least one leaf at each scale, in this second case, because of the particular value of the Gegenbauer frequency ( $\nu = 2^{-2}$ ), there exists scale for which the tree has no leaf (see scale  $j = 2$ ).



Figure 2: Basis for a Gegenbauer process, with  $\nu = 1/12$ .

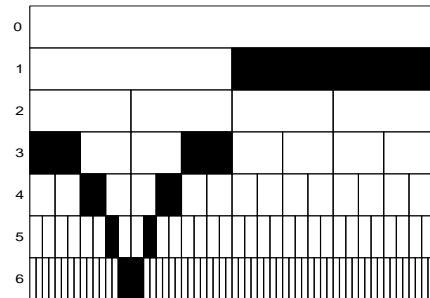


Figure 3: Basis for a Gegenbauer process, with  $\nu = 1/4$ .

### 3.4 The general case

In this section we are interested in the general case: the construction of the appropriate basis to simulate a  $k$ -factor Gegenbauer process. To achieve this goal, let's consider  $(X_t^1)_t$  and  $(X_t^2)_t$  as respectively a  $(k-1)$ -factor and a 1-factor Gegenbauer processes. We denote  $(d_1, \nu_1, \dots, d_{k-1}, \nu_{k-1})$  and  $(d_k, \nu_k)$  the parameters of  $(X_t^1)_t$  and  $(X_t^2)_t$ . Let  $\mathcal{B}_1$  and  $\mathcal{B}_2$  the best-ortho-bases of  $(X_t^1)_t$  and  $(X_t^2)_t$ . We denote respectively  $\mathcal{T}_1$  and  $\mathcal{T}_2$  the trees associated with the bases  $\mathcal{B}_1$  and  $\mathcal{B}_2$ .

Let  $(X_t)_t$  be a  $k$ -factor Gegenbauer process with parameters  $(d_1, \nu_1, \dots, d_k, \nu_k)$ . We denote  $\mathcal{B}$  the appropriate basis and  $\mathcal{T}$  the associated tree. Let  $\mathcal{B}'$  be the family equal to the union of the bases  $\mathcal{B}_1$  and  $\mathcal{B}_2$

and let  $\mathcal{T}'$  be the associated tree. We are now ready to state the following,

**Proposition 3.2** *Under the previous assumptions,*

1.  $\mathcal{B} \subset \mathcal{B}'$
2. Let  $(j, p)$  be node in the tree  $\mathcal{T}'$  such that there exists  $r^* = 1, \dots, J - j$  and  $s^* = 0, \dots, 2^{r^*} - 1$  such that  $(j + r^*, 2^{r^*}p + s^*)$  is also in the tree  $\mathcal{T}'$ . Then,

$$(j, p) \notin \mathcal{T} \quad \text{and} \quad (j + r^*, 2^{r^*}p + s) \in \mathcal{T}$$

**Proof:** See Appendix A.

According to this last proposition, the best-ortho-basis of a  $k$ -factor Gegenbauer process may be build using  $k$  well chosen best-ortho-bases of 1-factor Gegenbauer processes. The first part of this proposition defines a family where the appropriate basis is included and the second part is a pruning of this family to obtain the appropriate basis.

Now we propose an algorithm to build appropriate basis to simulate a  $k$ -factor Gegenbauer process. This algorithm lies on Algorithm 1 and results given in Proposition 3.2.

---

**Algorithm 2**  $k$ -factor Best-Basis Search Algorithm

---

**Require:** Gegenbauer frequencies  $\nu_i$  and sample size  $N = 2^J$ ,

**Initialization**

- 1: **for** Each Gegenbauer frequency  $\nu_i, i = 1, \dots, k$  **do**
- 2:   Construct the best-ortho-basis  $\mathcal{B}_i$  and associated tree  $Tree_i$  using Algorithm 1.
- 3: **end for**
- 4:  $Tree = \cup_{i=1}^k Tree_i$  (implemented using e.g. the logical OR operator under R or Matlab).

**Pruning**

- 5: **for**  $j = 1, \dots, J$  **do**
  - 6:   **for**  $p = 0, 2, \dots, 2^j - 2$  **do**
  - 7:     **if**  $Tree(j, p) = 1$  and  $\max_{r=1, \dots, J-j-1; s=0, \dots, 2^r-1} Tree(j + r, 2^r p + s) > 0$  **then**
  - 8:        $Tree(j, p) = 0$
  - 9:     **end if**
  - 10:   **end for**
  - 11: **end for**
- 

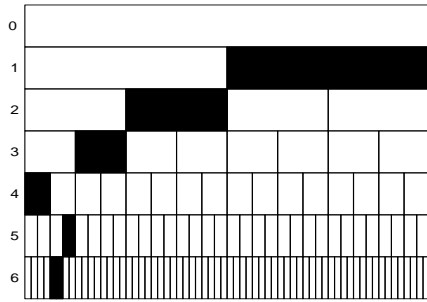
**Example** We here give an example of construction of the best-ortho-basis for a 2-factor Gegenbauer process  $(X_t)_t$  with Gegenbauer frequencies  $1/12$  and  $1/24$ . Figures 4.(a) and 4.(b) show the best-ortho-bases  $\mathcal{B}_1$  and  $\mathcal{B}_2$  of the processes  $(X_t^1)_t$  and  $(X_t^2)_t$  (see the previous Section for construction of these bases).

The family  $\mathcal{B}^*$  equal to  $\mathcal{B}_1 \cup \mathcal{B}_2$  is given Figure 4.(c). This family is not a basis, the intersections between its elements are not always empty. At depth  $j = 3$ , the elements at  $p = 0$  and  $p = 1$  are considered as elements of the best-ortho-basis which create some problem. Using the methodology developed above, we obtain an appropriate basis for the process  $(X_t)_t$ . This basis is represented Figure 4.(d).

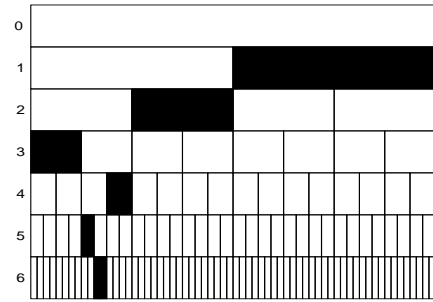
## 4 Analysis of decorrelation properties

One of the approximations done to simulate the Gegenbauer processes using wavelet is that the variance of the wavelet packet coefficients is exactly equal to the integral of the spectral density of the process over the corresponding dyadic interval. This equality would imply the non correlation of wavelet packet coefficients. In reality this equality is not verified which implies some correlation between wavelet packet coefficients. It is very interesting to evaluate this correlation to measure the error made. We give in the next proposition the asymptotic behavior of the covariance between wavelet packet coefficients for a 1-factor Gegenbauer process.

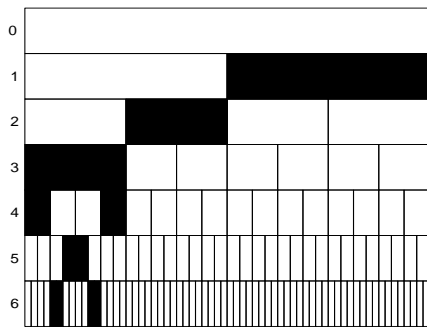




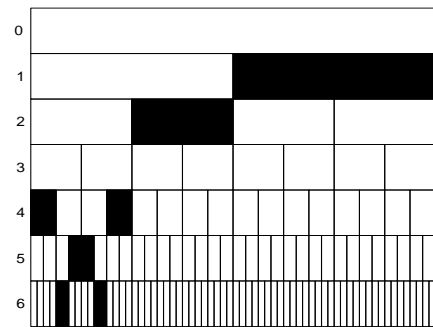
(a)



(b)



(c)



(d)

Figure 4: (a) Best-ortho-basis  $\mathcal{B}_1$ . (b) Best-ortho-basis  $\mathcal{B}_2$ . (c) Union of bases  $\mathcal{B}_1$  and  $\mathcal{B}_2$ . (d) Best-ortho-basis for the two factor Gegenbauer process  $X_t$ .

**Proposition 4.1** *If  $\psi$  has  $q \geq 1$  vanishing moments with support  $[(N_1 - N_2 + 1)/2, (N_2 - N_1 + 1)/2]$  and  $X(t)$  is a stationary 1-factor Gegenbauer process with Gegenbauer frequency  $\nu$ . Then the wavelet packet coefficients covariance  $\text{Cov}(W_{j_1}^{p_1}(k_1), W_{j_2}^{p_2}(k_2))$  decays as*

- $O(|2^{j_1}k_1 - 2^{j_2}k_2|^{2d-1-R_{p_1}-R_{p_2}})$ , if  $p_1 \neq 0$  and  $p_2 \neq 0$ ,
- $O(|2^{j_1}k_1 - 2^{j_2}k_2|^{2d-1-R_{\max(p_1, p_2)}})$ , if  $p_1 = 0$  or  $p_2 = 0$ ,
- $O(|2^{j_1}k_1 - 2^{j_2}k_2|^{2d-1})$ , if  $p_1 = p_2 = 0$ ,

for all,  $j_1, j_2, k_1$  and  $k_2$  such that  $|2^{j_1}k_1 - 2^{j_2}k_2| > (N^* + 1)(2^{j_1} + 2^{j_2})$ , with  $N^* = \max(N_1, N_2)$ , and  $R_p = q \sum_{k=0}^{j-1} p^k$  for  $p \neq 0$ , and  $p = (p^{j-1}p^{j-2} \dots p^1 p^0)_2$  is the binary representation of  $p$ . In the last case, we note that  $j_1 = j_2 = j$ .

**Proof:** See Appendix B.

This proposition generalized the results given by [Jen99] and [Jen00] for the case of the FARIMA process. Clearly the covariance between  $W_{j_1}^{p_1}(k_1)$  and  $W_{j_2}^{p_2}(k_2)$  decay exponentially over time and scale space. More precisely, the decay speed for  $p_1 \neq 0$  or  $p_2 \neq 0$ , depends on the regularity of the used wavelet, on the memory parameter of the process, and indirectly on the location of singularity through the frequency indices  $p_1$  and  $p_2$ . However, the larger  $q$  is, the wider the wavelet's support and the fewer are the number of wavelet packet coefficients that satisfy the support condition  $|2^{j_1}k_1 - 2^{j_2}k_2| > (N^* + 1)(2^{j_1} + 2^{j_2})$ . Thus, by choosing a wavelet with a large  $q$ , the rate of decay of autocovariance function increases, but over a subset of wavelet packet coefficients. However, the effective support of a wavelet is smaller than the provided bound (see Lemma 6.2), and we expect a rapid decay in the wavelet packet coefficient's covariance for translations and dilations satisfying  $|2^{j_1}k_1 - 2^{j_2}k_2| > (N^* + 1)(2^{j_1} + 2^{j_2})$ . The following simulation study confirms this remark.

## 5 Simulation results and discussion

### 5.1 Some simulations of Gegenbauer processes

This section is devoted to the illustration of some simulation examples of Gegenbauer processes. For comparative purposes, two methods are tested: ours and the one proposed in [Whi01]. Gegenbauer processes are synthesised with different frequencies and long memory parameters  $(d, \nu)$ :  $(0.4, 1/12)$ ,  $(0.2, 1/12)$ ,  $(0.3, 0.352)$  and  $(0.3, 0.016)$ . In each case, we consider the Daubechies wavelet 'db8' and we simulate  $M = 500$  time series with length  $N = 256$ . These different time series allow us to compute, for each process, the empirical mean of the autocovariance function  $\bar{\gamma}(\cdot)$  and its standard deviation  $\sigma_{\bar{\gamma}}^2(\cdot)$ , defined respectively by:

$$\bar{\gamma}(h) = \frac{1}{M} \sum_{i=1}^M \hat{\gamma}_i(h) \quad \text{and} \quad \sigma_{\bar{\gamma}}^2(h) = \frac{1}{M} \sum_{i=1}^M (\hat{\gamma}_i(h) - \bar{\gamma}(h))^2,$$

for the 64 first lags  $h = 1, \dots, 64$  and where  $\hat{\gamma}(\cdot)$  is the empirical autocovariance function of the  $i^{\text{th}}$  simulated time series. The results are depicted in Figure 5.

As revealed by the plots of the average autocovariance function Figure 5.(b), the quality of estimation of the autocovariance function obtained using our best-basis construction method and that of Whitcher are very close to each other. It appears that the simulations made using both our bases and those built with Whitcher's method tend to underestimate the autocovariance function. This underestimation seems to be more important with our bases but the difference between estimations for both methods are very small. This could be interpreted as a better performance of Whitcher's method, but the comparison in this context is not fair. Indeed, an important drawback of Whitcher's method is that it overpartitions the spectrum especially at fine scales leading to less correlation between the wavelet packet coefficients. Additionally, the standard deviation of the estimated autocovariance functions is smaller for simulations based on our bases than for those based Whitcher's best-bases. This implies that the synthesis method using our bases provide more stable estimation of the autocovariance function which is a desirable property (see in Figure 5.(b) for each different couple of parameters). To conclude we say that the small loss of quality of the estimation of autocovariance function when we use our method of construction of basis

is compensated by a more important stability of the estimations. Another clear advantage of our best-ortho-basis construction method will be within the framework of estimation and bootstrapping where overpartitioning of the spectrum will have disastrous consequences. This is the direction of our current investigations.

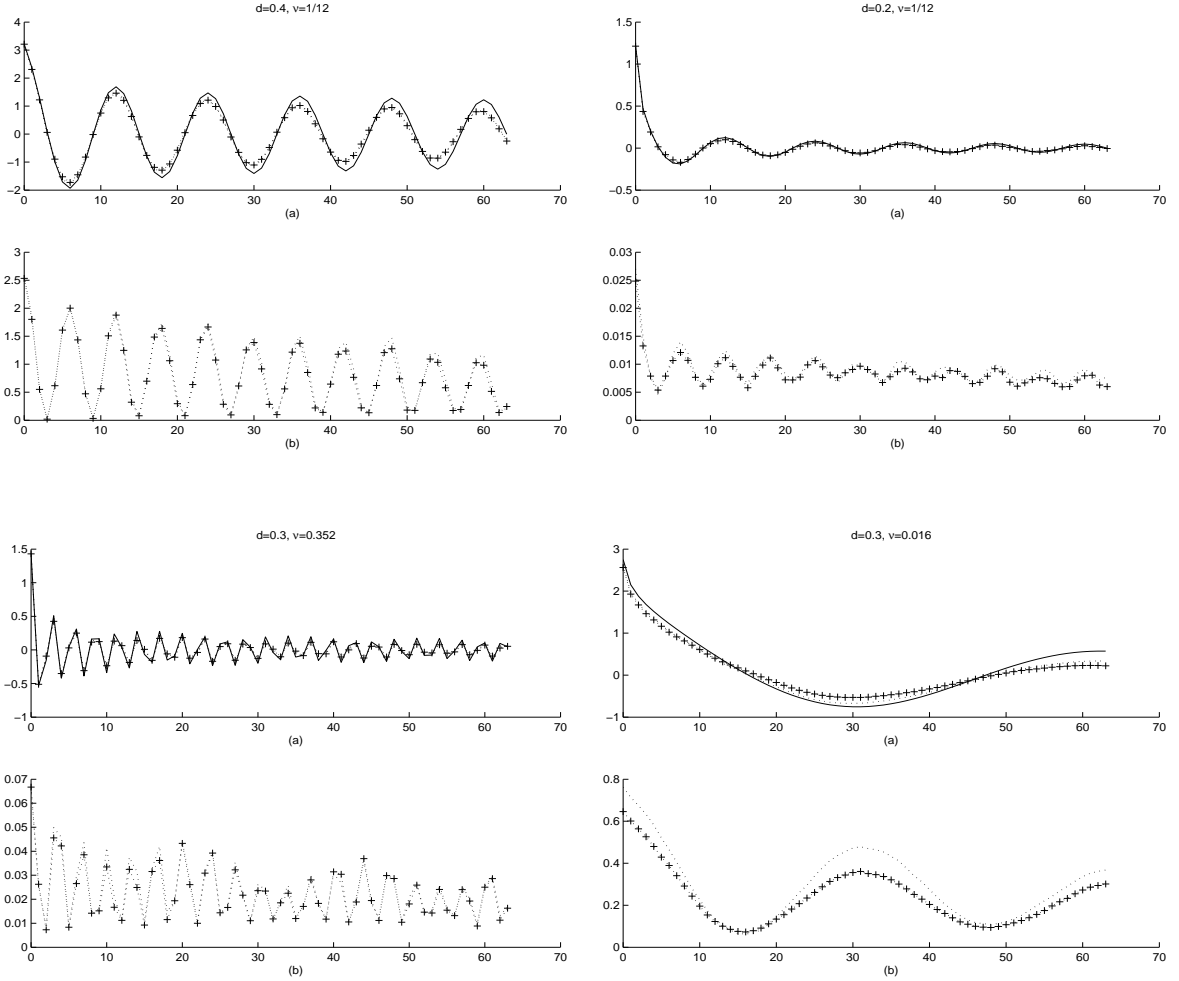


Figure 5: (a) Empirical mean of the autocovariance function for lags  $1, \dots, 64$  for Gegenbauer processes built using Whitcher's method (dotted line) and using the method we develop (dotted line with sign '+'). (b) Standard deviation of these estimations

## 5.2 Empirical analysis of the decorrelation

We now propose some simulations which allow to compare the decorrelation of the wavelet packet coefficients as a function of the wavelet properties (type, number of vanishing moments) and the process parameters ( $\nu$  and  $d$ ). The approximate diagonalizing capabilities of our algorithm are compared to those of [Whi01]. We then need to consider a criterion to measure the quality of non-correlation, such as the HS norm considered in Section 3.1, which measures the sum of squares of the off-diagonal elements of the covariance matrix in the best-ortho-basis. As explained above, the method of [Whi01] tends to overpartition the spectral axis yielding to too many packets. Hence, to penalize such configurations and make the comparison fair, we propose the following penalized criterion,

$$S(\mathcal{B}) = \|\Gamma[\mathcal{B}] - \Gamma_0\|_{HS}^2 + \lambda \#(\mathcal{B}) \quad (13)$$

where  $\Gamma[\mathcal{B}]$  is the correlation matrix of the process expressed in the basis  $\mathcal{B}$ ,  $\Gamma_0$  is the correlation matrix of a white noise, i.e. the identity matrix and  $\lambda$  is a weight parameter that we discuss latter.  $\#(\mathcal{B})$  stands for the number of wavelet packets (spectral axis partitions) in the basis. In other words, the basis  $\mathcal{B}$

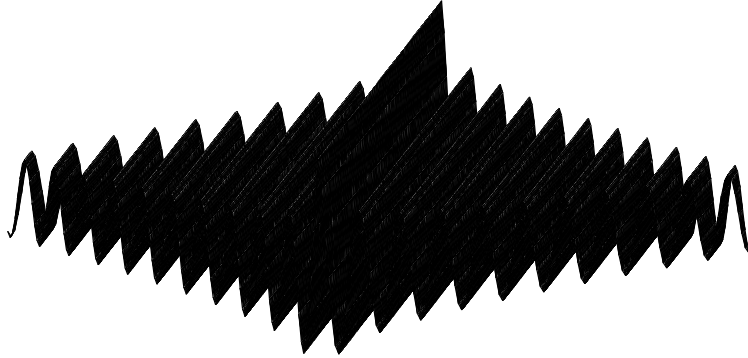


Figure 6: Correlation matrix of a Gegenbauer process with parameters  $d = 0.4$  and  $\nu = 1/12$ .

that we use to simulate the Gegenbauer process is optimal, in the sense that this basis is the one for which the norm  $\|\Gamma[\mathcal{B}] - \Gamma_0\|_{HS}^2$  is sufficiently good with the smallest number of partitions to penalize overpartitioned bases.

To determine the value of the constant  $\lambda$ , assume that we want to study the decorrelation of the  $2^J \times 2^J$  covariance matrix of a Gegenbauer process. We consider two extreme cases. In the one hand let  $\mathcal{B}_S$  be the Shannon basis. We can assume that the decorrelation of the covariance matrix of the Gegenbauer process in this basis is perfect; thus  $S(\mathcal{B}_S) = 2^J \lambda$ . The tree associated to this basis has too many leaves and then this kind of basis is not interesting. On the other hand we consider the basis  $\mathcal{B}_0$  composed with only one leaf. In this case there isn't any decorrelation of the covariance matrix, that is  $S(\mathcal{B}_0) = \|\Gamma - \Gamma_0\|_{HS}^2$ , with  $\Gamma$  the correlation matrix of the Gegenbauer process.

We consider that for this two extreme cases the statistic  $S(\cdot)$  has the same value, then

$$\begin{aligned} S(\mathcal{B}_0) &= S(\mathcal{B}_S) \\ \Leftrightarrow \|\Gamma - \Gamma_0\|_{HS}^2 + \lambda &= 2^J \lambda \\ \Leftrightarrow \lambda &= \frac{\|\Gamma - \Gamma_0\|_{HS}^2}{2^J - 1}. \end{aligned}$$

We compute the value of the statistic  $S(\cdot)$ , for different covariance matrix of long memory processes (see [And86] for details about the calculation of the exact autocovariance function of Gegenbauer processes) and different wavelets. To assess the quality of our best-ortho-basis, we systematically compare it to the ones obtained by the method of [Whi01]. For each process and each wavelet we give both the value of the criterion  $S$  and the value of the constant  $\lambda$ .

The values of the criterion  $S$  for each process and each wavelet are given in Table 1. The basis we propose gives at least better results than the basis given by Whitcher. One can remark that for a given process when the number of vanishing moments increase the criteria tend to be the same. Actually when the number of vanishing moments of a wavelet increases the basis proposed by Whitcher tends towards the basis we propose.

## 6 Conclusion

In this article, we provided a new method to build approximate diagonalizing bases for  $k$ -factor Gegenbauer processes. Exploiting the intuitive fact that a wavelet packets library contains the basis where a Gegenbauer process could be (almost) whitened, our best-ortho-basis search algorithm was formulated in the case of 1-factor process and the fast search algorithm of Coifman-Wickerhauser was adapted to find this best basis. Using this framework, our methodology was posed in a well principled way and the uniqueness of the basis was guaranteed. Furthermore, unlike Whitcher's approach, it is very fast (see

$d = 0.4, \nu = 1/12, \lambda = 20.7084$			$d = 0.2, \nu = 1/12, \lambda = 0.7428$			$d = 0.3, \nu = 0.352, \lambda = 2.8274$		
$q$	Whitcher basis	Our basis	$q$	Whitcher basis	Our basis	$q$	Whitcher basis	Our basis
Daubechies			Daubechies			Daubechies		
2	2728.6	1494.5	2	105.1	52.3	2	356.3	297.2
4	1116.7	686.2	4	47.3	31.1	4	169.1	158.2
6	750.4	441.8	6	31.9	23.3	6	128.5	109.1
8	632.7	352.4	8	28.9	20.2	8	101.2	87.5
10	421.7	308.2	10	21.9	18.4	10	96.8	76.5
Symmlet			Symmlet			Symmlet		
4	2211.0	677.0	4	86.4	30.8	4	288.8	157.1
5	1713.1	533.8	5	67.2	26.1	5	236.4	128.2
6	1371.0	444.7	6	54.2	23.4	6	191.9	109.4
7	1182.4	385.8	7	46.8	21.7	7	155.8	97.6
8	980.4	341.7	8	39.1	20.1	8	131.7	87.8
9	807.7	313.8	9	32.4	18.7	9	128.7	80.8
10	693.1	297.3	10	28.2	18.2	10	112.8	75.9
Coiflet			Coiflet			Coiflet		
2	2697.3	1081.1	2	106.0	42.3	2	362.2	231.3
4	1449.4	638.4	4	58.9	29.6	4	210.9	149.4
6	841.1	412.7	6	35.0	22.4	6	139.2	103.7
8	703.2	327.8	8	29.0	19.5	8	116.2	83.9
10	581.2	287.7	10	26.6	17.6	10	111.2	72.2
Battle-Lemarié			Battle-Lemarié			Battle-Lemarié		
2	3415.1	657.3	2	127.2	29.4	2	442.0	148.1
4	1303.6	267.5	4	49.0	16.1	4	172.3	59.7
6	667.9	247.8	6	25.9	14.4	6	101.1	48.2

$d = 0.3, \nu = 0.016, \lambda = 10.0526$			ARFIMA $d = 0.4, \lambda = 24.8313$		
$q$	Whitcher basis	Our basis	$q$	Whitcher basis	Our basis
Daubechies			Daubechies		
2	230.6	141.4	2	259.1	234.3
4	188.0	124.7	4	231.9	231.9
6	141.8	116.7	6	230.3	230.3
8	144.0	118.8	8	231.6	231.6
10	140.3	115.0	10	229.4	229.4
Symmlet			Symmlet		
4	214.4	120.9	4	255.4	230.9
5	212.7	119.1	5	231.3	231.3
6	210.4	116.2	6	230.1	230.1
7	211.4	117.4	7	230.1	230.1
8	178.6	114.7	8	230.1	230.1
9	138.7	113.4	9	229.0	229.0
10	138.8	113.5	10	229.6	229.6
Coiflet			Coiflet		
2	224.8	132.7	2	257.1	232.4
4	215.2	121.0	4	231.5	231.5
6	141.2	116.1	6	230.3	230.3
8	140.1	114.9	8	230.1	230.1
10	138.6	113.2	10	229.5	229.5
Battle-Lemarié			Battle-Lemarié		
2	294.1	121.3	2	304.7	230.7
4	206.8	111.8	4	229.1	229.1
6	136.0	110.7	6	228.9	228.9

Table 1:  $S$  score as a function of the number of vanishing moments for each wavelet family.

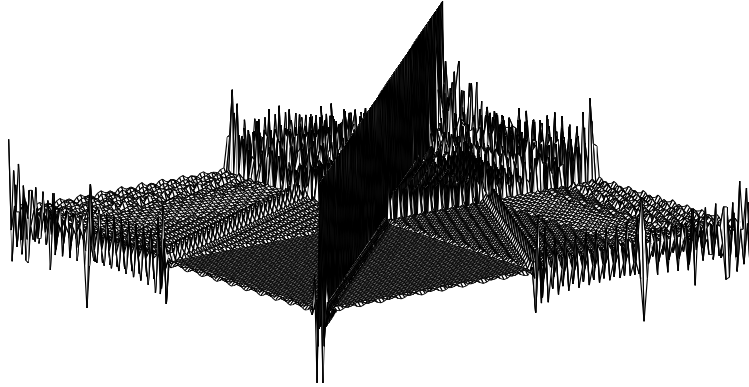


Figure 7: Correlation matrix  $\Gamma[\mathcal{B}]$  of the wavelet packet coefficients in the best-ortho-basis obtained from a Gegenbauer process with parameters  $d = 0.4$  and  $\nu = 1/12$  using '*db10*' filter.

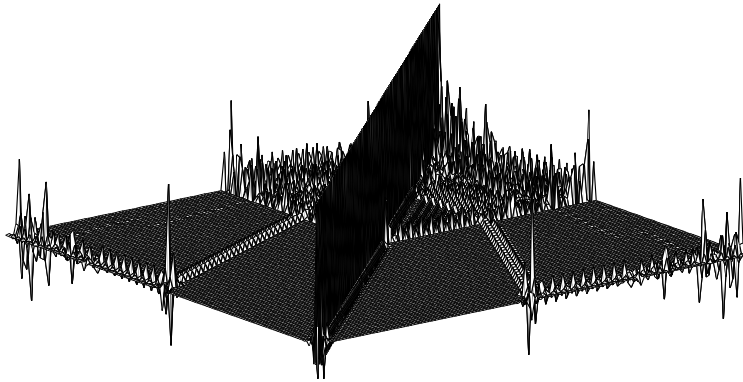


Figure 8: Correlation matrix  $\Gamma[\mathcal{B}]$  of the wavelet packet coefficients in the best-ortho-basis obtained from a Gegenbauer process with parameters  $d = 0.4$  and  $\nu = 1/12$  using '*sym10*' filter.

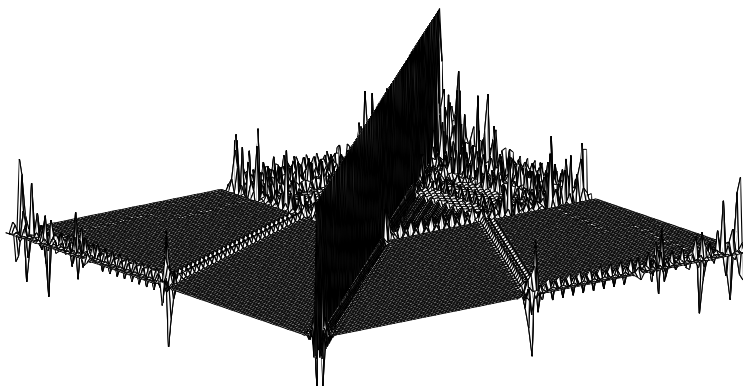


Figure 9: Correlation matrix  $\Gamma[\mathcal{B}]$  of the wavelet packet coefficients in the best-ortho-basis obtained from a Gegenbauer process with parameters  $d = 0.4$  and  $\nu = 1/12$  using '*coif5*' filter.

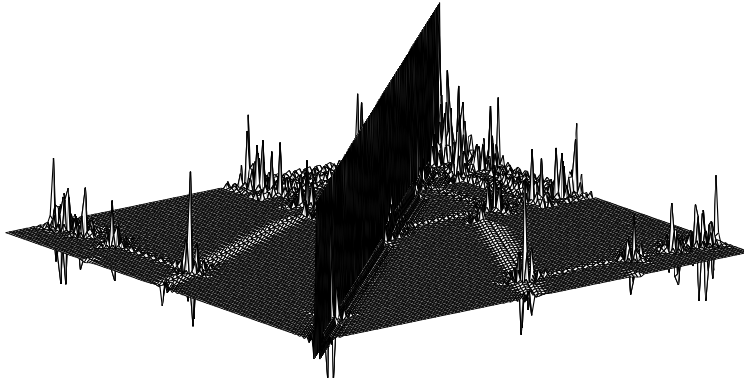


Figure 10: Correlation matrix  $\Gamma[\mathcal{B}]$  of the wavelet packet coefficients in the best-ortho-basis obtained from a Gegenbauer process with parameters  $d = 0.4$  and  $\nu = 1/12$  using 'bat6' filter.

simulations), does not depend on the wavelet choice, and is not very sensitive to the length of the time series. As the method construction of the best basis for simulation of  $k$ -factor Gegenbauer processes relies on the 1-factor construction method, the same conclusions hold.

Then, we studied the error of diagonalization in the best-ortho-basis. Towards this goal, we established the decay speed of the correlation between two wavelet packet coefficients. These results generalize the work of [Jen99] and [Jen00] provided in the case of FARIMA processes. A small empirical study shows how we improve, in the sense of the chosen criterion, the decorrelation of the process when it is decomposed in the provided basis. The comparison criterion we propose measures both the quality of the decorrelation and penalizes over-partitioned configurations. Owing to these nice theoretical and empirical properties, and given the simplicity of programming and running, we feel the general practitioner will be attracted to our simulator.

This new method of simulating Gegenbauer processes gives a new perspective for analyzing processes whose spectral density singularities occurs at any frequency in the Nyquist interval. In such analysis task, we have the basis by knowing the process parameters ( $\nu$  in particular). Our method has a direct application for bootstrap-based inference in the presence of Gegenbauer noise.

A remaining important open problem is how could we extend this work if the question of interest becomes: what are the parameters of a  $k$ -factor Gegenbauer given one or more sample paths of this process. This is an estimation problem that can be accomplished in a maximum likelihood framework once the diagonalizing basis is found. In this case, the best-ortho-basis cannot be found by a naive straightforward application of Algorithm 2. Nevertheless, we here give pointers to several promising directions that are now under investigation. First, in the 1-factor case, one can exploit the result of Proposition 3.1 by comparing the variance of the wavelet packet coefficients and only partition the packet whose variance is the largest. This first approach is simple but is not easy to generalize to a  $k$ -factor process. Alternatively, if we are given several time series of the Gegenbauer process, one could use a best-ortho-basis search procedure in the same vein as that proposed in [MZP98] and [DMvS98] for locally stationary processes. Additional research is still required and our current work is now directed towards extension of our methodology to estimation.

## Appendix A

### Proof of Proposition 3.1:

- We consider the node  $(j, p)$ . We compute the variance of the wavelet packet coefficients at the two children of this node:  $(j+1, 2p)$  and  $(j+1, 2p+1)$ . We assume that the frequency  $\nu$  is in the interval  $I_{j+1}^{2p} = [\frac{2p}{2^{j+1}}, \frac{2p+1}{2^{j+1}}]$ . Then a good approximation of the variance of the wavelet packet coefficient is given by the integral over the interval  $I_{j+1}^{2p}$  of the spectral density. On this interval we approximate the spectral density of the process  $f_X(\lambda) = \frac{\sigma^2}{2\pi} |2(\cos 2\pi\lambda - \cos 2\pi\nu)|^{-2d}$  by  $C_0|\lambda - \nu|^{-2d}$  with  $C_0$  a positive constant.

$$\begin{aligned}
\mathbb{V}[W_{j+1}^{2p}] &= \int_{\frac{2p}{2^{j+1}}}^{\frac{2p+1}{2^{j+1}}} \frac{\sigma^2}{2\pi} |2(\cos 2\pi\lambda - \cos 2\pi\nu)|^{-2d} d\lambda \\
&= \int_{\frac{2p}{2^{j+1}}}^{\frac{2p+1}{2^{j+1}}} C_0 |\lambda - \nu|^{-2d} d\lambda \\
&= C_0 \int_{\frac{2p}{2^{j+1}}}^{\nu} (\nu - \lambda)^{-2d} d\lambda + C_0 \int_{\nu}^{\frac{2p+1}{2^{j+1}}} (\lambda - \nu)^{-2d} d\lambda \\
&= \frac{C_0}{1-2d} \left( \left( \nu - \frac{2p}{2^{j+1}} \right)^{1-2d} + \left( \frac{2p+1}{2^{j+1}} - \nu \right)^{1-2d} \right) \\
&= \frac{C_0}{1-2d} \left( \left( \frac{2p+1}{2^{j+1}} - \nu - \frac{1}{2^{j+1}} \right)^{1-2d} + \left( \frac{2p+1}{2^{j+1}} - \nu \right)^{1-2d} \right) \\
&= C_0 \frac{\left( \frac{2p+1}{2^{j+1}} - \nu \right)^{1-2d}}{1-2d} \left( 1 + \left( 1 - \frac{1}{2^{j+1} \left( \frac{2p+1}{2^{j+1}} - \nu \right)} \right)^{1-2d} \right) \\
&= C_0 \frac{\left( \frac{2p+1}{2^{j+1}} - \nu \right)^{1-2d}}{1-2d} \left( 2 - \frac{1-2d}{2^{j+1} \left( \frac{2p+1}{2^{j+1}} - \nu \right)} + O\left( \frac{1}{\left( 2^{j+1} \left( \frac{2p+1}{2^{j+1}} - \nu \right) \right)^2} \right) \right) \\
&\sim C_0 \frac{\left( \frac{2p+1}{2^{j+1}} - \nu \right)^{1-2d}}{1-2d} \left( 2 - \frac{1-2d}{2^{j+1} \left( \frac{2p+1}{2^{j+1}} - \nu \right)} \right) \\
&= C_0 \frac{\left( \frac{2p+1}{2^{j+1}} - \nu \right)^{1-2d}}{(1-2d)2^{j+1} \left( \frac{2p+1}{2^{j+1}} - \nu \right)} \left( 2^{j+2} \left( \frac{2p+1}{2^{j+1}} - \nu \right) - 1 + 2d \right) \\
&\sim \frac{C_0}{2^{j+1}} \left( \frac{2p+1}{2^{j+1}} - \nu \right)^{-2d} \frac{1+2d}{1-2d},
\end{aligned}$$

To compute the variance of  $W_{j+1}^{2p+1}$  we denote  $\lambda^*$  the location of the maxima of the spectral density  $f_X$  over the interval  $I_{j+1}^{2p+1}$ .

$$\begin{aligned}
\mathbb{V}[W_{j+1}^{2p+1}] &= \int_{\frac{2p+1}{2^j}}^{\frac{2p+2}{2^{j+1}}} \frac{\sigma^2}{2\pi} |2(\cos 2\pi\lambda - \cos 2\pi\nu)|^{-2d} d\lambda \\
&\leq \frac{\sigma^2}{2\pi 2^{j+1}} |2(\cos 2\pi\lambda^* - \cos 2\pi\nu)|^{-2d}.
\end{aligned}$$

As previously we use an approximation of the spectral density. As  $\lambda^* = \frac{2p+1}{2^{j+1}}$ , we obtain that

$$\mathbb{V}[W_{j+1}^{2p+1}] \leq \frac{C_0}{2^{j+1}} \left( \frac{2p+1}{2^{j+1}} - \nu \right)^{-2d}.$$



Then we have that

$$\mathbb{V}[W_{j+1}^{2p+1}] \leq \frac{1+2d}{1-2d} \mathbb{V}[W_{j+1}^{2p}].$$

In this case we write that

$$\mathbb{V}[W_{j+1}^{2p+1}] \ll \mathbb{V}[W_{j+1}^{2p}],$$

and using the criterion defined in section 3.1 we obtain

$$\overline{\mathbb{V}[W_{j+1}^{2p+1}]} = 0.$$

Finally, at the node  $(j, p)$ , the algorithm (12) give us that,

$$\mathcal{B}_j^p = \mathcal{B}_{j+1}^{2p} \cup \mathcal{B}_{j+1}^{2p+1}.$$

- In the case where the frequency  $\nu$  is in the closure of the intervals  $I_{j+1}^{2p}$  and  $I_{j+1}^{2p+1}$ , we have not relationship as

$$\overline{\mathbb{V}[W_{j+1}^{2p}]} = 0 \quad \text{or} \quad \overline{\mathbb{V}[W_{j+1}^{2p+1}]} = 0,$$

and so it is not possible to conclude. However, at the depth  $j+2$ , we have

$$\overline{\mathbb{V}[W_{j+2}^{4p+1}]} = 0 \quad \text{and} \quad \overline{\mathbb{V}[W_{j+2}^{4p+2}]} = 0.$$

Then easily we obtain that for the algorithm (12),

$$\mathcal{B}_j^p = \mathcal{B}_{j+2}^{4p} \cup \mathcal{B}_{j+2}^{4p+1} \cup \mathcal{B}_{j+2}^{4p+2} \cup \mathcal{B}_{j+2}^{4p+3}.$$

### Proof of Proposition 3.2:

1. Let  $(j, p)$  be a node. We assume that this node is not in the tree  $\mathcal{T}'$ . It means that this node is not in the tree  $\mathcal{T}_1$  and neither in the tree  $\mathcal{T}_2$  and in terms of threshold, we have

$$\begin{aligned} \overline{\mathbb{V}[W_{j+1}^{2p}(1)]} = 0 & \quad \text{or} \quad \overline{\mathbb{V}[W_{j+1}^{2p+1}(1)]} = 0 \\ & \mathbf{and} \\ \overline{\mathbb{V}[W_{j+1}^{2p}(2)]} = 0 & \quad \text{or} \quad \overline{\mathbb{V}[W_{j+1}^{2p+1}(2)]} = 0. \end{aligned}$$

As the tree  $\mathcal{T}_2$  is associated to the best-ortho-basis of a 1-factor Gegenbauer process, the fact that the node  $(j, p)$  is node in the tree  $\mathcal{T}_2$  means that the frequency  $\nu_k$  is not in the interval  $I_j^p = [\frac{p}{2^j}, \frac{p+1}{2^j}]$ . Then in this interval, the function  $|2(\cos 2\pi\lambda - \cos 2\pi\nu_k)|^{-2d_k}$  is bounded and has a maximum at frequency  $\lambda^* \in I_j^p$ .

Then,

$$\begin{aligned} \mathbb{V}[W_{j+1}^{2p}] &= \frac{\sigma^2}{2\pi} \int_{\frac{2p}{2^{j+1}}}^{\frac{2p+1}{2^{j+1}}} |2(\cos 2\pi\lambda - \cos 2\pi\nu_k)|^{-2d_k} \prod_{i=1}^{k-1} |2(\cos 2\pi\lambda - \cos 2\pi\nu_i)|^{-2d_i} d\lambda \\ &\leq \frac{\sigma^2}{2\pi} |2(\cos 2\pi\lambda^* - \cos 2\pi\nu_k)|^{-2d_k} \int_{\frac{2p}{2^{j+1}}}^{\frac{2p+1}{2^{j+1}}} \prod_{i=1}^{k-1} |2(\cos 2\pi\lambda - \cos 2\pi\nu_i)|^{-2d_i} d\lambda \\ &= \frac{\sigma^2}{2\pi} |2(\cos 2\pi\lambda^* - \cos 2\pi\nu_k)|^{-2d_k} \mathbb{V}[W_{j+1}^{2p}(1)]. \end{aligned}$$

and

$$\begin{aligned} \mathbb{V}[W_{j+1}^{2p+1}] &= \frac{\sigma^2}{2\pi} \int_{\frac{2p}{2^{j+1}}}^{\frac{2p+2}{2^{j+1}}} |2(\cos 2\pi\lambda - \cos 2\pi\nu_k)|^{-2d_k} \prod_{i=1}^{k-1} |2(\cos 2\pi\lambda - \cos 2\pi\nu_i)|^{-2d_i} d\lambda \\ &\leq \frac{\sigma^2}{2\pi} |2(\cos 2\pi\lambda^* - \cos 2\pi\nu_k)|^{-2d_k} \int_{\frac{2p}{2^{j+1}}}^{\frac{2p+2}{2^{j+1}}} \prod_{i=1}^{k-1} |2(\cos 2\pi\lambda - \cos 2\pi\nu_i)|^{-2d_i} d\lambda \\ &= \frac{\sigma^2}{2\pi} |2(\cos 2\pi\lambda^* - \cos 2\pi\nu_k)|^{-2d_k} \mathbb{V}[W_{j+1}^{2p+1}(1)]. \end{aligned}$$

Finally, as  $\overline{\mathbb{V}[W_{j+1}^{2p}(1)]} = 0$  or  $\overline{\mathbb{V}[W_{j+1}^{2p+1}(1)]} = 0$ , we have

$$\mathbb{V}[W_{j+1}^{2p}] \gg \mathbb{V}[W_{j+1}^{2p+1}] \quad \text{or} \quad \mathbb{V}[W_{j+1}^{2p}] \gg \mathbb{V}[W_{j+1}^{2r p+1}],$$

it means,

$$\overline{\mathbb{V}[W_{j+1}^{2p}]} = 0 \quad \text{or} \quad \overline{\mathbb{V}[W_{j+1}^{2p+1}]} = 0.$$

Then,

$$\mathcal{B}_j^p = \mathcal{B}_{j+1}^{2p} \cup \mathcal{B}_{j+1}^{2p+1}$$

and so the node  $(j, p)$  is not in the tree  $\mathcal{T}$ . Finally,

$$\mathcal{B} \subset \mathcal{B}'.$$

2. Here  $(j, p)$  and  $(j + r^*, 2^{r^*} p + s^*)$  (for  $s^* = 0, \dots, 2^{r^*} - 1$ ) are in the tree  $\mathcal{T}'$ . We denote  $r$  the minimum value of  $r^*$

$$r = \min_{r=1, \dots, J-j-1} r^*, \quad (14)$$

for which there exists a  $s$  ( $s = 0, \dots, 2^r - 1$ ) such that the node  $(j + r, 2^r p + s)$  is in the tree  $\mathcal{T}'$ .

Then the fact that the nodes  $(j, p)$  and  $(j + r, 2^r p + s)$  are in the tree  $\mathcal{T}'$  means that  $(j, p)$  is in  $\mathcal{T}_1$  or in  $\mathcal{T}_2$  and  $(j + r, 2^r p + s)$  is in  $\mathcal{T}_2$  or in  $\mathcal{T}_1$  (it is important to remark that both  $(j, p)$  and  $(j + r, 2^r p + s)$  cannot be in  $\mathcal{T}_1$  or in  $\mathcal{T}_2$ ). Without loss of generality, we assume that  $(j, p)$  is in  $\mathcal{T}_2$  and  $(j + r, 2^r p + s)$  is in  $\mathcal{T}_1$ . All the calculations made in the following remain valid if we consider that  $(j, p)$  is in  $\mathcal{T}_1$  and  $(j + r, 2^r p + s)$  is in  $\mathcal{T}_2$ . To simplify the notations, we assume also that there exists a  $s$  which is even.

We denote  $W_j^p(1)$  and  $W_j^p(2)$ , for  $j = 0, \dots, J$  and  $p = 0, 2^j - 1$ , the wavelet packet coefficients of respectively the processes  $(X_t^1)_t$  and  $(X_t^2)_t$ . From these sub-processes, we have that for the algorithm CW,

- for the tree  $\mathcal{T}_1$ :

$$\mathcal{B}_{j+r-1}^{2^{r-1}p+\frac{s}{2}}(1) = \mathcal{B}_{j+r}^{2^r p+s}(1) \cup \mathcal{B}_{j+r}^{2^r p+s+1}(1)$$

because  $\overline{\mathbb{V}[W_{j+r}^{2^r p+s}(1)]} = 0$  or  $\overline{\mathbb{V}[W_{j+r}^{2^r p+s+1}(1)]} = 0$ ,

- for the tree  $\mathcal{T}_2$ :

$$\mathcal{B}_j^p(2) = \mathcal{B}_j^p(2)$$

because  $\overline{\mathbb{V}[W_{j+1}^{2p}(2)]} = \mathbb{V}[W_{j+1}^{2p}(2)]$  and  $\overline{\mathbb{V}[W_{j+1}^{2p+1}(2)]} = \mathbb{V}[W_{j+1}^{2p+1}(2)]$ ,

We consider the intervals  $I_{j+r}^{2^r p+s} = [\frac{2^r p+s}{2^{j+r}}, \frac{2^r p+s+1}{2^{j+r}}]$  and  $I_{j+r}^{2^r p+s+1} = [\frac{2^r p+s+1}{2^{j+r}}, \frac{2^r p+s+2}{2^{j+r}}]$ . As the node  $(j, p)$  is in the tree  $\mathcal{T}_2$ , and as  $\mathcal{T}_2$  the tree of a basis, the frequency  $\nu_k$  is not in the interval  $I_{j+r}^{2^r p+s} \cup I_{j+r}^{2^r p+s+1}$ .

We denote  $\lambda^*$  the location of the maximum of  $|2(\cos 2\pi\lambda - \cos 2\pi\nu_k)|^{-2d_k}$  in the interval  $I_{j+r-1}^{2^{r-1}p+s/2}$  (Note that the maximum is bounded). We have for the process  $(X_t)_t$ ,

$$\begin{aligned} \mathbb{V}[W_{j+r}^{2^r p+s}] &= \frac{\sigma^2}{2\pi} \int_{\frac{2^r p+s}{2^{j+r}}}^{\frac{2^r p+s+1}{2^{j+r}}} |2(\cos 2\pi\lambda - \cos 2\pi\nu_k)|^{-2d_k} \prod_{i=1}^{k-1} |2(\cos 2\pi\lambda - \cos 2\pi\nu_i)|^{-2d_i} d\lambda \\ &\leq \frac{\sigma^2 |2(\cos 2\pi\lambda^* - \cos 2\pi\nu_k)|^{-2d_k}}{2\pi} \int_{\frac{2^r p+s}{2^{j+r}}}^{\frac{2^r p+s+1}{2^{j+r}}} \prod_{i=1}^{k-1} |2(\cos 2\pi\lambda - \cos 2\pi\nu_i)|^{-2d_i} d\lambda \\ &= \frac{\sigma^2 |2(\cos 2\pi\lambda^* - \cos 2\pi\nu_k)|^{-2d_k}}{2\pi} \mathbb{V}[W_{j+r}^{2^r p+s}(1)], \end{aligned}$$

$$\begin{aligned}
\mathbb{V}[W_{j+r}^{2^r p+s+1}] &= \frac{\sigma^2}{2\pi} \int_{\frac{2^r p+s+1}{2^{j+r}}}^{\frac{2^r p+s+2}{2^{j+r}}} |2(\cos 2\pi\lambda - \cos 2\pi\nu_k)|^{-2d_k} \prod_{i=1}^{k-1} |2(\cos 2\pi\lambda - \cos 2\pi\nu_i)|^{-2d_i} d\lambda \\
&\leq \frac{\sigma^2 |2(\cos 2\pi\lambda^* - \cos 2\pi\nu_k)|^{-2d_k}}{2\pi} \int_{\frac{2^r p+s+1}{2^{j+r}}}^{\frac{2^r p+s+2}{2^{j+r}}} \prod_{i=1}^{k-1} |2(\cos 2\pi\lambda - \cos 2\pi\nu_i)|^{-2d_i} d\lambda \\
&= \frac{\sigma^2 |2(\cos 2\pi\lambda^* - \cos 2\pi\nu_k)|^{-2d_k}}{2\pi} \mathbb{V}[W_{j+r}^{2^r p+s+1}(1)].
\end{aligned}$$

Then, as  $\overline{\mathbb{V}[W_{j+r}^{2^r p+s}(1)]} = 0$  or  $\overline{\mathbb{V}[W_{j+r}^{2^r p+s+1}(1)]} = 0$ ,

$$\mathbb{V}[W_{j+r}^{2^r p+s}] \gg \mathbb{V}[W_{j+r}^{2^r p+s+1}] \quad \text{or} \quad \mathbb{V}[W_{j+r}^{2^r p+s+1}] \gg \mathbb{V}[W_{j+r}^{2^r p+s}],$$

it means,

$$\overline{\mathbb{V}[W_{j+r}^{2^r p+s}]} = 0 \quad \text{or} \quad \overline{\mathbb{V}[W_{j+r}^{2^r p+s+1}]} = 0.$$

Then,

$$\mathcal{B}_j^p = \bigcup_{i=0}^{2^r-1} \mathcal{B}_{j+r}^{2^r p+i},$$

because the choice of  $r$  is particular (see (14)). Finally, the node  $(j, p)$  is not in the tree  $\mathcal{T}$ .

However, the fact that the node  $(j+r, 2^r p+s)$  is in the tree  $\mathcal{T}_1$  means that,

$$\overline{\mathbb{V}[W_{j+r+1}^{2^{r+1} p+2s}(1)]} = W_{j+r+1}^{2^{r+1} p+2s}(1) \quad \text{and} \quad \overline{\mathbb{V}[W_{j+r+1}^{2^{r+1} p+2s+1}(1)]} = W_{j+r+1}^{2^{r+1} p+2s+1}(1).$$

As the frequency  $\nu_k$  is not in the interval  $I_{j+r}^{2^r p+s}$ , we obtain easily that

$$\overline{\mathbb{V}[W_{j+r+1}^{2^{r+1} p+2s}]} = W_{j+r+1}^{2^{r+1} p+2s} \quad \text{and} \quad \overline{\mathbb{V}[W_{j+r+1}^{2^{r+1} p+2s+1}]} = W_{j+r+1}^{2^{r+1} p+2s+1}.$$

Finally, the node  $(j+r, 2^r p+s)$  is in the tree  $\mathcal{T}$ . Using the argument, we show that the node  $(j+r^*, 2^{r^*} p+s^*)$  is in the tree  $\mathcal{T}$ . ■

## Appendix B

**Lemma 6.1** *Let  $\psi$  be a wavelet with  $q$  vanishing moments, and the associated high-pass QMF filter factorized as:*

$$\hat{g}(\omega) = (1 - e^{-i\omega})^q P(e^{i\omega}) \quad (15)$$

where  $P(\cdot)$  is a trigonometric polynomial bounded around  $\omega = 0$ . Then, for all  $j$  and  $p = 0, \dots, 2^j - 1$ , the moments of the wavelet packet function  $\psi_j^p$  are such that:

$$\mathcal{M}_{j,p}(r) = \int_{\mathbb{R}} t^r \psi_j^p(t) dt = \delta(r) \delta(p), \quad \text{for } 0 \leq r < R_p \quad (16)$$

where  $R_0 = 1$  and  $R_p = q \sum_{k=0}^{j-1} p_k$  for  $p \neq 0$ , and  $p = (p_{j-1} p_{j-2} \dots p_1 p_0)_2$  is the binary representation of  $p$ .

### Proof of Lemma 6.1:

Note that for  $p = 1$  (wavelet basis), our result specializes to the traditional relation  $\mathcal{M}_{j,1}(r) = 0$  for  $0 \leq r < q$ . The lemma can be proved either by induction in the original domain, or using explicit proof in the Fourier domain. We shall proceed according to the latter. Recall first that the Fourier transforms of  $\psi_{j+1}^{2^p}(t)$  and  $\psi_{j+1}^{2^{p+1}}(t)$  can be recursively written as:

$$\hat{\psi}_{j+1}^{2^p}(\omega) = \hat{h}(2^j \omega) \hat{\psi}_j^p(\omega) \quad (17)$$

$$\hat{\psi}_{j+1}^{2^{p+1}}(\omega) = \hat{g}(2^j \omega) \hat{\psi}_j^p(\omega) \quad (18)$$

Calculating  $\hat{\psi}_j^p(\omega)$  amounts to iterating the actions of the QMF filters  $h$  or  $g$ , from the root of the binary tree, to extract the appropriate range of frequencies. Thus, one can write that:

$$\hat{\psi}_j^p(\omega) = \left[ \prod_{k=0}^{j-1} \frac{F_{p_{k-1}}(2^{-k}\omega)}{\sqrt{2}} \right] \hat{\phi}(2^{-j}\omega) \quad (19)$$

$$\text{or equivalently } \hat{\psi}_j^p(\omega) = \left[ \prod_{k=0}^{j-1} F_{p_{j-k-1}}(2^k\omega) \right] \hat{\phi}(\omega) \quad (20)$$

where the sequence of filters  $F_{p_k}$  is chosen according to  $p$  :

$$F_{p_k} = \begin{cases} \hat{h} & \text{if } p_k = 0 \\ \hat{g} & \text{if } p_k = 1 \end{cases} \quad (21)$$

where  $p = 2^{j-1}p_{j-1} + 2^{j-2}p_{j-2} + \dots + 2p_1 + p_0$ , and  $\hat{\phi}(0) \neq 0$ . We point out that similar expressions as above can be deduced using the convolution-decimation operators associated with  $h$  and  $g$  (see Wickerhauser 94, Chap. 5[Wic94], Nielsen (1999)[Nie99]).

Let's take compactly supported wavelets with  $q$  vanishing moments, whose associated high-pass filters  $\hat{g}$  has  $q-1$  zeros at  $\omega = 0$  :

$$\hat{g}(\omega) = (1 - e^{-i\omega})^q P(e^{i\omega}) \quad (22)$$

where  $P(\cdot)$  is a trigonometric polynomial bounded around  $\omega = 0$ . The QMF low-pass filter  $\hat{h}$  has  $q-1$  zeros at  $\pi$ . The Daubechies filters satisfy these conditions. The number of vanishing moments of  $\psi_j^p(t)$  is equivalently given by the number of vanishing derivatives of  $\hat{\psi}_j^p(\omega)$  at  $\omega = 0$ , that is:

$$\mathcal{M}_{j,p}(r) = \left[ \left( \frac{1}{i} \partial_\omega \right)^r \hat{\psi}_j^p(\omega) \right]_{\omega=0} \quad \text{for } r = 0, \dots, R_p - 1 \quad (23)$$

- If  $p = 0$ , it is easy to see that  $\hat{\psi}_j^p(\omega)$  is just the product of low-pass filters, and  $\psi_j^0(t) = \phi_j(t)$  the scaling function at depth  $j$ . Then,  $\mathcal{M}_{j,0}(r) = \hat{\phi}_j(0)$ , which is non-zero with  $R_0 = 1$ . It is worth noting that if additional constraints are imposed on the wavelet choice (e.g. Coiflets),  $\mathcal{M}_{j,0}(r)$  might be zero for  $1 \leq r < q$ .
- If  $p \neq 0$ , from (20) we can write:

$$\hat{\psi}_j^p(\omega) = \prod_{k|p_{j-k-1}=1} (1 - e^{-i2^k\omega})^q Q(\omega) \quad (24)$$

where  $Q(\cdot)$  is again bounded around  $\omega = 0$ . The number of vanishing moments is then given by the number of zeros at  $\omega = 0$  which is  $R_p = q \sum_k p_k$ . The lemma follows.

**Lemma 6.2** *If the QMF  $h$  has a support in  $[N_1, N_2]$ , then the support of the wavelet packet function  $\psi_j^p(t)$  at each node  $(j, p)$  in the WP binary tree is always included in  $[-2^j(N^* + 1), 2^j(N^* + 1)]$ , with  $N^* = \max(|N_1|, |N_2|)$ .*

**Proof of Lemma 6.2:**

This is proved by induction. We begin by proving it at the first scale, where it is easy to show (see e.g. Mallat (1998)[Mal98], Proposition 7.2) that  $\psi_0$  will be supported in the interval  $[N_1, N_2]$  and  $\psi_1$  in  $[\frac{N_1-N_2+1}{2}, \frac{N_2-N_1+1}{2}]$ . The support inclusion statement is obviously true for  $\psi_0$ . For  $\psi_1$ , we first assume that there exists a finite strictly positive integer  $L$  such that  $L = |N_1 - N_2| \geq 1$ . Then, for all  $L$  even or odd:

$$\begin{aligned} \left| \frac{N_1 - N_2 + 1}{2} \right| &\leq \left\lfloor \frac{L}{2} \right\rfloor + 1 \\ &\leq \left\lfloor \frac{|N_1| + |N_2|}{2} \right\rfloor + 1 \\ &\leq \max(|N_1|, |N_2|) + 1 \end{aligned} \quad (25)$$

which proves the support inclusion for  $\psi_1$ .

Assume now that  $\text{Supp } \psi_j^{p_1} \subseteq [-2^j(N^* + 1), 2^j(N^* + 1)]$ . Then, using the recursive definition of the children functions  $\psi_{j+1}^{2p_1}$  and  $\psi_{j+1}^{2p_1+1}$ ,

$$\text{Supp } \psi_{j+1}^{2p_1} \subseteq [-2^j(N^* + 1) + 2^j N_1, 2^j(N^* + 1) + 2^j N_2] \quad (26)$$

$$\text{Supp } \psi_{j+1}^{2p_1+1} \subseteq [-2^j(N^* + 1) + 2^j(1 - N_2), 2^j(N^* + 1) + 2^j(1 - N_1)] \quad (27)$$

where the support properties of the QMF filters  $h$  and  $g$  were also used. It follows that:

$$\begin{aligned} |2^j(N^* + 1) + 2^j N_l| &\leq 2^j(N^* + 1) + 2^j \max_{l=1,2} |N_l| \\ &\leq 2^j(N^* + 1) + 2^j + 2^j N^* \\ &\leq 2^{j+1}(N^* + 1) \end{aligned} \quad (28)$$

$$\begin{aligned} |2^j(N^* + 1) + 2^j(1 - N_l)| &\leq 2^j(N^* + 1) + 2^j + 2^j \max_{l=1,2} |N_l| \\ &\leq 2^{j+1}(N^* + 1) \end{aligned} \quad (29)$$

for  $l = 1, 2$ . This completes the proof.

**Lemma 6.3** *Let  $\mathcal{I}$  be a collection of disjoint dyadic intervals  $I_p^j$  whose union is the positive half line, and  $\mathcal{B} = \{\psi_j^p(t - 2^j k) : 0 \leq k < 2^{J-j}, I_p^j \in \mathcal{I}\}$  is the associated orthonormal basis. Let  $h$  and  $g$  the QMFs as defined in Eq.22. The vanishing moments  $\mathcal{M}_{j_1, j_2}^{p_1, p_2}(m)$  of the intercorrelation function  $\Lambda_{j_1, j_2}^{p_1, p_2}(h)$  of  $\psi_{j_1}^{p_1}(t)$  and  $\psi_{j_2}^{p_2}(t) \in \mathcal{B}$  satisfy:*

1.  $p_1 \neq 0$  and  $p_2 \neq 0$ :

$$\mathcal{M}_{j_1, j_2}^{p_1, p_2}(m) = 0, \quad \text{for } 0 \leq m < R_{p_1} + R_{p_2}. \quad (30)$$

2.  $p_1 \neq 0$  or  $p_2 \neq 0$ :

$$\mathcal{M}_{j_1, j_2}^{p_1, p_2}(m) = 0, \quad \text{for } 0 \leq m < R_{\max(p_1, p_2)}. \quad (31)$$

3.  $p_1 = p_2 = 0$ :

$$\mathcal{M}_{j_1, j_2}^{p_1, p_2}(m) = 0, \quad \text{for } 1 \leq m < 2q. \quad (32)$$

Furthermore, the support of  $\Lambda_{j_1, j_2}^{p_1, p_2}(h)$  is included in  $[-(N^* + 1)(2^{j_1} + 2^{j_2}), (N^* + 1)(2^{j_1} + 2^{j_2})]$ .

### Proof of Lemma 6.3:

By definition of the intercorrelation function, we have:

$$\Lambda_{j_1, j_2}^{p_1, p_2}(h) = \int \psi_{j_1}^{p_1}(t) \psi_{j_2}^{p_2}(t - h) dt \quad (33)$$

Given the fact that these wavelet packet functions belong to the orthonormal basis  $\mathcal{B}$ , we immediately obtain that at integer lags:

$$\Lambda_{j_1, j_2}^{p_1, p_2}(n) = \delta(j_1 - j_2) \delta(p_1 - p_2) \delta(n) \quad (34)$$

As far as the support is concerned, it is not a difficult matter to see, using Lemma 6.2, that  $\Lambda_{j_1, j_2}^{p_1, p_2}(h)$  is supported in  $[-(N^* + 1)(2^{j_1} + 2^{j_2}), (N^* + 1)(2^{j_1} + 2^{j_2})]$ .

Let's now turn to the moments of  $\Lambda_{j_1, j_2}^{p_1, p_2}(h)$ .

1.  $p_1 \neq 0$  and  $p_2 \neq 0$ :

In this case, we know from Lemma 6.1 that  $\psi_{j_1}^{p_1}$  and  $\psi_{j_2}^{p_2}$  have respectively  $R_{p_1} = \sum_{k=0}^{j_1-1} p_1^k$  and  $R_{p_2} = \sum_{k=0}^{j_2-1} p_2^k$  vanishing moments. Then,  $\Lambda_{j_1, j_2}^{p_1, p_2}(h)$  will have  $R_{p_1} + R_{p_2}$  vanishing moments since,

$$\begin{aligned} \int h^m \Lambda_{j_1, j_2}^{p_1, p_2}(h) dh &= \int \int h^m \psi_{j_1}^{p_1}(t) \psi_{j_2}^{p_2}(t-h) dt dh \\ &= - \int \int (v-u)^m \psi_{j_1}^{p_1}(v) \psi_{j_2}^{p_2}(u) du dv \\ &= - \sum_{n=0}^m (-1)^n \binom{m}{n} \int v^{m-n} \psi_{j_1}^{p_1}(v) dv \int u^n \psi_{j_2}^{p_2}(u) du \\ &= 0, \quad \text{for } 0 \leq m < R_{p_1} + R_{p_2}. \end{aligned} \quad (35)$$

where we used uniform convergence and continuity to invert the order of summation and integration. Note that the Fubini theorem allows us to invert the order of integrals.

2.  $p_1 \neq 0$  or  $p_2 \neq 0$ :

Without loss of generality, assume that  $p_1 \neq 0$  and  $p_2 = 0$ . The same reasoning as above can be adopted to conclude that:

$$\begin{aligned} \int h^m \Lambda_{j_1, j_2}^{p_1, 0}(h) dh &= - \sum_{n=0}^m (-1)^n \binom{m}{n} \int v^{m-n} \psi_{j_1}^{p_1}(v) dv \int u^n \phi_{j_2}(u) du \\ &= 0, \quad \text{for } 0 \leq m < R_{p_1}. \end{aligned} \quad (36)$$

3.  $p_1 = p_2 = 0$ :

In an orthonormal basis of wavelet packets, this situation is not possible unless  $j_1 = j_2 = j$ . Thus,

$$\begin{aligned} \int h^m \Lambda_{j, j}^{0, 0}(h) dh &= 2^{-j} \int \int h^m \phi(2^{-j}t) \phi(2^{-j}(t-h)) dt dh \\ &= 2^{j(m+1)} \int \int u^m \phi(v) \phi(v-u) du dv \\ &= 0, \quad \text{for } 1 \leq m < 2q. \end{aligned} \quad (37)$$

where the latter result is proved in [Bey92].

#### Proof of Proposition 4.1:

Here we are interested in the covariance between the wavelet packet coefficients  $W_{j_1}^{p_1}(k_1)$  and  $W_{j_2}^{p_2}(k_2)$ . We have

$$\text{Cov} [W_{j_1}^{p_1}(k_1), W_{j_2}^{p_2}(k_2)] = \int \int \mathbb{E}[X(t)X(s)] \psi_{j_1}^{p_1}(t - 2^{j_1}k_1) \psi_{j_2}^{p_2}(s - 2^{j_2}k_2) dt ds \quad (38)$$

$$= \int \int \cos(\nu(t-s)) |t-s|^{2d-1} \psi_{j_1}^{p_1}(t - 2^{j_1}k_1) \psi_{j_2}^{p_2}(s - 2^{j_2}k_2) dt ds \quad (39)$$

With a change of variables,  $u = t - 2^{j_1}k_1$  and  $v = s - 2^{j_2}k_2$ ,

$$\text{Cov} [W_{j_1}^{p_1}(k_1), W_{j_2}^{p_2}(k_2)] = \int \int \cos(\nu(u + 2^{j_1}k_1 - v - 2^{j_2}k_2)) |u + 2^{j_1}k_1 - v - 2^{j_2}k_2|^{2d-1} \psi_{j_1}^{p_1}(u) \psi_{j_2}^{p_2}(v) du dv$$

With an other change of variables,  $u = t$  and  $v = t - h$

$$\begin{aligned} \text{Cov} [W_{j_1}^{p_1}(k_1), W_{j_2}^{p_2}(k_2)] &= \int \int \cos(\nu(h + 2^{j_1}(k_1 - 2^{j_2-j_1}k_2))) |h + 2^{j_1}(k_1 - 2^{j_2-j_1}k_2)|^{2d-1} \psi_{j_1}^{p_1}(t) \psi_{j_2}^{p_2}(t-h) dt dh \\ &= \int \cos(\nu(h + 2^{j_1}(k_1 - 2^{j_2-j_1}k_2))) |h + 2^{j_1}(k_1 - 2^{j_2-j_1}k_2)|^{2d-1} \Lambda_{j_1, j_2}^{p_1, p_2}(h) dh, \end{aligned} \quad (41)$$

where  $\Lambda_{j_1, j_2}^{p_1, p_2}(h) = \int \psi_{j_1}^{p_1}(t) \psi_{j_2}^{p_2}(t-h) dt$ .

Denoting  $\alpha = 2^{j_1}(k_1 - 2^{j_2-j_1}k_2)$ , we obtain,

$$\text{Cov} [W_{j_1}^{p_1}(k_1), W_{j_2}^{p_2}(k_2)] = \int \cos(\nu(h + \alpha)) |h + \alpha|^{2d-1} \Lambda_{j_1, j_2}^{p_1, p_2}(h) dh. \quad (42)$$

According to the support of  $\psi$ , from Lemma 6.2 we know that the support of  $\Lambda_{j_1, j_2}^{p_1, p_2}(h)$  is included in  $[-(2^{j_1} + 2^{j_2})(N^* + 1), (2^{j_1} + 2^{j_2})(N^* + 1)]$ . As  $h$  is in the support of  $\Lambda_{j_1, j_2}^{p_1, p_2}$  and by assumption  $\alpha > (N^* + 1)(2^{j_1} + 2^{j_2})$ , we have  $h/\alpha < 1$ . Then,

- if we denote  $f(h + \alpha) = |h + \alpha|^{2d-1}$ , we have

$$f(h + \alpha) = |h + \alpha|^{2d-1} = |\alpha|^{2d-1} \left| 1 + \frac{h}{\alpha} \right|^{2d-1} = |\alpha|^{2d-1} \left\{ 1 + \sum_{i=1}^{\infty} \binom{2d-1}{i} \left( \frac{h}{\alpha} \right)^i \right\}. \quad (43)$$

- when  $\alpha$  is large,  $h/\alpha$  tends to 0, and then,

$$\cos(\nu(\alpha + h)) = \cos \left( \nu\alpha \left( 1 + \frac{h}{\alpha} \right) \right) \sim \cos(\nu\alpha). \quad (44)$$

Hence, from (43)-(44), it follows that, for large  $\alpha$

$$\text{Cov} [W_{j_1}^{p_1}(k_1), W_{j_2}^{p_2}(k_2)] \sim |\alpha|^{2d-1} \cos(\nu\alpha) \left\{ \int \Lambda_{j_1, j_2}^{p_1, p_2}(h) dh + \sum_{i=1}^{\infty} \binom{2d-1}{i} \int \left( \frac{h}{\alpha} \right)^i \Lambda_{j_1, j_2}^{p_1, p_2}(h) dh \right\}. \quad (45)$$

We must then provide an upper bound on the integrals inside the braces. In the following we distinguish three different cases depending on the number of vanishing moments of  $\Lambda_{j_1, j_2}^{p_1, p_2}$  according to Lemma 6.3, that is:

1. If  $p_1 \neq 0$  and  $p_2 \neq 0$ , then  $\mathcal{M}_{p_1, p_2}^{j_1, j_2}(m) = 0$ , for  $0 \leq m < R_{p_1} + R_{p_2}$ . We denote  $q^* = R_{p_1} + R_{p_2}$ . Then, using the fact that the  $q^*$  first moments of  $\Lambda_{j_1, j_2}^{p_1, p_2}$  are null,

$$\text{Cov} [W_{j_1}^{p_1}(k_1), W_{j_2}^{p_2}(k_2)] \sim C_1 |\alpha|^{2d-1-q^*} + R_{q^*+1}, \quad (46)$$

with

$$C_1 = \cos(\nu\alpha) \frac{(2d-1)!}{q^*!(2d-1-q^*)!} \int h^{q^*} \Lambda_{j_1, j_2}^{p_1, p_2}(h) dh \quad (47)$$

and

$$R_{q^*+1} = \cos(\nu\alpha) |\alpha|^{2d-1} \sum_{i=q^*+1}^{\infty} \binom{2d-1}{i} \int \left( \frac{h}{\alpha} \right)^i \Lambda_{j_1, j_2}^{p_1, p_2}(h) dh. \quad (48)$$

$$|R_{q^*+1}| \leq |\alpha|^{2d-1} \binom{2d-1}{q^*} \left| \sum_{i=q^*+1}^{\infty} \int \int \left( \frac{h}{\alpha} \right)^i \psi_{j_1}^{p_1}(t) \psi_{j_2}^{p_2}(t-h) dt dh \right| \quad (49)$$

$$= |\alpha|^{2d-1} \binom{2d-1}{q^*} \left| \sum_{i=q^*+1}^{\infty} \int \int \left( \frac{t-h}{\alpha} \right)^i \psi_{j_1}^{p_1}(t) \psi_{j_2}^{p_2}(h) dt dh \right| \quad (50)$$

$$\leq |\alpha|^{2d-1} \sum_{i=q^*+1}^{\infty} \binom{2d-1}{q^*} \int \int \left( \sup_{t, h} \left| \frac{t-h}{\alpha} \right| \right)^i |\psi_{j_1}^{p_1}(t) \psi_{j_2}^{p_2}(h)| dt dh \quad (51)$$

$$= |\alpha|^{2d-1} \binom{2d-1}{q^*} \int \int |\psi_{j_1}^{p_1}(t) \psi_{j_2}^{p_2}(h)| dt dh \sum_{i=q^*+1}^{\infty} \left( \sup_{t, h} \left| \frac{t-h}{\alpha} \right| \right)^i \quad (52)$$

$$= C_2 |\alpha|^{2d-1} \sum_{i=1}^{\infty} \beta^{q^*+i}, \quad (53)$$

$$(54)$$

where

$$\beta = \sup_{t,h} \left| \frac{t-h}{\alpha} \right| \quad \text{and} \quad C_2 = \binom{2d-1}{q^*} \int |\psi_{j_1}^{p_1}(t)| dt \int |\psi_{j_2}^{p_2}(h)| dh. \quad (55)$$

Then,

$$|R_{q^*+1}| \leq C_3 |\alpha|^{2d-1-q^*-1}, \quad (56)$$

where  $C_3$  is a finite constant. Finally,

$$\text{Cov} [W_{j_1}^{p_1}(k_1), W_{j_2}^{p_2}(k_2)] = O \left( |2^{j_1} k_1 - 2^{j_2} k_2|^{2d-1-q^*} \right), \quad (57)$$

for  $|2^{j_1} k_1 - 2^{j_2} k_2| > (N^* + 1)(2^{j_1} + 2^{j_2})$  and  $q^* = R_{p_1} + R_{p_2}$ .

2. If  $p_1 \neq 0$  or  $p_2 \neq 0$  then  $\mathcal{M}_{p_1, p_2}^{j_1, j_2}(m) = 0$ , for  $0 \leq m < R_{\max(p_1, p_2)}$ . Using the same argument as previously, we find that

$$\text{Cov} [W_{j_1}^{p_1}(k_1), W_{j_2}^{p_2}(k_2)] = O \left( |2^{j_1} k_1 - 2^{j_2} k_2|^{2d-1-q^*} \right), \quad (58)$$

with  $|2^{j_1} k_1 - 2^{j_2} k_2| > (N^* + 1)(2^{j_1} + 2^{j_2})$  and  $q^* = R_{\max(p_1, p_2)}$ .

3.  $p_1 = p_2 = 0$ :  $\mathcal{M}_{p_1, p_2}^{j_1, j_2}(m) = 0$  for  $1 \leq m < 2q$ . In this particular case, we have necessary  $j_1 = j_2 = j$ . We then upper-bound the covariance as follows,

$$\text{Cov} [W_j^0(k_1), W_j^0(k_2)] \sim |\alpha|^{2d-1} \left\{ \int \cos(\nu(\alpha + h)) \Lambda_{j,j}^{0,0}(h) dh + \cos(\nu\alpha) \sum_{i=1}^{\infty} \binom{2d-1}{i} \int \left( \frac{h}{\alpha} \right)^i \Lambda_{j,j}^{0,0}(h) dh \right\}. \quad (59)$$

Then,

$$\text{Cov} [W_j^0(k_1), W_j^0(k_2)] \sim C_0 |\alpha|^{2d-1} + C_1 |\alpha|^{2d-1-2q} + R_{2q+1}, \quad (60)$$

where

$$C_0 = \int \cos(\nu(h + \alpha)) \Lambda_{j,j}^{0,0}(h) dh, \quad C_1 = \cos(\nu\alpha) \frac{(2d-1)!}{(2q)!(2d-1-2q)!} \int h^{2q} \Lambda_{j,j}^{0,0}(h) dh, \quad (61)$$

and

$$|R_{2q+1}| \leq |\cos(\nu\alpha)| |\alpha|^{2d-1} \sum_{i=2q+1}^{\infty} \binom{2d-1}{i} \int \left( \frac{h}{\alpha} \right)^i \Lambda_{j,j}^{0,0}(h) dh \quad (62)$$

$$= O(|\alpha|^{2d-1-2q-1}) \quad (63)$$

As previously, we remark that when  $\alpha$  is large, equation (44) holds and then,

$$C_0 \sim \cos(\nu\alpha) \int \int \phi_j(t) \phi_j(t-h) dt dh \quad (64)$$

$$= 2^j \cos(\nu\alpha) |\Phi(0)|^2 = 2^j \cos(\nu\alpha). \quad (65)$$

Finally, using a similar argument as in the previous cases, we find that

$$\text{Cov} [W_j^0(k_1), W_j^0(k_2)] = O(|2^j(k_1 - k_2)|^{2d-1}). \quad (66)$$

for  $|k_1 - k_2| > 2(N^* + 1)$ . ■



## References

- [And86] J. Andel. Long memory time series. *Kybernetika*, 22:105–123, 1986.
- [Ber94] J. Beran. *Statistics for Long-Memory Processes*. Chapman and Hall, London, 1994.
- [Bey92] G. Beylkin. On the representation of operators in bases of compactly supported wavelets. *SIAM Journal on Numerical Analysis*, 29(6):1716–1740, 1992.
- [Chu96] C.F. Chung. Estimating a generalized long-memory process. *Journal of Econometrics*, 73:237–259, 1996.
- [CMW92] R.R. Coifman, Y. Meyer, and M.V. Whickerhauser. Wavelet analysis and signal processing. In *Wavelets and their applications, Boston.*, pages 153–178, 1992.
- [CW92] R.R. Coifman and M.V. Whickerhauser. Entropy-based algorithms for best basis selection. *IEEE Transactions on Information Theory*, 38:713–718, 1992.
- [DH87] R.B. Davies and D.S. Harte. Tests for hurst effect. *Biometrika*, 74:95–102, 1987.
- [Dij94] P. Dijkerman. On teh correlation structure of the wavelet coefficients of fractional brownian motion. *IEEE Trans. Inf. Theory*, 45:1609–1612, 1994.
- [DMvS98] D. Donoho, S. Mallat, and R. von Sachs. Estimating covariances of locally stationary processes: rates of convergence of best basis methods. *Technical Report, University of Berkeley*, 1998.
- [DT93] M. Deriche and A. Tewfik. Maximum likelihood estimation of the parameters of discrete fractionally differenced gaussian noise process. *IEEE Transactions on Signal Processing*, 41:2977–2989, 1993.
- [Fla92] P. Flandrin. Wavelet analysis and synthesis of fractional brownian motion. *IEEE Trans. Inf. Theory*, 38:910–917, 1992.
- [GJ80] C.W.J. Granger and R. Joyeux. An introduction to long-memory time series models and fractional differencing. *JTSA*, 1:15–29, 1980.
- [GZW89] H.L. Gray, N.F. Zhang, and W.A Woodward. On generalized fractional processes. *JTSA*, 10:233–257, 1989.
- [Hos81] J.R.M. Hosking. Fractional differencing. *Biometrika*, 68:165–176, 1981.
- [Jen99] M.J. Jensen. Using wavelets to obtain a consistent ordinary least squares estimator of the long-memory parameter. *Journal of Forecasting*, 18:17–32, 1999.
- [Jen00] M.J. Jensen. An alternative maximum likelihood estimator of long-memory processes using compactly supported wavelets. *Journal of Economic Dynamics and Control*, 24:361–387, 2000.
- [Mal98] S.G. Mallat. *A wavelet tour in signal processing*. Academic Press, 1998.
- [MW96] M.C. McCoy and A.T. Walden. Wavelet analysis and synthesis of stationary long-memory processes. *Journal of Computational and Graphical Statistics*, 5:1–31, 1996.
- [MZP98] S. Mallat, Z. Zhang, and G. Papanicolaou. Adaptive covariance estimation of locally stationary processes. *Annals of Statistics*, 26:1–47, 1998.
- [Nie99] M. Nielsen. *Size Properties of Wavelet Packets*. PhD thesis, Washington University, St. Louis, 1999.
- [PW00] D.B. Percival and A.T. Walden. *Wavelet Methods for Time Series Analysis*. Cambridge: Cambridge University Press, 2000.
- [TK92] A.H. Tewfik and M. Kim. Correlation structure of the discrete wavelet coefficient of fractional brownian motion. *IEEE Transactions on Information Theory*, 38:904–909, 1992.

- [Whi01] B. Whitcher. Simulating gaussian stationary processes with unbounded spectra. Journal of Computational and Graphical Statistics, 10:112–134, 2001.
- [Wic91] M.V. Wickerhauser. Fast approximate factor analysis. In Curves and Surfaces in Computer Vision and Graphics II, volume 1610, pages 23–32. SPIE, 1991.
- [Wic94] M.V. Wickerhauser. Adapted Wavelet Analysis from Theory to Software. IEEE Press, 1994.
- [Wor90] G. W. Wornell. A karhunen-loeve-like expansion for 1/f process via wavelets. IEEE Trans. Inf. Theory, 36(4):859–861, 1990.
- [Wor96] G.W. Wornell. Signal Processing with Fractals: A Wavelet-Based Approach. Prentice Hall, 1996.