16th European Signal Processing Conference (EUSIPCO 2008), Lausanne, Switzerland, August 25-29, 2008, copyright by EURASIP

LIMIT DISTRIBUTIONS FOR WAVELET PACKET COEFFICIENTS OF BAND-LIMITED STATIONARY RANDOM PROCESSES

Abdourrahmane M. Atto[1], *Dominique Pastor*[2]

Institut TELECOM, TELECOM Bretagne
LAB-STICC, CNRS UMR 3192
Technopôle Brest-Iroise, CS 83818
29238 Brest Cedex 3, FRANCE

ABSTRACT

This paper addresses the limit distribution of wavelet packet coefficients obtained by decomposing band-limited random processes. When the wavelet decomposition filters satisfy a certain property of regularity, strictly stationary band-limited random processes yield wavelet packet coefficients that are asymptotically uncorrelated and Gaussian distributed when the resolution level increases. For any given path, the variance of the limit distribution is the value of the power spectral density of the input process at a specific frequency. Experimental results are presented to assess the convergence rate when Daubechies filters are used.

1. INTRODUCTION

The Discrete Wavelet Packet Transform (DWPT) allows many possible representations of functions by providing various Hilbertian bases. It is not computationally expensive and has some remarkable properties such as the sparse representation it provides for smooth signals [1, 2] or the 'whitening effect' it asymptotically yields for a large class of random processes [3, 4]. For applications in signal processing, time-series analysis and communication systems, the DWPT is a powerful tool and, as such, has received much interest in recent years. The DWPT is computed via projections on wavelet packet spaces, which are Hilbertian functional spaces hereafter denoted $\mathbf{W}_{j,n}$ (see section 2 for details). In this notation, $j \in \mathbb{N}$ is the resolution level and the shift parameter n is valued in $\{0, 1, \ldots, 2^j - 1\}$.

Let X be a second-order centred strictly stationary random process, continuous in quadratic mean. Assume that X has finite cumulants and has a polyspectrum $\gamma_N(\omega_1, \cdots, \omega_N)$ for every natural number N and every $(\omega_1, \cdots, \omega_N) \in \mathbb{R}^N$. When $N = 1$, γ_1 is the spectrum of X and is simply denoted γ. Let $c_{j,n} = (c_{j,n}[k])_{k\in\mathbb{Z}}$ stand for the sequence of the wavelet packet coefficients of X with respect to $\mathbf{W}_{j,n}$. According to [3, Proposition 12], if the shift parameter n is a constant and j tends to infinity, the sequence $(c_{j,n})_{j\in\mathbb{N}}$ converges, in a 'distributional' sense specified below, to a discrete white Gaussian process with variance $\gamma(0)$.

However, most of the paths of an DWPT tree involve nodes (j, n) whose shift parameters $n = n(j)$ vary with j. In [4], by taking this dependency into account, and by using specific DWPT filters, the asymptotic decorrelation of the wavelet packet coefficients is established for band-limited wide-sense stationary random processes. The limit variance of the wavelet packet coefficients is the value of the spectrum of the input random process at a frequency associated to the path followed in the decomposition tree. In the present paper, the focus is the limit distribution of the wavelet packet coefficients of a strictly random process, for any path of the DWPT. The limit distribution still depends on the DWPT filters and the path considered in the decomposition tree. In this respect, the results presented below complement those given in [3, 4].

The paper is organized as follows. Section 2 recalls some basics about the DWPT. After presenting preliminary results concerning the DWPT of strictly stationary random processes (see section 3), the main result of this paper is proposition 1. It is stated in section 4.3 on the basis of the material given in sections 4.1 and 4.2. Experimental results are presented in section 5 to illustrate and confirm proposition 1. The conclusion of this paper announces a theoretical extension of this work.

2. THE DWPT

A DWPT is performed by using decomposition filters with impulse responses h_ε, $\varepsilon = 0, 1$, that satisfy the following properties. First, each filter h_ε is an element of $\ell^2(\mathbb{Z})$ and its Fourier transform is hereafter defined by

$$H_\varepsilon(\omega) = \frac{1}{\sqrt{2}} \sum_{\ell\in\mathbb{Z}} h_\varepsilon[\ell] \exp(-i\ell\omega). \tag{1}$$

Second, the matrix $\begin{pmatrix} H_0(\omega) & H_1(\omega) \\ H_0(\omega + \frac{\pi}{2}) & H_1(\omega + \frac{\pi}{2}) \end{pmatrix}$ is unitary for every real number ω. The unitary nature of this matrix implies that $|H_\varepsilon(\omega)| \leqslant 1$ for every $\varepsilon = 0, 1$ and every $\omega \in \mathbb{R}$.

Let Φ be a function such that $\{\tau_k\Phi \ : k \in \mathbb{Z}\}$ is an orthonormal system of $L^2(\mathbb{R})$, where $\tau_k\Phi \ : t \longmapsto \Phi(t-k)$. Let $\mathbf{U}$ be the closure of the space spanned by this orthonormal system. Let us define the following sequence of elements of $L^2(\mathbb{R})$ by recursively setting, for $n \in \mathbb{N}$ and $\varepsilon \in \{0, 1\}$,

$$\left\{ \begin{array}{lcl} W_\varepsilon(t) & = & \sqrt{2}\sum_{\ell\in\mathbb{Z}} h_\varepsilon[\ell]\Phi(2t-\ell) \\ W_{2n+\varepsilon}(t) & = & \sqrt{2}\sum_{\ell\in\mathbb{Z}} h_\varepsilon[\ell]W_n(2t-\ell). \end{array} \right. \tag{2}$$

The function Φ in Eq. (2) is not necessarily the scaling function associated with h_0. If Φ is this scaling function, we have $W_0 = \Phi$ and Eq. (2) holds true even if $n = 0$. For any pair (j, n) of natural numbers and $k \in \mathbb{Z}$, we define

$$W_{j,n,k}(t) = W_{j,n}(t - 2^j k) = 2^{-j/2} W_n(2^{-j}t - k). \tag{3}$$

Then, $\{W_{j,n,k} : k \in \mathbb{Z}\}$ is an orthonormal system of $L^2(\mathbb{R})$. The closure of the functional space spanned by this system is called the *wavelet packet space* $\mathbf{W}_{j,n}$. Any $W_{j,n,k}$ is called a *wavelet packet function*. The DWPT decomposition of the function space $\mathbf{U}$ then consists in the following recursive splitting of $\mathbf{U}$ into orthogonal subspaces:

$$\left\{ \begin{array}{lcl} \mathbf{U} & = & \mathbf{W}_{1,0} \oplus \mathbf{W}_{1,1}, \\ \mathbf{W}_{j,n} & = & \mathbf{W}_{j+1,2n} \oplus \mathbf{W}_{j+1,2n+1}, \end{array} \right.$$

[1] am.atto@telecom-bretagne.eu
[2] dominique.pastor@telecom-bretagne.eu

for every natural number j and every $n = 0, 1, 2, \ldots, 2^j - 1$. According to the foregoing, $\mathbf{U}$ can be split into orthogonal sums of wavelet packet spaces. Thus, for any given Hilbertian random process $X(t)$, the coefficients of the projection of X on $\mathbf{W}_{j,n}$ define a random sequence $(c_{j,n}[k])_{k\in\mathbb{Z}}$ where

$$c_{j,n}[k] = \int_{\mathbb{R}} X(t) W_{j,n,k}(t) dt. \quad (4)$$

3. PRELIMINARY RESULTS ABOUT THE DWPT OF A STRICTLY STATIONARY RANDOM PROCESS

Henceforth, $X(t)$ stands for some centred, strictly stationary random process. It is also assumed that X has finite cumulants and has a polyspectrum $\gamma_N(\omega_1, \omega_2, \cdots, \omega_N)$ for every natural number N and every $(\omega_1, \omega_2, \cdots, \omega_N) \in \mathbb{R}^N$. The polyspectrum is the Fourier transform of the $(N+1)$-th cumulant of X. When $N = 1$, γ_1 is the spectrum of X. This spectrum is hereafter denoted γ. From now on, it is assumed that γ_N belongs to $L^\infty(\mathbb{R}^N)$ and that there exist two positive real numbers ν and C such that $|\mathcal{F}\Phi(\omega)| \leqslant C/(1 + |\omega|^{1+\nu})$, $\omega \in \mathbb{R}$, where $\mathcal{F}f$ denotes the Fourier transform of $f \in L^1(\mathbb{R}) \cup L^2(\mathbb{R})$ and is given by $\mathcal{F}f(\omega) = \int f(t) \exp(-i\omega t) dt$ if $f \in L^1(\mathbb{R})$. The $(N+1)$-th cumulant of the random process $c_{j,n}$ has the following integral form (see [5, Proposition 1]):

$$\begin{aligned} &\mathsf{cum}_{j,n}[k_1, k_2, \cdots, k_N] \\ &\quad = \tfrac{2^{-j(N-1)/2}}{(2\pi)^N} \int_{\mathbb{R}^N} d\omega_1 d\omega_2 \cdots d\omega_N \\ &\qquad \exp\left(-i(k_1\omega_1 + k_2\omega_2 + \cdots + k_N\omega_N)\right) \\ &\qquad \gamma_N(-\omega_1 2^{-j}, -\omega_2 2^{-j}, \cdots, -\omega_N 2^{-j}) \\ &\qquad \mathcal{F}W_n(-\omega_1 - \omega_2 - \cdots - \omega_N) \\ &\qquad \mathcal{F}W_n(\omega_1)\mathcal{F}W_n(\omega_2) \cdots \mathcal{F}W_n(\omega_N). \end{aligned} \quad (5)$$

If the shift parameter n is constant, it follows from Lebesgue's dominated convergence theorem that, for any natural number $N > 1$, $\mathsf{cum}_{j,n}[k_1, k_2, \cdots, k_N]$ tends to 0 uniformly in $k_1, k_2, \ldots, k_N$ when j tends to ∞. This is a consequence of [3, Proposition 11].

Now, from Eq. (5), it follows that when $N = 1$, the cumulant $\mathsf{cum}_{j,n}[k]$ or order 2 of X, that is the autocorrelation function $R_{j,n}[k]$ of the random process $c_{j,n}$, is

$$R_{j,n}[k] = \frac{1}{2\pi} \int_{\mathbb{R}} \gamma(\frac{\omega}{2^j}) |\mathcal{F}W_n(\omega)|^2 \exp(ik\omega)\, d\omega. \quad (6)$$

If $\gamma \in L^\infty(\mathbb{R})$ and is continuous at 0, the integrand on the right hand side (rhs) of Eq. (6) is integrable and the limit of $\gamma(\omega/2^j)$ is $\gamma(0)$ when j tends to ∞. Therefore, for every given natural number n, it follows from Lebesgue's dominated convergence theorem applied to Eq. (6) (see for instance [3, Corollary 5]) that $\lim_{j\to+\infty} R_{j,n}[k] = \gamma(0)\delta[k]$, where $\delta[\cdot]$ is the standard Kronecker symbol. According to the foregoing, when the shift parameter n is a constant function of the resolution level j and j tends to ∞, the sequence $(c_{j,n})_{j\in\mathbb{N}}$, associated with the strictly stationary random process X considered above, converges to a discrete white Gaussian process with standard deviation $\gamma(0)$ in the following 'distributional' sense: Given any natural number N and any N-uple $k_1, k_2, \ldots, k_N$ of integers, the distribution of the random vector $(c_{j,n}[k_1], c_{j,n}[k_2], \ldots, c_{j,n}[k_N])$ converges, when j tends to infinity, to the centred N-variate normal distribution $\mathcal{N}(0, \gamma(0)\mathrm{I}_N)$ with covariance matrix $\gamma(0)\mathrm{I}_N$, where I_N is the $N \times N$ identity matrix.

When the shift parameter $n = n(j)$ varies with j, which is the case for most paths of the DWPT tree, Lebesgue's dominated convergence theorem does not apply to Eqs. (5), (6) and the analysis of the statistical behaviour of $c_{j,n(j)}$ when j tends to infinity becomes more intricate. This behaviour is studied in section 4 after introducing some material in sections 4.1 and 4.2.

4. MAIN RESULTS

The main result of this paper is Proposition 1, given in section 4.3 below. To understand this theoretical result, the following material is needed. This material concerns the representation of DWPT paths by means of binary sequences (see section 4.1) and the Shannon DWPT of band-limited functions (see section 4.2). In addition, it is more convenient to write the cumulant given by Eq. (5) in the following equivalent form

$$\begin{aligned} &\mathsf{cum}_{j,n}[k_1, k_2, \cdots, k_N] \\ &\quad = \tfrac{1}{(2\pi)^N} \int_{\mathbb{R}^N} d\omega_1 d\omega_2 \cdots d\omega_N \\ &\qquad \exp\left(-i2^j(k_1\omega_1 + k_2\omega_2 + \cdots + k_N\omega_N)\right) \\ &\qquad \gamma_N(-\omega_1, -\omega_2, \cdots, -\omega_N) \\ &\qquad \mathcal{F}W_{j,n}(-\omega_1 - \omega_2 - \cdots - \omega_N) \\ &\qquad \mathcal{F}W_{j,n}(\omega_1)\mathcal{F}W_{j,n}(\omega_2) \cdots \mathcal{F}W_{j,n}(\omega_N). \end{aligned} \quad (7)$$

This equality derives from a straightforward change of variable and the relation $\mathcal{F}W_{j,n}(\omega) = 2^{j/2}\mathcal{F}W_n(2^j\omega)$, which is a consequence of the second equality in Eq. (3). In the same way, the autocorrelation function (6) equals

$$R_{j,n}[k] = \frac{1}{2\pi} \int_{\mathbb{R}} \gamma(\omega) |\mathcal{F}W_{j,n}(\omega)|^2 \exp\left(i2^j k\omega\right) d\omega. \quad (8)$$

4.1 Binary representations of the paths of the DWPT decomposition tree

With the same notations as in section 2, a given wavelet packet path $\mathcal{P}$ is described by a sequence of nested functional subspaces: $\mathcal{P} = (\mathbf{U}, \{\mathbf{W}_{j,n(j)}\}_{j\in\mathbb{N}})$, where $\mathbf{W}_{j,n(j)} \subset \mathbf{W}_{j-1,n(j-1)}$. By construction, each $\mathbf{W}_{j,n(j)}$ is obtained by recursively decomposing $\mathbf{U}$ with a particular sequence of filters $(h_{\varepsilon_\ell})_{\ell=1,2,\cdots,j}$ where each $\varepsilon_\ell \in \{0,1\}$. Therefore, the shift parameter is

$$n(j) = \sum_{\ell=1}^{j} \varepsilon_\ell 2^{j-\ell} \in \{0, 1, \ldots, 2^j - 1\} \quad (9)$$

at every resolution level j. Note also the easy relation

$$n(j) = 2n(j-1) + \varepsilon_j, \quad (10)$$

for $j \in \mathbb{N}$, with the convention $n(0) = 0$. Thus, path $\mathcal{P}$ can be assigned to the binary sequence $\lambda = (\varepsilon_\ell)_{\ell\geqslant 1}$ of elements of $\{0,1\}$. Conversely, any binary sequence $\lambda = (\varepsilon_\ell)_{\ell\in\mathbb{N}}$ where each ε_ℓ is an element of $\{0,1\}$ specifies a unique path $\mathcal{P}_\lambda$ of the decomposition tree. At each node of this path, the shift parameter n depends on j and λ via Eq. (9) so that the notation $n = n_\lambda(j)$ will hereafter be used to indicate this dependence.

Let us consider a path $\mathcal{P}_\lambda = (\mathbf{U}, \{\mathbf{W}_{j,n_\lambda(j)}\}_{j\in\mathbb{N}})$ of the DWPT decomposition tree associated with a binary sequence λ of elements of $\{0,1\}$. At each resolution level j, the shift parameter n is the function $n = n_\lambda(j)$ of j. We then have two cases. First, if n_λ is a constant function of j, it derives from Eq. (10) that λ is the null sequence. In this case, the shift parameter is 0 at each resolution level and the DWPT of X consists of an infinite sequence of low-pass filtering. The limit distribution is then derived from Lebesgue's dominated convergence theorem applied to Eqs. (5) and (6). The second case is that of a function n_λ which is not constant with j. For instance, consider the sequence $\lambda = (1, 1, \ldots)$ for which $n_\lambda(j) = 2^j - 1$ so that the nodes of $\mathcal{P}_\lambda$ are $(j, 2^j - 1)$. As mentioned at the end of section 3, when the shift parameter $n = n(j)$ varies with j, Lebesgue's dominated convergence theorem does not apply to Eqs. (5) and (6). The analysis of the statistical behaviour of $c_{j,n(j)}$ when j tends to infinity is described in section 4.3.

4.2 Shannon DWPT and the Paley-Wiener space of π band-limited functions

We start by considering the case where the DWPT is performed via the Shannon DWPT filters. The Shannon filters are hereafter denoted $h^{\mathsf{S}}_\varepsilon$ for $\varepsilon = 0, 1$. These filters are ideal low-pass and high-pass filters. The Fourier transform of $h^{\mathsf{S}}_\varepsilon$ is

$$H^{\mathsf{S}}_\varepsilon(\omega) = \sum_{\ell\in\mathbb{Z}} \mathbb{1}_{\left[-\frac{(\varepsilon+1)\pi}{2}, -\frac{\varepsilon\pi}{2}\right] \cup \left[\frac{\varepsilon\pi}{2}, \frac{(\varepsilon+1)\pi}{2}\right]} (\omega - 2\pi\ell). \tag{11}$$

The scaling function Φ^{S} associated with these filters is $\Phi^{\mathsf{S}}(t) = \mathsf{sinc}(t) = \sin(\pi t)/\pi t$, $t \in \mathbb{R}$, with $\Phi^{\mathsf{S}}(0) = 1$. The Fourier transform of this scaling function is $\mathcal{F}\Phi^{\mathsf{S}} = \mathbb{1}_{[-\pi,\pi]}$, where $\mathbb{1}_K$ denotes the indicator function of a given set K: $\mathbb{1}_K(x) = 1$ if $x \in K$ and $\mathbb{1}_K(x) = 0$, otherwise. The closure $\mathbf{U}^{\mathsf{S}}$ of the space spanned by the orthonormal system $\{\tau_k\Phi^{\mathsf{S}} : k \in \mathbb{Z}\}$ is then the Paley-Wiener (PW) space of those elements of $L^2(\mathbb{R})$ that are π band-limited in the sense that their Fourier transform is supported within $[-\pi, \pi]$. The PW space $\mathbf{U}^{\mathsf{S}}$ is the natural representation space of π band-limited and second-order Wide-Sense Stationary (WSS) random processes (see [4, Appendix D]). Any element of this space satisfies Shannon's sampling theorem. Therefore, the DWPT of any band-limited and second-order WSS random process can be initialized with the samples of this proces. Form now on, the decomposition space is the PW space $\mathbf{U}^{\mathsf{S}}$.

Let us consider the Shannon DWPT of the PW space $\mathbf{U}^{\mathsf{S}}$. The wavelet packet functions $W^{\mathsf{S}}_{j,n}$ of this DWPT can be computed via Eqs. (2), (3), by setting $\Phi = \Phi^{\mathsf{S}}$ and $h_\varepsilon = h^{\mathsf{S}}_\varepsilon$, $\varepsilon = 0, 1$. The Fourier transforms of these wavelet packet functions are given by ([6, Proposition 8.2, p. 328])

$$\mathcal{F}W^{\mathsf{S}}_{j,n} = 2^{j/2}\mathbb{1}_{\left[-\frac{(G(n)+1)\pi}{2^j}, -\frac{G(n)\pi}{2^j}\right] \cup \left[\frac{G(n)\pi}{2^j}, \frac{(G(n)+1)\pi}{2^j}\right]}, \tag{12}$$

for every non-negative integer j and every $n \in \{0, \dots, 2^j - 1\}$. The map G is defined by $G(0) = 0$ and by recursively setting, for $\varepsilon = 0, 1$ and $\ell = 0, 1, 2, \dots$

$$G(2\ell+\varepsilon) = \begin{cases} 2G(\ell)+\varepsilon & \text{if } G(\ell) \text{ is even,} \\ 2G(\ell)-\varepsilon+1 & \text{if } G(\ell) \text{ is odd.} \end{cases} \tag{13}$$

The restriction of G to the set $\{0, 1, \dots, 2^j - 1\}$ is a permutation of this set. This permutation induces a frequency re-ordering of the Shannon wavelet packets $\mathcal{F}W^{\mathsf{S}}_{j,n}$, $n = 0, 1, \dots, 2^j - 1$.

When the wavelet packet functions are the functions $W^{\mathsf{S}}_{j,n}$, it follows from Eqs. (7) and (12) that the cumulant $\mathsf{cum}^{\mathsf{S}}_{j,n}[k_1, k_2, \cdots, k_N]$ of the discrete random process returned at node (j, n) by the Shannon DWPT of X is such that $|\mathsf{cum}^{\mathsf{S}}_{j,n}[k_1, k_2, \cdots, k_N]| \leqslant \|\gamma_N\|_\infty 2^{-j(N-1)/2}$. Given any natural number $N > 1$, the rhs of the latter inequality does not depend on $n, k_1, \dots, k_N$ and vanishes when j tends to ∞. Thus for every natural number $N > 1$, $\mathsf{cum}^{\mathsf{S}}_{j,n}[k_1, k_2, \cdots, k_N]$ tends to zero uniformly in $n, k_1, k_2, \cdots, k_N$, when j tends to infinity. In addition, the autocorrelation function $R^{\mathsf{S}}_{j,n}$ resulting from the projection of X on $\mathbf{W}^{\mathsf{S}}_{j,n}$ derives from Eqs. (8), (12), and is given by (see [4])

$$R^{\mathsf{S}}_{j,n}[k] = \frac{2^j}{\pi}\int_{\left[\frac{G(n)\pi}{2^j}, \frac{(G(n)+1)\pi}{2^j}\right]} \gamma(\omega)\cos(2^j k\omega)d\omega. \tag{14}$$

If the spectrum γ of X is continuous at point $a(\lambda)$, then, given any path $\mathcal{P}_\lambda = (\mathbf{U}^{\mathsf{S}}, \{\mathbf{W}^{\mathsf{S}}_{j,n_\lambda(j)}\}_{j\in\mathbb{N}})$ associated with a binary sequence $\lambda = (\varepsilon_\ell)_{\ell\in\mathbb{N}}$ of elements of $\{0, 1\}$, $\lim_{j\to+\infty} R^{\mathsf{S}}_{j,n_\lambda(j)}[k] = \gamma(a(\lambda))\delta[k]$ uniformly in $k \in \mathbb{Z}$, where

$$a(\lambda) = \lim_{j\to+\infty} G(n_\lambda(j))\pi/2^j. \tag{15}$$

The following result summarizes the foregoing analysis.

Lemma 1 *If γ is continuous at $a(\lambda)$, then, when j tend to infinity, the sequence $\left(c^{\mathsf{S}}_{j,n_\lambda(j)}\right)_j$ converges in distribution to a white Gaussian process with variance $\gamma(a(\lambda))$, the convergence being in the following sense: For every $x \in \mathbb{R}^N$ and every $\eta > 0$, there exists $j_0 = j_0(x, \eta) > 0$ such that, for every $j \geqslant j_0$ the absolute value of the difference between the value at x of the probability distribution of the random vector*

$$(c^{\mathsf{S}}_{j,n_\lambda(j)}[k_1], c^{\mathsf{S}}_{j,n_\lambda(j)}[k_2], \dots, c^{\mathsf{S}}_{j,n_\lambda(j)}[k_N])$$

and the value at x of the centred N-variate normal distribution $\mathcal{N}(0, \gamma(a(\lambda))\mathrm{I}_N)$ with covariance matrix $\gamma(a(\lambda))\mathrm{I}_N$ is less than η.

4.3 Central limit theorems

Lemma 1 concerns ideal DWPT filters. In order to obtain a similar result for filters of more practical interest, the DWPT is now assumed to be performed by using decomposition filters $h^{[r]}_\varepsilon$, $\varepsilon = 0, 1$, that depend on a non-negative integer or real value r such that

$$\lim_{r\to\infty} H^{[r]}_\varepsilon = H^{\mathsf{S}}_\varepsilon \quad \text{(a.e.)}, \tag{16}$$

where $H^{[r]}_\varepsilon$ is the Fourier transform of $h^{[r]}_\varepsilon$ and $H^{\mathsf{S}}_\varepsilon$ is given by Eq. (11); r is called the order of the DWPT filters. When r tends to ∞, the DWPT filters with impulse responses $\{h^{[r]}_\varepsilon\}_{\varepsilon=0,1}$ converge in the sense specified by Eq. (16) to the Shannon DWPT filters $\{h^{\mathsf{S}}_\varepsilon\}_{\varepsilon=0,1}$. On the other hand, Eq. (16) can be regarded as a property of regularity for the following reasons. According to [7], the Daubechies filters satisfy Eq. (16) when r is the number of vanishing moments of the Daubechies wavelet function; according to [8], Battle-Lemarié filters also satisfy Eq. (16) when r is the spline order of the Battle-Lemarié scaling function.

Let us consider decomposition filters satisfying Eq. (16). Let λ be a binary sequence of elements of $\{0, 1\}$. The following result, similar to Lemma 1, describes the asymptotic distribution of the discrete random process $c^{[r]}_{j,n_\lambda(j)}$ returned at node $(j, n_\lambda(j))$ when the resolution level j and the order r of the filters increase. With the same assumptions and notations as those used so far:

Proposition 1 *Assume that γ is continuous at $a(\lambda)$. Then, when j and r tend to infinity, the sequence $\left(c^{[r]}_{j,n_\lambda(j)}\right)_{r,j}$ converges in distribution to a white Gaussian process with variance $\gamma(a(\lambda))$ in the following sense: For every $x \in \mathbb{R}^N$ and every $\eta > 0$, there exists $j_0 = j_0(x, \eta) > 0$ and $r_0 = r_0(x, j_0, \eta)$ such that, for every $j \geqslant j_0$ and every $r \geqslant r_0$, the absolute value of the difference between the value at x of the probability distribution of the random vector*

$$(c^{[r]}_{j,n_\lambda(j)}[k_1], c^{[r]}_{j,n_\lambda(j)}[k_2], \dots, c^{[r]}_{j,n_\lambda(j)}[k_N])$$

and the value at x of the centred N-variate normal distribution $\mathcal{N}(0, \gamma(a(\lambda))\mathrm{I}_N)$ with covariance matrix $\gamma(a(\lambda))\mathrm{I}_N$ is less than η.

Basically, this result is a consequence of the following two facts. For every given natural number j and every $n \in \{0, 1, \dots, 2^j - 1\}$, let $\mathsf{cum}^{[r]}_{j,n}[k_1, k_2, \dots, k_N]$ stand for the cumulant of order $N+1$ of the wavelet packet coefficients of X with respect to the packet $\mathbf{W}^{[r]}_{j,n}$. First, if $N > 1$, we have

$$\lim_{j\to+\infty}\left(\lim_{r\to+\infty} \mathsf{cum}^{[r]}_{j,n_\lambda(j)}[k_1, k_2, \dots, k_N]\right) = 0, \qquad (17)$$

uniformly in $k_1, k_2, \dots, k_N$. Second, if γ is continuous at $a(\lambda)$, then

$$\lim_{j\to+\infty}\left(\lim_{r\to+\infty} R^{[r]}_{j,n_\lambda(j)}[k]\right) = \gamma(a(\lambda))\delta[k], \qquad (18)$$

uniformly in $k \in \mathbb{Z}$, with $a(\lambda)$ given by Eq. (15). Note that the latter statement follows from [4, Theorem 1].

Remark 1 *If $n = n_\lambda(j)$ is a constant function of j, the behaviour of $\mathsf{cum}^{[r]}_{j,n}[k_1, k_2, \cdots, k_N]$ when j tends to ∞ straightforwardly derives from Lebesgue's dominated convergence theorem applied to Eqs.* (5), (6). *However, Eqs.* (17),(18) *suggest that r may play a role in the convergence of the cumulant. The experimental results of the next section highlights that this convergence actually accelerates when r increases.*

5. EXPERIMENTAL RESULTS

Since Proposition 1 is asymptotic, our purpose is to experimentally study how well the tendency to Gaussianity is satisfied when the input process is non-Gaussian and the DWPT is performed with finite values for the resolution level and the order of decomposition filters. The Daubechies filters are used to perform the DWPT. They converge to the Shannon filters when the number r of vanishing moments of the Daubechies mother wavelet increases. As above, $X(t)$ stands for the centred Hilbertian random process to decompose. In our experiments, $X(t)$ is Generalized Gaussian (GG). This means that, for every $t \in \mathbb{R}$, $X(t)$ follows the Generalized Gaussian Distribution (GGD) with *scale* α, *shape* β and zero mean. For each $t \in \mathbb{R}$, the Probability Density Function (PDF) of $X(t)$ is $f_{\alpha,\beta}$ defined for every real value x by $f_{\alpha,\beta}(x) = \frac{\beta}{2\alpha\Gamma(1/\beta)} \exp\left(-(|x|/\alpha)^\beta\right)$ where Γ is the standard Gamma function. The value of the GGD standard deviation is $\sigma = \alpha\sqrt{\Gamma(3/\beta)/\Gamma(1/\beta)}$. In what follows, $\alpha = \sqrt{\Gamma(1/\beta)/\Gamma(3/\beta)}$ so that $\sigma = 1$. When the shape parameter β equals 2, the GGD is Gaussian; when β decreases (from 2 to 0), the PDF of the GGD is sharper, and sharper (see figure 1); when $\beta = 1$, the GGD is the Laplacian distribution. Moreover, in our experiments, the samples $X(1)$, $X(2)$, $\dots$, $X(N)$ of the GG process $X(t)$ are correlated. In fact, these samples are synthesized by filtering a discrete sequence of independent and identically GG distributed random variables through an auto-regressive (AR) filter of order 1, and such that the spectrum of $X(t)$ is $\gamma(\omega) = (1-\mu)^2/|1 - \mu\exp(-i\omega)|^2$ where $0 < \mu < 1$. If α and β are the parameters of the GG random variables used to synthesize the samples of $X(t)$, we henceforth say that the output discrete process $X(t)$ is AR(1)-GG. Experimental tests are carried out with $\mu = 0.5$, 0.75, 0.9, 0.95. The spectra corresponding to these values are those of figure 1.

The experiments are conducted with 100 independent random copies of the random vector formed by the N samples $X(1)$, $X(2)$, $\dots$, $X(N)$ with $N = 2^{20}$. Each copy is used as an input of the DWPT. We then consider the four wavelet packet paths associated with the sequences $\lambda_q = (\delta[q-\ell])_{\ell\in\mathbb{N}}$, for $q = 0, 1, 2$ and 3. For these sequences, and taking into account Eq. (9), we have $n_{\lambda_0}(\ell) = 0$ for every natural number

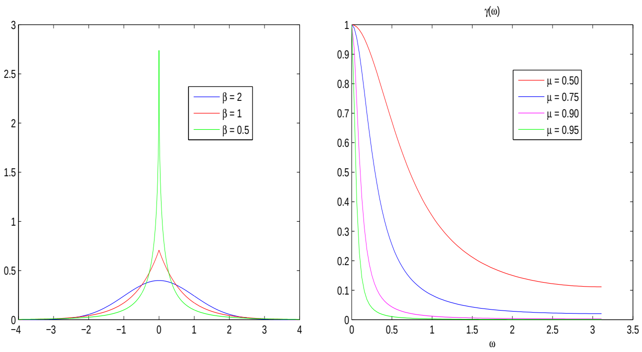

Figure 1: Left: GGD with unit variance and shape $\beta = 2, 1, 0.5$. Right: coloured GG Spectrum for some values of μ.

ℓ, and for $q = 1, 2, 3$:

$$n_{\lambda_q}(\ell) = \begin{cases} 0 & \text{for} \quad \ell = 1, 2, \cdots, q-1, \\ 2^{\ell-q} & \text{for} \quad \ell = q, q+1, \cdots. \end{cases}$$

It follows that $G(n_{\lambda_0}(\ell)) = 0$ and that

$$G(n_{\lambda_q}(\ell)) = \begin{cases} 0 & \text{for} \quad \ell = 1, 2, \cdots, q-1, \\ 2^{\ell-q+1} - 1 & \text{for} \quad \ell = q, q+1, \cdots, \end{cases}$$

for $q = 1, 2, 3$. According to Eq. (15), $a(\lambda_0) = 0$ and $a(\lambda_q) = \pi/2^{q-1}$ for $q = 1, 2$ and 3. Table 1 gives the values $\gamma(a(\lambda))$ for the four test sequences. For every path λ among those introduced above, the Kolmogorov-Smirnov (KS) test with significant level 5% is used to decide whether the samples $(c^{[r]}_{j,n_\lambda(j)}[k]/\sqrt{\gamma(a(\lambda))})_k$, returned by the DWPT for a given copy, satisfy the null hypothesis (that is, follow the normal distribution $\mathcal{N}(0,1)$), or not (alternative hypothesis). The KS acceptance rates obtained are presented in table 2. By increasing the resolution level j when the order of the filters is constant and equals $r = 1$, the KS acceptance rate increases for most of the DWPT paths. When the resolution is fixed to $j = 6$, it suffices to increase the order r to also increase this acceptance rate. For the sequences under consideration and for AR(1)-GG processes with $1 \leqslant \beta \leqslant 2$ and $0 < \mu < 0.9$, normality can reasonably be considered to be attained when the resolution level j is 6 and the order of the Daubechies filters is $r = 7$. The less satisfactory results occur for large values of μ or small values of β. When μ is large, the spectrum becomes rather sharp around the null frequency (see figure 1 for $\mu = 0.90, 0.95$); on the other hand, when β is small, the PDF of the GGD is still sharper at the origin (see figure 1 for $\beta = 0.5$). However, even for large values of μ and small values of β, increasing both the order of the filters and the resolution level improves the KS acceptance rate (see table 3).

As an illustration, figure 2 shows histograms of the DWPT coefficients obtained at resolution level 6, by using Daubechies filters of order 7. The decomposition concerns the samples of an AR(1)-GG process with $\beta = 1$ and $\mu = 0.75$. These histograms are compared with the PDF of the Gaussian limit distribution.

It follows from the results above that significant acceptance rates are attained by increasing first the resolution level and then the order of the filters. This confirms the theoretical results. The order of the filters speeds up the convergence to normality even for the path associated with the null sequence λ_0 (see remark 1). This can be noticed by comparing, at resolution level $j = 6$, the acceptance rates obtained for $r = 1$ to those obtained for $r = 7$ in tables 2 and 3 for $\mathcal{P}_{\lambda_0}$.

Table 1: Values $\gamma(a(\lambda))$ for the four test sequences.

Path	$\mu = 0.5$	$\mu = 0.75$	$\mu = 0.9$	$\mu = 0.95$
$\mathcal{P}_{\lambda_0}$	1	1	1	1
$\mathcal{P}_{\lambda_1}$	0.1111	0.0204	0.0028	0.0007
$\mathcal{P}_{\lambda_2}$	0.2052	0.0412	0.0057	0.0014
$\mathcal{P}_{\lambda_3}$	0.4798	0.1332	0.0201	0.0048

Table 2: KS test acceptance rates for the normal distribution $\mathcal{N}(0,1)$ of the DWPT coefficients returned at resolution level $j = 3, 6$ for different DWPT paths. The DWPT input process is AR(1)-GG with α such that $\sigma = 1$.

Coloured GG process, with $\beta = 1.5$									
	$\mu = 0.5$			$\mu = 0.75$			$\mu = 0.9$		
	$j = 3$	$j = 6$	$j = 6$	$j = 3$	$j = 6$	$j = 6$	$j = 3$	$j = 6$	$j = 6$
Path	$r = 1$	$r = 1$	$r = 7$	$r = 1$	$r = 1$	$r = 7$	$r = 1$	$r = 1$	$r = 7$
$\mathcal{P}_{\lambda_0}$	0%	95%	98%	0%	42%	99%	0%	0%	19%
$\mathcal{P}_{\lambda_1}$	0%	91%	98%	0%	52%	96%	0%	0%	94%
$\mathcal{P}_{\lambda_2}$	0%	95%	88%	0%	37%	86%	0%	0%	91%
$\mathcal{P}_{\lambda_3}$	0%	0%	86%	0%	14%	65%	0%	0%	53%
Coloured GG process, with $\beta = 1$									
	$\mu = 0.5$			$\mu = 0.75$			$\mu = 0.9$		
	$j = 3$	$j = 6$	$j = 6$	$j = 3$	$j = 6$	$j = 6$	$j = 3$	$j = 6$	$j = 6$
Path	$r = 1$	$r = 1$	$r = 7$	$r = 1$	$r = 1$	$r = 7$	$r = 1$	$r = 1$	$r = 7$
$\mathcal{P}_{\lambda_0}$	0%	84%	94%	0%	31%	96%	0%	0%	21%
$\mathcal{P}_{\lambda_1}$	0%	94%	96%	0%	67%	93%	0%	0%	92%
$\mathcal{P}_{\lambda_2}$	0%	95%	82%	0%	56%	78%	0%	0%	89%
$\mathcal{P}_{\lambda_3}$	0%	0%	71%	0%	3%	50%	0%	0%	41%

Table 3: KS test acceptance rates for the normal distribution $\mathcal{N}(0,1)$ of the DWPT coefficients returned at resolution level $j = 6, 7$ for different DWPT paths. The DWPT input process is AR(1)-GG with α such that $\sigma = 1$.

Coloured GG process, with $\beta = 0.5$, $\mu = 0.95$			
	$j = 6$	$j = 6$	$j = 7$
Path	$r = 7$	$r = 20$	$r = 20$
$\mathcal{P}_{\lambda_0}$	0%	0%	45%
$\mathcal{P}_{\lambda_1}$	78%	86%	91%
$\mathcal{P}_{\lambda_2}$	2%	30%	79%
$\mathcal{P}_{\lambda_3}$	0%	4%	35%

6. CONCLUSION

In this paper, the tendency to normality of the wavelet packet coefficients of a strictly stationary random process has been studied. We have considered DWPT filters whose Fourier transforms converge almost everywhere to the Fourier transform of the Shannon filters. This type of filters makes it possible to state results that are valid for any path of the DWPT. Daubechies and Battle-Lemarié filters are examples of such filters. The asymptotic distribution of the wavelet packet coefficients is normal with variance equal to the value taken by the input process spectrum at some specific frequency. This frequency can be computed with respect to the nested supports of the Fourier transforms of the wavelet packets associated with the chosen path. The results of this paper may thus be applicable to several signal processing fields, data analysis or communication applications. Detailed proofs and comments of the results presented above will be given in a forthcoming paper in the general framework of M-band wavelet packet transforms. A preliminary version of this paper is downloadable from `http://fr.arxiv.org/abs/0802.0797`.

REFERENCES

[1] D. Donoho and I. Johnstone, "Ideal spatial adaptation by wavelet shrinkage," *Biometrica*, vol. 81, no. 3, pp. 425–455, Aug. 1994.

[2] I. Johnstone, "Wavelets and the theory of non-parametric function estimation," *Journal of the Royal Statistical Society*, vol. A, no. 357, pp. 2475–2493, 1999.

[3] D. Leporini and J. Pesquet, "High-order wavelet packets and cumulant field analysis," *IEEE Transactions on Information Theory*, vol. 45, no. 3, pp. 863–877, Apr. 1999.

[4] A. 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.

[5] S. Touati and J. 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.

[6] S. Mallat, *A wavelet tour of signal processing, second edition.* Academic Press, 1999.

[7] J. Shen and G. Strang, "Asymptotics of daubechies filters, scaling functions, and wavelets," *Applied and Computational Harmonic Analysis*, vol. 5, no. HA970234, pp. 312+, 1998.

[8] A. Aldroubi, M. Unser, and M. Eden, "Cardinal spline filters: Stability and convergence to the ideal sinc interpolator," *Signal Process*, vol. 28, no. 8, pp. 127–138, Aug. 1992.

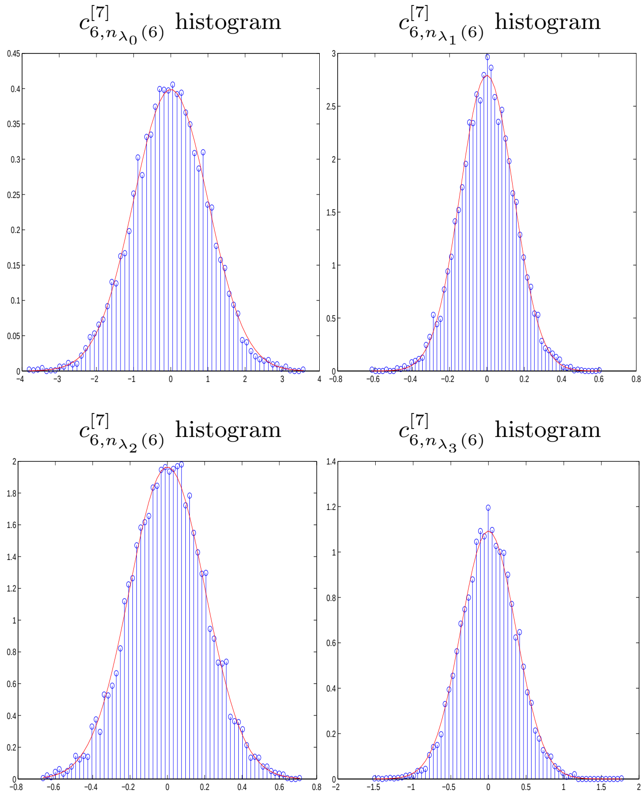

Figure 2: Histograms of the DWPT coefficients, at resolution level 6, using Daubechies filters of order 7. The decomposition is applied to samples of an AR(1)-GG process with $\beta = 1$ and $\mu = 0.75$. The limit distribution $\mathcal{N}(0, \gamma(a(\lambda)))$ where $a(\lambda_0) = 0$ and $a(\lambda_q) = \pi/2^{q-1}$ for $q = 1, 2$ and 3 is represented in (red) continuous line.