Decorrelation of wavelet coefficients for long-range dependent processes

Jan Mielniczuk and Piotr Wojdyllo

***Abstract*— We consider a discrete time stationary long-range dependent process $(X_k)_{k\in Z}$ such that its spectral density equals $\varphi(|\lambda|)^{-2d}$, where φ is a smooth function such that $\varphi(0) = \varphi''(0) = 0$ and $\varphi(\lambda) \geq c\lambda$ for $\lambda \in [0, \pi]$. Then for any wavelet ψ with N vanishing moments the lag k within-level covariance of wavelet coefficients decays as $\mathcal{O}(k^{2d-2N-1})$ when $k \to \infty$. The result applies to Fractionally Integrated ARMA processes as well as to fractional Gaussian noise.**

***Index Terms*— Decorrelation, long-range dependence, spectral density, wavelet coefficients.**

I. Introduction

Let $(X_k)_{k\in\mathbf{Z}}$ be a discrete time mean zero stationary time series such that its second moments are finite and denote by f its spectral density defined as $r(k) := \operatorname{cov}(X_1, X_{1+k}) = \int_{-\pi}^{\pi} f(\lambda)e^{ik\lambda}\, d\lambda$. In the paper we focus on the case when $(X_k)_{k\in\mathbf{Z}}$ is long-range dependent i.e. $f(\cdot)$ has a pole of order $0 < \gamma < 1$ at 0:

$$f(\lambda) \sim C_\gamma \lambda^{-\gamma} \qquad \text{when } \lambda \to 0, \tag{1}$$

where $a(\lambda) \sim b(\lambda)$ means that the ratio $a(\lambda)/b(\lambda)$ tends to 1.

The equivalence (1) translates into a condition on the covariance function $r(k) \sim C_r k^{\gamma-1}$ when $k \to \infty$ under mild assumptions on $r(\cdot)$ (cf. [19], Theorem III-14). Observe that this implies $\sum_{k=1}^{\infty} |r(k)| = \infty$, which is another commonly adopted definition of long-range dependence.

There has been a surge of interest in long-range dependence over the last fifteen years due mainly to the fact that the property (1) is believed to hold for time series commonly occurring in empirical studies e.g. in economics and finance, see e.g. [3], [8], [14]. Moreover, it turns out that the behaviour of estimators based on such sequences is qualitatively different from those based on independent or weakly dependent ones, e.g. Central Limit Theorem with the usual $\sqrt{n}$ standardization frequently fails to hold. Therefore, investigation of such estimators usually requires different tools than those developed for the weak dependence. In this context it appears to be potentially useful to observe that despite the strong dependence, the wavelet coefficients are usually *weakly* dependent within a fixed level resolution level. This, among others, greatly facilitates studying properties of wavelet-based estimators for long-range dependent data (we refer to Chapter 9 of [13] for discussion of whitening effect of wavelet transform). As a typical example we mention a wavelet estimator of γ based on a wavelet spectrum [1], see discussion section for more details. The frequently stated result for a continuous time process (cf e.g. [2], formula (28)) is as follows. Let ψ be a wavelet with N vanishing moments i.e. $\int_{\mathbf{R}} s^i \psi(s)\, ds = 0$ for $i = 0, 1, \ldots, N-1$ and let $d_{j,k}$ be the wavelet coefficient at resolution j and location k defined in the next section. Then for any $j \in \mathbf{Z}$

$$r_j(k-k') = \operatorname{Cov}(d_{jk}, d_{jk'}) = O(|k-k'|^{\gamma-2N-1}), \tag{2}$$

when $|k-k'| \to \infty$. Heuristic justification of (2) is based on the observation that N zero moments imply that that the Fourier transform $\hat{\psi}$ of ψ is $\mathcal{O}(|\lambda|^N)$ and the behaviour of the covariance is governed by $|\hat{\psi}(\lambda)|^2|\lambda|^{-\gamma}$ at the origin. Thus the pole of $|\lambda|^{-\gamma}$ is balanced by the regularity $|\lambda|^{2N}$ contributed by the wavelet, resulting in the order of $\mathcal{O}(|\lambda|^{2N-\gamma})$ at 0. This in turn suggests that (2) holds. This line of reasoning turns out to be hard to formalize and to our knowledge no formal proofs of (2) for a general class of long-range dependent processes exist. However, there are results on closely related problems for specific processes. In particular, [9] and [16] proved that the covariance function at lag k for the fractional Brownian motion is $\mathcal{O}(|k|^{\gamma-2N+1})$ and Mielniczuk and Wojdyllo [12] extending Tewfik and Kim's reasoning in [16] proved that (2) indeed holds for the continuous time fractional Gaussian noise (fGn). Also the analogous result was established there for FARIMA$(0, d, 0)$ processes.

In the correspondence we provide a rigorous justification of (2) for a wide class of discrete time long-range dependent stationary sequences satisfying (1) and for a general wavelet having N vanishing moments. More generally, it is sufficient to assume that f equals $L(|\lambda|)\varphi(|\lambda|)^{-2d}$, where L and φ are smooth functions such that $\varphi(0) = \varphi''(0) = 0$ and $\varphi(\lambda) \geq c\lambda$ for $\lambda \in [0, \pi]$.

In particular the result holds for FARIMA(p, d, q) process and the fGn process filtered by an arbitrary smooth filter. We focus here on within-scale decorrelation property; for some results on between-scales decorrelation, see [5].

II. Preliminaries

Let $\psi \in \mathcal{L}^2(\mathbf{R})$ be a real-valued wavelet pertaining to a multiresolution analysis i.e. a function such that $\int \psi(s)\, ds = 0$ and a family $\psi_{jk}(t) = 2^{-j/2}\psi(2^{-j}t - k)$, $j, k \in \mathbf{Z}$ of its rescaled and translated versions constitutes an orthonormal basis in $\mathcal{L}^2(\mathbf{R})$ (see [6], [18]). Then there exists a scaling function $\phi(\cdot)$ pertaining to ψ (cf. [17], Section 3.3) with $\phi_{jk}(t) = 2^{-j/2}\phi(2^{-j}t - k)$. For such a function, in particular,

Manuscript received December 7, 2005. Revised January 15, 2007.

J. Mielniczuk is with Institute of Computer Science, Polish Academy of Sciences, Ordona 21, 01-237 Warszawa, POLAND. E-mail: miel@ipipan.waw.pl.

P. Wojdyllo is with Institute of Mathematics, Polish Academy of Sciences, Sniadeckich 8, 00-956 Warszawa 10, PO Box 21, POLAND. E-mail: pwoj@impan.gov.pl.

$\{\phi_{jk}\}_{k\in\mathbf{Z}}$ is an orthonormal system for any $j \in \mathbf{Z}$ and, moreover, $S_j \bigoplus W_j = S_{j-1}$, where W_j (resp. S_j) is a closure of a linear subspace spanned by $\{\psi_{jk}\}_{k\in\mathbf{Z}}$ (resp. $\{\phi_{jk}\}_{k\in\mathbf{Z}}$). We denote by v_k the coefficients of the function $\psi_{1,0} \in W_1 \subset S_0$ with respect to the orthonormal basis $\{\phi_{0,k}\}$ of S_0

$$v_k = 2^{-1/2} \int \psi(t/2)\phi(t-k)\,dt, \qquad k \in \mathbf{Z}, \tag{3}$$

and analogously for $\phi_{1,0}$

$$u_k = 2^{-1/2} \int \phi(t/2)\phi(t-k)\,dt, \qquad k \in \mathbf{Z}. \tag{4}$$

Let $(X_k)_{k\in\mathbf{Z}}$ be an arbitrary mean zero stationary time series. Define the following embedding $\tilde{X}_t = \sum_n X_n\phi(t-n)$ for $t \in \mathbf{R}$, which can be considered a continuous time approximation of X_n. In the simplest case of the Haar scaling function $\phi(s) = I_{[0,1)}(s)$, $\tilde{X}_t$ is a piecewise constant approximation of X_n: $\tilde{X}_t = X_n$ for $t \in [n, n+1)$. Let $a_{j,k} = \int \tilde{X}_t\phi_{j,k}(t)\,dt$ and $d_{j,k} = \int \tilde{X}_t\psi_{j,k}(t)\,dt$. As $\{\psi_{jk}\}_{j,k\in\mathbf{Z}}$ form a basis in $L^2(\mathbf{R})$ coefficients $d_{j,k}$ contain all information about the sample path of (X_n). Thus, it is of importance to study their stochastic properties. It is easily seen that coefficients $d_{j,k}$ can be alternatively defined through the following recursive relations. Let $a_{0,n} = X_n$ and

$$a_{j,k} = \sum_{n\in\mathbf{Z}} u_n a_{j-1,2k+n} \tag{5}$$

$$d_{j,k} = \sum_{n\in\mathbf{Z}} v_n a_{j-1,2k+n} \tag{6}$$

The objective of the note is to study approximate decorrelation of wavelet coefficients $d_{j,k}$. It can be easily proved that $(d_{j,\cdot})$ and $(a_{j,\cdot})$ form stationary sequences for each $j \in \mathbf{Z}$. Denote their spectral densities by f_j^d and f_j^a respectively; the same notation is used for their 2π- extension to the whole line. As it is seen from (5)-(6) the recursive scheme of calculating $d_{j,k}$ consists solely in filtering and decimation (cf e.g. [13]), it follows that $f_0^a = f$ and for any $i \geq 1$

$$f_i^a(\lambda) = \frac{1}{2}(U(\lambda/2)f_{i-1}^a(\lambda/2) + U(\lambda/2+\pi)f_{i-1}^a(\lambda/2+\pi))$$

while

$$f_i^d(\lambda) = \frac{1}{2}(V(\lambda/2)f_{i-1}^a(\lambda/2) + V(\lambda/2+\pi)f_{i-1}^a(\lambda/2+\pi)),$$

where $V(\lambda) = |\sum_k v_k e^{ik\lambda}|^2$ and $U(\lambda) = |\sum_k u_k e^{ik\lambda}|^2$. Moreover ([13], equation (348b))

$$\int_{-\pi}^{\pi} f_j^d(\lambda)e^{i\lambda l}\,d\lambda = \int_{-\pi}^{\pi} f(\lambda)V_j(\lambda)e^{i2^j\lambda l}\,d\lambda \tag{7}$$

where $V_j(\lambda) = V(2^{j-1}\lambda)\prod_{l=0}^{j-2} U(2^l\lambda)$.

In the following we consider two main examples of LRD processes.

(i) Fractionally Integrated Moving Average FARIMA$(0,d,0)$ with $0 < d < 1/2$ is a process (X_n) such that $(1-B)^d X_n = \varepsilon_n$, where $(\varepsilon_n)_{-\infty}^{\infty}$ is i.i.d. Gaussian $N(0,\sigma^2)$ sequence. Moreover, $(1-B)^d$ is the fractional differencing operator defined by $(1-B)^d = \sum_{k=0}^{\infty}\binom{d}{k}(-1)^k B^k$, where $\binom{d}{k} = \Gamma(d+1)/(\Gamma(k+1)\Gamma(d-k+1))$, $\Gamma(\cdot)$ is the gamma function and $B^k X_n = X_{n-k}$. A general FARIMA(p,d,q) process $\bar{X}_n$ is a process such that $\bar{X}_n = (1-B)^d X_n$ is ARMA(p,q) i.e. $P(B)\bar{X}_n = Q(B)\varepsilon_n$, where $P(s) = 1 - a_0 s - \cdots - a_p s^p$ and $Q(s) = 1 - b_0 s - \cdots - b_q s^q$, $p, q \in \mathbf{N}$. It follows that if $P(\cdot)$ has no roots in the complex unit circle than the spectral density of FARIMA(p,d,q) is

$$f(\lambda) = \sigma^2(2\sin(|\lambda|/2))^{-2d}\frac{|Q(e^{i\lambda})|^2}{|P(e^{i\lambda})|^2}. \tag{8}$$

In particular, $f(\lambda) = \sigma^2(2\sin(|\lambda|/2))^{-2d}$ for FARIMA$(0,d,0)$. Thus for such processes γ in (1) equals $2d$ and the specific form of the spectral density of the FARIMA$(0,d,0)$ naturally extends to the form $f(\lambda) = L(\lambda)\varphi(|\lambda|)^{-2d}$ considered in Theorem 3.1 by choosing $\varphi(\lambda) = 2\sigma^{-1/d}\sin\left(\frac{\lambda}{2}\right)$.

(ii) Fractional Gaussian Noise (fGn) X_n is defined as the first order difference of the fractional Brownian motion B_t^H i.e. the H- self-similar Gaussian process with stationary increments $X_n = B_{n+1}^H - B_n^H$. It follows that ([15], Proposition 7.2.9) for $1/2 < H < 1$, (X_n) is long-range dependent and its spectral density equals

$$f(\lambda) = \sigma^2 C(H)\sin^2(|\lambda|/2)\Big(\sum_{k\in Z}\frac{1}{|\lambda+2\pi k|^{2H+1}}\Big) \tag{9}$$

Thus it follows that $f(\lambda)$ satisfies (1) with $\gamma = 2H-1$.

III. Results

Our main result is

Theorem 3.1: Let φ be a smooth function such that $\varphi(0) = \varphi''(0) = 0$ and $\varphi(\lambda) \geq c\lambda$ for some $c > 0$ and $\lambda \in [0,\pi]$. Assume the squared gain function V is $2N+1$ times differentiable, $V^{(i)}(0) = 0$ for $0 \leq i \leq 2N-1$ and $V^{(2N+1)}(\cdot)$ is bounded. Let $(X_k)_{k\in Z}$ be a mean zero stationary process with the spectral density given by $f(\lambda) = \varphi(|\lambda|)^{-2d}$ for $\lambda \in [-\pi,\pi]$. Then for $j \geq 1$ the lag-k covariance $r_j(k)$ of its wavelet coefficients $d_{j,k}$ pertaining to the wavelet filter with the squared gain function $V(\lambda)$ satisfies $r_j(k) = \mathcal{O}(k^{-2N+2d-1})$ when $k \to \infty$.

Observe that the assumption $\varphi(\lambda) \geq c\lambda$ for $\lambda \in [0,\pi]$ implies in particular that $\varphi'(0) > 0$ and thus $f(\lambda) \sim \lambda^{-2d}$ for $\lambda \to 0$. Moreover, for a concave φ the assumption follows from $\varphi(\pi) > 0$. The result allows for the following generalization.

Theorem 3.2: Assume that L is a $2N$ times differentiable function with the bounded derivatives and the spectral density of the process $(X_k)_{k\in Z}$ is given by $f(\lambda) = \varphi(|\lambda|)^{-2d}L(|\lambda|)$ for $\lambda \in [-\pi,\pi]$. Then Theorem 3.1 holds provided its remaining assumptions are valid.

Note that the only case of (1) not covered by Theorem 3.2 is the situation when for all φ fulfilling the requirements of Theorem 3.1 $L(\lambda) := f(\lambda)/\varphi(\lambda)^{-\gamma}$ does not satisfy its assumptions. Also in the theorems above it is assumed that the stationary process has mean zero. However, as ψ has N vanishing moments the results hold true for any process (X_k) being a sum of a stationary process (X_k^0) and a polynomial trend of degree less or equal $N-1$ when f is the spectral density of (X_k^0) (cf [13], Section 9.4).

Corollary 3.3: If compactly supported wavelet ψ has N vanishing moments and f is as in Theorem 3.2, then Theorem 3.1 holds for the wavelet coefficients $d_{j,k}$ pertaining to ψ.

Proof: Denote $v(\lambda) = \sum_{j=0}^{K-1} v_j e^{ij\lambda}$. A simple inductive argument (cf. [17] p. 83) shows that vanishing of first N moments of ψ is equivalent to the condition $v^{(i)}(0) = 0$ for $0 \le i \le N-1$. Also since ψ is compactly supported, ϕ also is and only a finite number of v_i, $i = 0, 1, \ldots, K-1$ is nonzero. As functions v and V are trigonometric polynomials, they have derivatives of all orders.

The condition $V^{(i)}(0) = 0$, $i = 0, 1, \ldots, 2N-1$ is implied by

$$v^{(i)}(0) = 0, \text{ for } \ i = 0, 1, \ldots, N-1. \tag{10}$$

This follows by noting that (10) translates to $\sum_{k=0}^{K-1} k^j v_k = 0$ for $0 \le j \le N-1$. Thus

$$\sum_{0 \le k,l \le K-1} (k-l)^j v_k v_l = \sum_{s=0}^{j} \binom{j}{s} \sum_{k=0}^{K-1} k^s v_k \sum_{l=0}^{K-1} l^{j-s} v_l = 0$$

for $0 \le j \le 2N-1$, which is equivalent to $V^{(j)}(0) = 0$ for $0 \le j \le 2N-1$. Hence, the assumptions of Theorem 3.1 are fulfilled and by Theorem 3.2 the assertion follows. ■

Corollary 3.4: Theorem 3.2 holds true for Daubechies wavelet and for the least asymmetric wavelet with the support of size $2N-1$ as well as for the coiflet with the support of size $3N-1$ for even N.

Proof: As all the wavelets listed above are compactly supported and have N vanishing moments (cf. pp. 168, 254-255, 259 in [6], respectively), the assertion follows by Corollary 3.3. ■

Corollary 3.5: The result of Theorem 3.2 holds for the following processes: (i) $FARIMA(p,d,q)$ process with $0 < d < 1/2$ provided roots of $P(\cdot)$ are not contained in the unit circle; (ii) fGn process; (iii) any process of the form $Y_k = \sum_{i=0}^{\infty} a_i X_{k-i}$ if (X_k) satisfies the assumptions of Theorem 3.2 and $A(\lambda) = |\sum_k a_k e^{ik\lambda}|^2$ satisfies assumption of the Theorem 3.2 for function $L(\cdot)$; (iv) process (Y_k), where $Y_k = X_k + Z_k$, where X_k and Z_k are independent, (X_k) is as in (iii) and (Z_k) is a short-range dependent process such that its spectral density is $2N+2$ times differentiable and its derivative of order $2N+2$ is absolutely integrable. Moreover, if (Z_k) satisfies the assumptions of Theorem 3.2 with the spectral density $\tilde{\varphi}(|\lambda|)^{-2d'}$, where $d' \le d$, the result still holds.

Proof: Only parts (ii) and (iv) need justification, (i) and (iii) follow directly from Theorem 3.2. It follows from (9) that the spectral density of the fGn process can be written as $\sigma^2 C(H)|\lambda|^{-2d} L(|\lambda|)$, where for $\lambda \in [0, \pi]$

$$L(\lambda) = \sum_{k \in Z} \frac{\sin^2(\lambda/2)\lambda^{2d}}{|\lambda + 2\pi k|^{2d+2}}$$

and $d = H - 1/2$. It is trivial to check that the summands in definition of $L(\lambda)$ for $k \ne 0$ are differentiable arbitrary number of times and that their derivatives of any order s are absolutely summable. Moreover, it is routinely checked that the summand for $k = 0$ which equals $\sin^2(|\lambda|/2)/|\lambda|^2$ has bounded derivatives of all orders in $(0, \pi]$ (cf. Remark 5.4(iii)). Indeed, this easily follows from equality $2 \sin \lambda\pi/\lambda = \int_{-\pi}^{\pi} \cos \lambda s \, ds$. Part (iv) follows by noting that that the spectral density of (Y_k) is the sum of the spectral densities of X and Z processes and order of decorrelation of wavelet coefficients of Z process is $\mathcal{O}(k^{-2N-2})$ due to Lemma 5.2(b). ■

IV. Discussion

Decorrelation property stated in Theorems 3.1 and 3.2 may be put to two uses. The first one is simulation of long-range dependent processes, as discussed e.g. in [10]. The second one concerns statistical inference for such processes and is briefly considered here. One of the vital problems in this field of study is estimation of exponent γ of the spectral density at 0 or, equivalently, its Hurst exponent $H = (\gamma+1)/2$. For review and comparison of performance of various estimators of H see e.g. [11]. One of the most promising proposals is the wavelet estimator defined in [1] which is based on the asymptotic equivalence $\mu_j = \mathbf{E} d_{jk}^2 \sim 2^{j\gamma} C_\gamma \int |\lambda|^{-\gamma} |\hat{\psi}(\lambda)|^2 \, d\lambda$ when $\lambda \to 0$ for f satisfying (1) and with $\hat{\psi}$ denoting the Fourier transform of ψ. Thus defining $\hat{\mu}_j$ as an empirical mean of squared wavelet coefficients at the level j, estimator of γ may be constructed by regressing $\log_2 \hat{\mu}_j$ on j. This reasoning, with some refinements taking care of bias term and heteroscedasticity of errors, leads to the wavelet estimator of γ. Its properties are studied in [1] under simplifying assumption that within-scale wavelet coefficients are independent. It is conjectured that Theorems 3.1 and 3.2 can be used to prove properties of the wavelet estimator under more realistic assumptions of their weak dependence. Moreover, Corollary 3.5 lists the long-range processes for which approach in [1] is legitimate. The same observation concerns other estimators of γ which use decorrelation property, see e.g. [5] and [20]. In particular, in section 9.3 of [13] estimation of parameters d and σ^2 based on a sample from Farima$(0, d, 0)$ Gaussian process is discussed. Instead of considering maximum likelihood estimators, maximizers of an approximate version of the likelihood are found. More specifically, the idea is to replace the covariance matrix Σ in the likelihood by its approximation $W^T \Lambda W$, where W is Discrete Wavelet Transform (DWT) matrix, W^T is the transposed matrix W and Λ is the covariance matrix of the wavelet coefficients. In view of decorrelation property of within and between-scale coefficients, off-diagonal terms in Λ are disregarded.

In order to appreciate yet another application of decorrelation property consider estimation of regression function m based on observations $y_i = m(t_i) + \varepsilon_i$, where $t_i = 1, \ldots, n$ are fixed points and ε_i are zero mean errors with variance σ^2. In [7] a wavelet shrinkage estimator of $m(\cdot)$ is discussed when (ε_i) is the Gaussian noise. The idea is to use the wavelet representation of m with plugged-in estimated coefficients $\hat{d}_{j,k}$ instead of $d_{j,k}$. The terms pertaining to small $\hat{d}_{j,k}$ are omitted. In case of hard thresholding only terms satisfying $|\hat{d}_{j,k}| > \delta$ for some threshold δ are retained. The proposed form of the threshold $\delta = (2 \log n)^{1/2} \sigma$ is based on the result from extreme value theory applied to wavelet coefficients which states that $(\sqrt{2 \log n}\, \sigma)^{-1} \max_{i=1,\ldots,n} Z_i \to 1$ in probability for Gaussian i.i.d. innovations Z_i. We conjecture that in view of the decorrelation property this type of threshold should also be applicable for long-range dependent Gaussian errors. Indeed, Theorem 9.2.2 in [4] states that the above result on $\max_{i=1,\ldots,n} Z_i$ holds provided (Z_i) is a Gaussian stationary sequence such that $r(i) \log n \to 0$.

V. Proofs

We begin with three auxiliary lemmas. All functions are defined on $[-\pi,\pi]$, whenever needed their 2π-periodic extension to $\mathbf{R}$ is considered. C will denote a generic constant, a value of which may change from line to line.

Lemma 5.1: Let $g(\lambda)$ be an s times differentiable function on the interval $[-\pi,\pi]$. The sth derivative of $g(\lambda)^\alpha$ equals to

$$[g(\lambda)^\alpha]^{(s)} = \sum_{(i_j)} \prod_{j=0}^{s} c_j \left[g^{(j)}(\lambda)\right]^{i_j},$$

where the summation extends over all $(s+1)$-dimensional multiindices (i_j) such that $\sum_j i_j = \alpha$, $\sum_j j i_j = s$ and c_j are some constants not depending on g.

Proof: We proceed by induction on s. For $s=1$ the assertion holds with $i_0 = \alpha-1$ and $i_1 = 1$. Denote by $[i_0 \ i_1 \ i_2 \ldots]$ the term $\prod_{j=0}^{s}\left[g^{(j)}(\lambda)\right]^{i_j}$ pertaining to the multiindex $I = (i_0 \ i_1 \ i_2 \ldots)$. The induction step follows from observation that

$$[i_0 \ i_1 \ i_2 \ \ldots]' = \sum_{j: i_j \neq 0} i_j \left[D_j\left(i_0 \ i_1 \ i_2 \ \ldots\right)\right],$$

where

$$D_j\left(i_0 \, i_1 \, \ldots \, i_j \, i_{j+1} \, \ldots\right) = \left(i_0 \, i_1 \, \ldots \, i_j - 1 \, i_{j+1}+1 \, \ldots\right).$$

Moreover, if $I' = D_j I$, where $\sum_j j I_j = s$, then $\sum_j j I'_j = s+1$. ■

Lemma 5.2: If a function $A(\lambda)$ is $2N$ times differentiable on $[-\pi,\pi]$ then the cosine Fourier coefficients of $B(\lambda) = A\left(\frac{\lambda}{2}\right) + A\left(\frac{\lambda}{2}+\pi\right)$ satisfy:

$$(i) \qquad \int_{-\pi}^{\pi} B(\lambda)\cos k\lambda \, d\lambda = 2\int_{-\pi}^{\pi} A(\lambda)\cos 2k\lambda \, d\lambda.$$

$$(ii) \qquad \int_{-\pi}^{\pi} B(\lambda)\cos k\lambda \, d\lambda = \left(-\frac{1}{k^2}\right)^N \int_{-\pi}^{\pi} B^{(2N)}(\lambda)\cos k\lambda \, d\lambda.$$

Proof: **(ii).** It is enough to prove the assertion for $N=1$ as the general case easily follows by recursion. We integrate the LHS by parts, considering $\cos k\lambda$ as the derivative of $\frac{\sin k\lambda}{k}$, while the boundary term vanishes because $\sin k\pi = \sin(-k\pi) = 0$. The second integration by parts using $\sin k\lambda = \left(\frac{-\cos k\lambda}{k}\right)'$ yields the result, since now the boundary term vanishes due to $B'(\pi) = B'(-\pi)$.

(i). By the substitutions $\lambda' := \lambda/2$ and $\lambda'' := \lambda/2+\pi$ in the components corresponding to the terms $A\left(\frac{\lambda}{2}\right)$ and $A\left(\frac{\lambda}{2}+\pi\right)$, respectively, we obtain that

$$\int_{[-\pi,\pi]} B(\lambda)\cos k\lambda \, d\lambda = 2\int_{[-\pi/2,\pi/2]} A(\lambda')\cos 2k\lambda' \, d\lambda' + 2\int_{[\pi/2,3\pi/2]} A(\lambda'')\cos k(2\lambda''-2\pi) \, d\lambda''.$$

Summing the above and using periodicity of A,

$$\int_{[-\pi,\pi]} B(\lambda)\cos k\lambda \, d\lambda \quad = \quad 2\int_{[-\pi,\pi]} A(\lambda)\cos 2k\lambda \, d\lambda$$

■

Lemma 5.3: Let $1<\alpha<2$, φ be a three times continuously differentiable function such that $\varphi(0) = \varphi''(0) = 0$, $\varphi(\lambda) \geq c\lambda$ for $\lambda \in [0,\pi]$, and H is bounded. Then

$$\int_{[0,\pi]} \varphi(\lambda)^{-\alpha} H(\lambda) \, \sin k\lambda \, d\lambda = \mathcal{O}(k^{\alpha-1}). \tag{11}$$

Proof: Let us approximate the function $\varphi(\lambda)^{-\alpha}$ by $(c_1\lambda)^{-\alpha}$, where $c_1 = \varphi'(0)$. We have

$$\left|\varphi(\lambda)^{-\alpha} - (c_1\lambda)^{-\alpha}\right| \leq \leq \alpha \max\left(\varphi(\lambda)^{-\alpha-1}, (c_1\lambda)^{-\alpha-1}\right)\left|\varphi(\lambda) - c_1\lambda\right|.$$

Because of $\varphi(\lambda) \geq c\lambda$ for $\lambda \in [0,\pi]$, the maximum above is of order $\lambda^{-\alpha-1}$. Then for some ξ_λ lying between $c_1\lambda$ and $\varphi(\lambda)$

$$\left|\varphi(\lambda)^{-\alpha} - (c_1\lambda)^{-\alpha}\right| \leq C\lambda^{-\alpha-1}|\varphi'''(\xi_\lambda)|\lambda^3 \leq C\lambda^{-\alpha+2},$$

as assumptions imply that $\varphi'''(\xi_\lambda) = \mathcal{O}(1)$. Thus the difference of the approximated and approximating integrals can be upper bounded by

$$\int_{[0,\pi]} \left|\varphi(\lambda)^{-\alpha} - (c_1\lambda)^{-\alpha}\right| |H(\lambda)| \, d\lambda \leq \tag{12}$$

$$\leq \int_{[0,\pi]} \left|C\lambda^{-\alpha+2}\right| |H(\lambda)| \, d\lambda. \tag{13}$$

Since $1<\alpha<2$, the difference is bounded and thus $\mathcal{O}(1)$. Now we are left with the evaluation of the order of coefficients in the approximating integral. Substituting $\eta := k\lambda$ we get

$$\frac{1}{k}\int_{[0,k\pi]} \left[c_1\frac{\eta}{k}\right]^{-\alpha} H\left(\frac{\eta}{k}\right) \, \sin\eta \, d\eta = = \frac{c_1^{-\alpha}}{k} k^\alpha \int_{[0,k\pi]} \eta^{-\alpha} H\left(\frac{\eta}{k}\right) \, \sin\eta \, d\eta = \mathcal{O}\left(k^{\alpha-1}\right).$$

The last equality follows from the observation that the integrand is $\mathcal{O}\left(\eta^{-\alpha}\right)$ for large η as the remaining terms are bounded and is $\mathcal{O}\left(\eta^{-\alpha+1}\right)$ for small η in view of $\sin\eta = \mathcal{O}(\eta)$. ■

Proof: Theorem 3.1. Assume first that the resolution level $j=1$. Let $C(\lambda) = \varphi(|\lambda|)^{-2d} V(|\lambda|)$. In view of (7) and Lemma 5.2(b) it is enough to prove that

$$\int_{[-\pi,\pi]} \left\{C\left(\frac{\lambda}{2}\right) + C\left(\frac{\lambda}{2}+\pi\right)\right\}^{(2N)} \cos k\lambda \, d\lambda \ = \mathcal{O}(k^{2d-1}). \tag{14}$$

Observe that in view of Lemma 5.2(a), LHS of (14) equals to

$$2^{-2N+1}\int_{[-\pi,\pi]} C^{(2N)}(\lambda) \ \cos 2k\lambda \, d\lambda \ .$$

Since $C^{(2N)}(\cdot)$ is even, the equation (14) will follow from

$$2^{-2N+2}\int_{[0,\pi]} C^{(2N)}(\lambda) \ \cos 2k\lambda \, d\lambda \ = \mathcal{O}(k^{2d-1}). \tag{15}$$

Observe that Lemma 5.1 yields for $\lambda \in [0,\pi]$

$$C^{(2N)}(\lambda) = \left[\varphi(\lambda)^{-2d} V(\lambda)\right]^{(2N)} = \quad (16)$$

$$= \sum \varphi(\lambda)^{-2d-\sum_{j\geq 1} i_j} \times \quad (17)$$

$$\times \prod_{j=1}^{s_1} c_j \left[\varphi^{(j)}(\lambda)\right]^{i_j} V^{(s_2)}(\lambda), \quad (18)$$

where $s_1 + s_2 = 2N$. Assume first that $\sum_{j\geq 1} i_j = s_1$. As $\sum_{i_j\geq 1} j i_j = s_1$, this is only possible when $i_1 = s_1$ and $i_j = 0$ for $j \geq 2$. Thus the product involves only the first derivative of φ i.e. the summand in (18) is of the form

$$\varphi(\lambda)^{-2d-s} V^{(2N-s)}(\lambda)\left[\varphi'(\lambda)^s\right] = \quad (19)$$

$$= \varphi(\lambda)^{-2d}\left[\frac{V^{(2N-s)}(\lambda)}{\varphi(\lambda)^s}\right]\left[\varphi'(\lambda)^s\right], \quad (20)$$

where $s_1 = s$, $s_2 = 2N - s$. Define $F_i(\lambda) = \left[\varphi(\lambda)^{-i}\right] V^{(2N-i)}(\lambda)$ and $G_s(\lambda) = \left[\varphi'(\lambda)^s\right]$. Then (19) equals $\varphi(\lambda)^{-2d} F_s(\lambda) G_s(\lambda)$.

In the remaining case $\sum_{j\geq 1} i_j < s_1$, the corresponding term in $C^{(2N)}(\lambda)$ is a product of $\varphi(\lambda)^{k-2d} F_s(\lambda)$ ($k \geq 1$) and of the higher order derivatives of φ. As its derivative is integrable, straightforward integration by parts implies that it is $\mathcal{O}\left(k^{-1}\right)$.

Thus it is enough to consider the term of $C^{(2N)}(\lambda)$ corresponding to the highest order singularity. It equals to ($s = 0, 1, \ldots, 2N$)

$$\int_{[0,\pi]} \varphi(\lambda)^{-2d} G_s(\lambda) F_s(\lambda) \cos 2k\lambda \, d\lambda \,. \quad (21)$$

Integrating by parts,

$$\int_{[0,\pi]} \varphi(\lambda)^{-2d} G_s(\lambda) F_s(\lambda) \left(\frac{\sin 2k\lambda}{2k}\right)' d\lambda =$$

$$= \varphi(\lambda)^{-2d} G_s(\lambda) F_s(\lambda) \left.\frac{\sin 2k\lambda}{2k}\right|_0^\pi$$

$$- \int_{[0,\pi]} \left[\varphi(\lambda)^{-2d} G_s(\lambda) F_s(\lambda)\right]' \frac{\sin 2k\lambda}{2k} d\lambda.$$

Since the boundary term vanishes, (21) reduces to the sum of three components

$$-\frac{1}{2k}\int_{[0,\pi]} \varphi(\lambda)^{-2d-1} \varphi'(\lambda) G_s(\lambda) F_s(\lambda) \sin 2k\lambda \, d\lambda + \quad (22)$$

$$-\frac{1}{2k}\int_{[0,\pi]} \varphi(\lambda)^{-2d} G_s'(\lambda) F_s(\lambda) \sin 2k\lambda \, d\lambda + \quad (23)$$

$$-\frac{1}{2k}\int_{[0,\pi]} \varphi(\lambda)^{-2d} G_s(\lambda) F_s'(\lambda) \sin 2k\lambda \, d\lambda \,. \quad (24)$$

Observe that in view of assumptions on V and φ functions F_i and G_i are bounded for $i = 0, 1, \ldots, 2N$. Indeed, Taylor expansion of $V^{(2N-s)}(\lambda)$ up to the term $\mathcal{O}(\lambda^s)$ yields $V^{(2N-s)}(\lambda) = \mathcal{O}(\lambda^s)$ and $\varphi^{-s}(\lambda) = \mathcal{O}(\lambda^{-s})$ in view of assumptions. Thus by Lemma 5.3 the part (22) has the order $\mathcal{O}(k^{2d-1})$. Since $G_s'(\lambda) F_s(\lambda)$ is bounded on $[0,\pi]$ and $\sin k\lambda = \mathcal{O}(\lambda)$ then the integrand in (23) is bounded and the integral is $\mathcal{O}(k^{-1})$. In order to handle (24) observe that the derivative F_s', whenever exists, can be expressed as

$$F_s'(\lambda) = \varphi(\lambda)^{-1}\left(F_{s-1}(\lambda) - sF_s\varphi'(\lambda)\right).$$

Since the function $G_s(\lambda)\left(F_{s-1}(\lambda) - sF_s\varphi'(\lambda)\right)$ is bounded for $s \geq 1$, using Lemma 5.3 again we obtain that the integral in (24) is $\mathcal{O}(k^{2d-1})$. For $s = 0$ observe that $F_0'(\lambda) = V^{(2N+1)}(\lambda)$. As $V^{(2N+1)}(\cdot)$ is bounded, the integral (24) is $\mathcal{O}(k^{-1})$ for $s = 0$. Thus (15) is proved.

The case $j > 1$ is obtained by an analogous argument noting that in view of (7) function $C(\lambda)$ has to be replaced now by $C_j(\lambda) = \varphi(\lambda)^{-2d} V_j(\lambda)$. Then if any of $2N$ derivatives falls onto U-term in the product V_j, the resulting integrand function has no singularity. Thus we can reduce the problem as previously considering a modified function $G(\lambda) = \varphi'(\lambda)^s \prod_{l=0}^{j-2} U(2^l\lambda)$. Note that differentiability of U up to the order $2N$ follows from $U(\lambda) = V(\lambda + \pi)$. ■

Proof: Theorem 3.2. The proof proceeds as that of the previous result by noting that one obtains the decorrelation of $(d_{1,k})$ pertaining to the spectral density $f(\lambda) = \varphi(|\lambda|)^{-2d} L(|\lambda|)$, setting $C(\lambda) = \varphi(|\lambda|)^{-2d} L(|\lambda|) V(\lambda)$. Then if any of $2N$ derivatives falls onto L-term in the product on the interval $[0,\pi]$, the resulting integrand function has no singularity and the problem is reduced as previously with function $G(\lambda) = \left[\varphi'(\lambda)\right]^s L(\lambda)$. ■

Remark 5.4: The following observations follow immediately from the proofs of Theorems 3.1 and 3.2.

(i) The results hold true when φ is a $2N$ times differentiable function such that $\varphi(\lambda) = c_1\lambda + \mathcal{O}(\lambda^3)$ for $\lambda \to 0$ and $\varphi(\lambda) \geq c\lambda$ for $\lambda \in [0,\pi]$.

(ii) One can restate the results in terms of functions $F_i(\lambda) = \left[\varphi(\lambda)^{-i}\right] V^{(2N-i)}(\lambda)$. Namely, it is enough to assume that F_i' exist for all $i = 0, 1, \ldots, 2N$ except the finite number of points and F_i are bounded.

(iii) Condition on $V^{(2N+1)}(\cdot)$ may be weakened to $V^{(2N+1)}(\lambda) = \mathcal{O}(\lambda^{-1})$ when $\lambda \to 0$. Moreover, it is sufficient that derivatives $L^{(i)}(\cdot), i = 0, 1, \ldots, 2N$ exist except a finite number of points and are uniformly bounded in $(0,\pi]$.

(iv) The result still holds true for the spectral density $f(\lambda) = \varphi(|\lambda|)^{-2d} L(|\lambda|)$, where $L(\lambda) = \log(1/\lambda)$. The reasoning is as follows. Since the derivative $L^{(k)}(\lambda)$ equals $C_k\lambda^{-k}$ for $\lambda > 0$, the term $\varphi(\lambda)^{-2d-\sum_{j\geq 1} i_j}$ in (18) is replaced with

$$\varphi(\lambda)^{-2d} C_k \lambda^{-k} \varphi(\lambda)^{-\sum_{j\geq 1} i_j},$$

where $\sum j i_j = s_1 - k$, F_s is now $F_s(\lambda) = \frac{V^{(2N-s)}(\lambda)}{\lambda^k \varphi(\lambda)^{s-k}}$, and $G_s(\lambda) = C_k\left[\varphi'(\lambda)\right]^s$. The reasoning remains the same and the argument for $j > 1$ is analogous.

References

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

[2] Abry, P., Flandrin, P., Taqqu, M.S. and Veitch, D. Self-similarity and long-range dependence through the wavelet lens. In: [8].

[3] Beran, J. *Statistics for Long Memory Processes*, Chapman and Hall, New York, 1994.

[4] Berman, S.M. *Sojourns and Extremes of Stochastic Processes*, Wadsworth, 1992.

[5] Craigmile, P.F. and Percival, D.B. Asymptotic decorrelation of between-scale wavelet coefficients. *IEEE Trans. Inform. Theory* **51**, 1039-1048, 2005.

[6] Daubechies, I. *Ten Lectures on Wavelets* Vol. 61, CBMS-NSF Regional Conference Series in Applied Mathematics. Philadelphia, SIAM, 1992.

[7] Donoho, D.L. and Johnstone, I.M. Ideal wavelet adaptation by wavelet shrinkage. *Biometrika* **81**, 425-455, 1994.

[8] Doukhan, P., Oppenheim, G. and M.S. Taqqu (eds.) *Theory and Applications of Long-Range Dependence*, Birkhäuser, 2003.

[9] Flandrin, P. Wavelet analysis and synthesis of fractional Brownian motion. *IEEE Trans. Inform. Theory* **38**, 910-917, 1992.

[10] McCoy, E. J. and Walden, A. T. Wavelet analysis and synthesis of stationary long-memory processes. *J. Comput. Graph. Statist.* **5**, 26-56, 1996.

[11] Mielniczuk, J. and Wojdyllo, P. Estimation of Hurst exponent revisited. *Computational Statistics and Data Analysis*, to appear.

[12] Mielniczuk, J. and Wojdyllo, P. Wavelets for time series data: review and new results. *Control & Cybernetics*, **34**, 1093-1126, 2005.

[13] Percival, D.B. and Walden, A.T. *Wavelet Methods for Time Series Analysis*, Cambridge, Cambridge University Press, 2000.

[14] Robinson, P.M. *Time Series with Long Memory*, Oxford University Press, 2003.

[15] Samorodnitsky, G. and Taqqu, M. *Stable Non-Gaussian Random Processes*, New York, Chapman and Hall, 1994.

[16] Tewfik, A.H. and Kim, M. Correlation structure of discrete wavelet coefficients of fractional Brownian motion. *IEEE Trans. Inform. Theory* **38**, 904-909, 1992.

[17] Vidakovic, B. *Statistical Modelling by Wavelets*, Wiley, 1999.

[18] Wojtaszczyk, P. *A Mathematical Introduction to Wavelets*, Cambridge, Cambridge University Press, 1997.

[19] Yong, C.H. *Asymptotic Behaviour of Trigonometric Series*, Chinese Univ. Hong Kong, 1974.

[20] Zhang, L., Bao, P. and Wu, X. Wavelet estimation of fractional Brownian motion in a noisy environment *IEEE Trans. Inform. Theory* **50**, 2194-2200, 2004.

Jan Mielniczuk Jan Mielniczuk received the M.Sc. and Ph.D. degrees in mathematics from Warsaw University, Poland, in 1981 and 1985, respectively, and the D.Sc. degree from the Institute of Mathematics of Polish Academy of Sciences in 1996.

In 1981 he joined Institute of Computer Science, Polish Academy of Sciences, Warsaw, where he is now an Associate Professor. He is also an Associate Professor at the Faculty of Mathematics and Information Sciences of Warsaw University of Technology. He held several visiting appointments at the University of Michigan and Rice University. His research interests are in time series and dependence analysis, statistical modeling, nonparametric inference and curve estimation. He is an associate editor for Statistics.

Piotr Wojdyllo Piotr Wojdyllo received M.Sc. summa cum laude and Ph.D. degrees from Department of Mathematics, Mechanics, and Computer Science, Warsaw University in 1995 and 2001, respectively. Since 2004 he holds an Assistant Professor position at Institute of Mathematics, Polish Academy of Sciences working on the functional analytic and algebraic aspects of wavelet and Gabor systems as well as on applications to signal processing and data analysis.