# Whittle estimation with (quasi-)analytic wavelets

Sophie Achard, Irène Gannaz

# Local Whittle estimation with (quasi-)analytic wavelets

Sophie Achard

Univ. Grenoble Alpes, CNRS, Inria, Grenoble INP, LJK, France

Irène Gannaz

Univ Lyon, INSA Lyon, UJM, UCBL, ECL, ICJ, UMR5208,
69621 Villeurbanne, France

June 2022

**Abstract**

In the general setting of long-memory multivariate time series, the long-memory characteristics are defined by two components. The long-memory parameters describe the autocorrelation of each time series. And the long-run covariance measures the coupling between time series, with general phase parameters. This wide class of models provides time series not necessarily Gaussian nor stationary. It is of interest to estimate the parameters: long-memory, long-run covariance and general phase. This inference is not possible using real wavelets decomposition or Fourier analysis. Our purpose is to define an inference approach based on a representation using quasi-analytic wavelets. We first show that the covariance of the wavelet coefficients provides an adequate estimator of the covariance structure including the phase term. Consistent estimators based on a Whittle approximation are then proposed. Simulations highlight a satisfactory behavior of the estimation on finite samples on linear time series and on multivariate fractional Brownian motions. An application on a real dataset in neuroscience is displayed, where long-memory and brain connectivity are inferred.

**Keywords.** Multivariate processes, long-memory, covariance, phase, wavelets, cerebral connectivity

## 1 Introduction

Multivariate processes are often observed nowadays thanks to the recordings of multiple sensors simultaneously. Numerous examples can be cited such as hydrology [Whitcher and Jensen, 2000],

finance [Gençay et al., 2001] or neuroscience [Achard and Gannaz, 2016]. When in addition the time series have the property of long-memory, the definition of the model is complicated and several definitions can be proposed. A simple model is defined by long-memory parameters for each univariate process with a covariance structure added between the processes [Robinson, 1995a]. However phenomena such as co-integration cannot be modeled using this simple definition. An alternative was proposed by Robinson [2008], Kechagias and Pipiras [2014], where a phase-term is added to the covariance structure.

Let $\mathbf{X}(t)$ denote a multivariate long-memory dependence process $\mathbf{X}(t) = \begin{bmatrix} X_1(t) & \dots & X_p(t) \end{bmatrix}^T$, $t \in \mathbb{Z}$ with long memory parameters $\mathbf{d} = (d_1, d_2, \dots, d_p)$, $\mathbf{d} \in (-0.5, +\infty)^p$. We will denote by $\mathbb{L}$ the difference operator, $(\mathbb{L}\mathbf{X})(t) = \mathbf{X}(t+1) - \mathbf{X}(t)$. The $k$-th difference operator, $\mathbb{L}^k$, $k \in \mathbb{N}$, is defined by $k$ recursive applications of $\mathbb{L}$. For $\mathbf{D} = \lfloor \mathbf{d} + 1/2 \rfloor$, we assume that the multivariate process $\mathrm{Diag}\left(\mathbb{L}^{D_\ell}, \ell = 1, \dots, p\right)\mathbf{X}$ is covariance stationary with a spectral density matrix given by

$$\text{(M-1)}\quad \boldsymbol{f}^{\mathbf{D}}(\lambda) = \left(\mathrm{Diag}\left(|\lambda|^{-d_1^*}, \dots, |\lambda|^{-d_p^*}\right) \boldsymbol{\Theta}\, \mathrm{Diag}\left(|\lambda|^{-d_1^*}, \dots, |\lambda|^{-d_p^*}\right)\right) \circ \boldsymbol{f}^S(\lambda), \quad \text{for all } \lambda > 0,$$

where $\circ$ denotes the Hadamard product, and $d_m^* = d_m - D_m \in (-0.5, 0.5)$ for all $m$. The process $\mathbb{L}^{D_m} X_m$ is said to have long-memory if $d_m^* \in (0, 0.5)$, and to be anti-persistent if $d_m^* \in (-0.5, 0)$ (see for instance Lobato [1999], Shimotsu [2007]). For simplicity of notation, we will use the term long-memory call parameters $\mathbf{d}$ throughout the paper.

The function $\boldsymbol{f}(\cdot)$ is defined by

$$\boldsymbol{f}(\lambda) = (\boldsymbol{\Lambda}(\lambda)\, \boldsymbol{\Theta}\, \boldsymbol{\Lambda}(\lambda)) \circ \boldsymbol{f}^S(\lambda), \quad \text{for all } \lambda > 0,$$

with $\boldsymbol{\Lambda}(\lambda) = \mathrm{Diag}\left(|\lambda|^{-d_1}, \dots, |\lambda|^{-d_p}\right)$. Under the condition (M-1), the function $\boldsymbol{f}(\cdot)$ is called the generalized spectral density of the multivariate process $\{\mathbf{X}(t),\, t \in \mathbb{Z}\}$.

The function $\boldsymbol{f}^S(\cdot)$ represents the short-range memory of $\boldsymbol{f}(\cdot)$. In order to get identifiability, it is necessary to assume $\boldsymbol{f}^S(0) = \mathbf{1}$. The following assumption is also needed to control the regularity.

$$\text{(M-2)}\quad \text{There exists } C_f > 0 \text{ and } \beta > 0 \text{ such that } \sup_{0<\lambda<\pi} \sup_{\ell,m=1,\dots,N} \frac{\left|f^S_{\ell,m}(\lambda) - 1\right|}{\lambda^\beta} \leqslant C_f\ .$$

In particular, our definition agrees with the one given in Kechagias and Pipiras [2014] if $D_m = 0$ for

all $m$. Our definition (M-1) encompass both stationary and non-stationary time series. Therefore it has the advantage of covering multivariate fractional Brownian motion [Coeurjolly et al., 2013].

The major interest of this model is the introduction of the matrix $\mathbf{\Theta}$. This provides a generalisation of multivariate long-memory models used in Lobato [1997], Shimotsu [2007], Achard and Gannaz [2016]. Indeed, the matrix $\mathbf{\Theta}$ can be written as,

$$\Theta_{\ell,m} = \Omega_{\ell,m} \mathrm{e}^{\mathrm{i}\phi_{\ell,m}} \ ,$$

with $\mathbf{\Omega} = (\Omega_{\ell,m})_{\ell,m=1,\dots,p}$ a real symmetric non-negative semi-definite matrix and $\mathbf{\Phi} = (\phi_{\ell,m})_{\ell,m=1,\dots,p}$ an anti-symmetric matrix. Let the bar above denote the conjugate operator. The matrix $\mathbf{\Theta}$ satisfies $\mathbf{\Theta}^T = \overline{\mathbf{\Theta}}$ since $\boldsymbol{f}^T(\cdot) = \overline{\boldsymbol{f}(\cdot)}$. We will use $\|\mathbf{\Omega}\|$ to denote the infinity norm, that is, $\|\mathbf{\Omega}\| = \max_{\ell,m=1,\dots,p} \Omega_{\ell,m}$. In Lobato [1997], Shimotsu [2007], Achard and Gannaz [2016], the phase term was defined by $\phi_{\ell,m} = \pi(d_\ell - d_m)/2$.

In a univariate setting, the main parameter of interest is the long-memory parameter or equivalently the Hurst parameter. In this particular case, three main families of Fourier-based estimation have already been proposed: the average periodogram estimation [Robinson, 1994], the log periodogram regression [Geweke and Porter-Hudak, 1983, Robinson, 1995a] and semiparametric estimation based on Whittle approximation [Künsch, 1987, Robinson, 1995b]. Estimation with a wavelet representation of time series was proposed in Abry and Veitch [1998] with a log-scalogram approach similar to log-periodogram estimation and in Moulines et al. [2008] with a wavelet-based local Whittle estimation.

In a multivariate setting, estimation procedures have also been proposed using either Fourier or wavelets. For a general phase term, Sela and Hurvich [2012] proposed an estimation based on the average periodogram and Robinson [2008] and Baek et al. [2020] developed a Fourier-based local Whittle estimation. For a fixed phase term, $\phi_{\ell,m} = \pi(d_\ell - d_m)/2$, estimation of both the covariance structure and the long-memory was proposed by Lobato [1999] and Shimotsu [2007], with a Fourier-based local Whittle estimation, and by Achard and Gannaz [2016] with a similar procedure based on a real wavelet representation.

The objective of this work is to propose an estimation procedure in the general framework described above, with a general phase, based on a wavelet representation of the processes rather than a Fourier representation. We do not consider here the fractional co-integration and we refer to Baek et al.

[2020] for a Fourier-based local Whittle estimation.

Introducing wavelets is motivated by their flexibility for real data applications. It enables in particular to consider non-stationary processes thanks to an implicit differentiation. The introduction of a general phase term challenges the choice of the wavelet filters. Because of condition (M-1), we need to consider analytic filters especially for identifying the imaginary part of $\mathbf{\Theta}$. Indeed, as illustrated in Gannaz et al. [2017], with real wavelet filters it is not possible to recover both the real and imaginary part of the matrix $\mathbf{\Theta}$. Quasi-analytic wavelet filters are described in Section 2. Main properties of the filters are displayed and an approximation of the covariance of wavelet coefficients is derived in Section 3. Section 4 recalls the definition of the wavelet local Whittle estimators and gives their consistency and their convergence rate, as well as their asymptotic distribution. Section 5 reports some simulation results, on ARFIMA linear models and on multivariate fractional Brownian motions. Section 6 provides an empirical application on neuroscience data. The detailed proofs are provided in Appendix.

## 2 Transform of the multivariate process

We first define the filters used to transform the multivariate time series $\{\mathbf{X}(t),\, t \in \mathbb{Z}\}$.

Let $(h^{(L)}, h^{(H)})$ and $(g^{(L)}, g^{(H)})$ denote two pairs of respectively low-pass and high-pass filters. Let $(\varphi_h(\cdot), \psi_h(\cdot))$ be respectively the father and mother wavelets associated to $(h^{(L)}, h^{(H)})$. They can be defined through their Fourier transforms as

$$\widehat{\varphi}_h(\lambda) = 2^{-1/2} \prod_{j=1}^{\infty} \left[ 2^{-1/2} \widehat{h}^{(L)}(2^{-j}\lambda) \right] \quad \text{and} \quad \widehat{\psi}_h(\lambda) = 2^{-1} \widehat{h}^{(H)}(\lambda/2)\, \widehat{\varphi}_h(\lambda/2). \tag{1}$$

Let us define similarly $(\varphi_g, \psi_g)$ the father and the mother wavelets associated with the wavelet filters $g^{(L)}$ and $g^{(H)}$. Their Fourier transforms are

$$\widehat{\varphi}_g(\lambda) = 2^{-1/2} \prod_{j=1}^{\infty} \left[ 2^{-1/2} \widehat{g}^{(L)}(2^{-j}\lambda) \right] \quad \text{and} \quad \widehat{\psi}_g(\lambda) = 2^{-1} \widehat{g}^{(H)}(\lambda/2)\, \widehat{\varphi}_g(\lambda/2). \tag{2}$$

The complex father and mother wavelets $(\varphi(\cdot), \psi(\cdot))$ are then defined by

$$\widehat{\varphi}(\lambda) = \widehat{\varphi}_h(\lambda) + \mathrm{i}\,\widehat{\varphi}_g(\lambda) \quad \text{and} \quad \widehat{\psi}(\lambda) = \widehat{\psi}_h(\lambda) + \mathrm{i}\,\widehat{\psi}_g(\lambda). \tag{3}$$

### 2.1 Selesnick's common factor filters

From Heisenberg's uncertainty principle, analytic filters with finite support do not exist. We choose here to consider the quasi-analytic filters introduced by Thiran [1971], Selesnick [2002]. The *common-factor* wavelets, defined by Selesnick [2002], are parametrized by a degree $L$ on the analytic property of the derived complex wavelet. We refer the reader to Selesnick [2001, 2002], Achard et al. [2020] for a fuller description of the construction of the wavelets and their properties.

Let $\widehat{d}_L(\lambda)$ be defined by

$$\widehat{d}_L(\lambda) = \mathrm{e}^{\mathrm{i}\,\lambda(-L/2+1/4)} \left[\cos(\lambda/4)^{2L+1} + \mathrm{i}\,(-1)^{L+1} \sin(\lambda/4)^{2L+1}\right].$$

Next, filters $(\widehat{h}^{(L)}, \widehat{h}^{(H)})$, and $(\widehat{g}^{(L)}, \widehat{g}^{(H)})$ are defined by

$$\widehat{h}^{(L)}(\lambda) = 2^{-M+1/2} \left(1 + e^{-i\lambda}\right)^M \widehat{q}_{L,M}(\lambda)\, \widehat{d}_L(\lambda) \quad \text{and} \quad \widehat{h}^{(H)}(\lambda) = \overline{\widehat{h}^{(L)}(\lambda+\pi)} e^{-i\lambda}, \tag{4}$$

$$\widehat{g}^{(L)}(\lambda) = 2^{-M+1/2} \left(1 + e^{-i\lambda}\right)^M \widehat{q}_{L,M}(\lambda)\, \overline{\widehat{d}_L(\lambda)} e^{-i\lambda L} \quad \text{and} \quad \widehat{g}^{(H)}(\lambda) = \overline{\widehat{g}^{(L)}(\lambda+\pi)} e^{-i\lambda}, \tag{5}$$

with $\widehat{q}_{L,M}(\lambda)$ a real polynomial of $(e^{-i\lambda})$ such that $\widehat{q}_{L,M}(0) = 1$. Observe that the normalization of the filters is different from the one defined in Achard et al. [2020].

Common-factor wavelets are introduced with $\widehat{q}_{L,M}$ such that the wavelet decomposition satisfies the perfect reconstruction condition. This condition is classically used for deriving wavelet bases $2^{1/2}\psi_{g\,j,k} = 2^{-1/2}2^{j/2}\psi_g(2^j\cdot - k)$ and $2^{1/2}\psi_{h\,j,k} = 2^{-1/2}2^{j/2}\psi_h(2^j\cdot - k)$, $j, k \in \mathbb{Z}$, which are orthonormal bases of $L^2(\mathbb{R})$. In that case, $\widehat{q}_{L,M}$ is defined as a solution of

$$|\widehat{q}_{L,M}(\lambda)|^2\, s(\lambda) + |\widehat{q}_{L,M}(\lambda+\pi)|^2\, s(\lambda+\pi) = 1\,, \tag{6}$$

where $s(\lambda) = \frac{2^{4L-1}}{(2L+1)^2} 2^{-M} (1+\cos(\lambda))^M \left|\widehat{d}_L(\lambda)\right|^2$. The existence of $\widehat{q}_{L,M}$ is proved in Achard et al. [2020]. However, to the best of our knowledge, under perfect reconstruction, no explicit expression of $\widehat{q}_{L,M}$ is easy to obtain. Sine perfect reconstruction is not necessary for deriving estimation

procedures, we can assume that $\widehat{q}_{L,M}$ is a constant equal to 1.

**Definition 1** (Common-Factor Wavelets (CFW)). *Let $M$, $L$ be strictly positive integers. Let $(\varphi, \psi)$ be a family of Common-Factor wavelets defined by equations (1), (2), (3), and (4), (5). If the filter $\widehat{q}_{L,M}$ satisfies perfect reconstruction condition (6), the pair $(\varphi, \psi)$ will be denoted by CFW-PR(M,L). If $\widehat{q}_{L,M}$ is a constant polynomial equal to 1, $(\varphi, \psi)$ will be denoted by CFW-C(M,L) filters.*

Let us first recall the main result established in Achard et al. [2020].

**Theorem 2** (Achard et al. [2020]). *For all $\lambda \in \mathbb{R}$, for all $\widehat{q}_{L,M}$ real polynomial of $(e^{-i\lambda})$,*

$$\widehat{\psi}(\lambda) = \widehat{\psi}_h(\lambda) + \mathrm{i}\,\widehat{\psi}_g(\lambda) = \left(1 - \mathrm{e}^{\mathrm{i}\eta_L(\lambda)}\right)\widehat{\psi}_h(\lambda)\,,$$

$$\text{with} \quad a_L(\lambda) = 2(-1)^L \operatorname{atan}\left(\tan^{2L+1}(\lambda/4)\right) \quad \text{and} \quad \eta_L(\lambda) = -a_L(\lambda/2+\pi) + \sum_{j=1}^{\infty} a_L(2^{-j-1}\lambda). \tag{7}$$

*Additionally, for all $\lambda \in \mathbb{R}$,*

$$\left|\widehat{\psi}_h(\lambda) + \mathrm{i}\,\widehat{\psi}_g(\lambda) - 2\mathbb{1}_{\mathbb{R}_+}(\lambda)\,\widehat{\psi}_H(\lambda)\right| = U_L(\lambda)\left|\widehat{\psi}_H(\lambda)\right|\,,$$

*where $U_L$ is a $\mathbb{R} \to [0,2]$ function satisfying, for all $\lambda \in \mathbb{R}$,*

$$U_L(\lambda) \leqslant 2\sqrt{2}\left(\log_2\left(\frac{\max(4\pi, |\lambda|)}{2\pi}\right) + 2\right)\left(1 - \frac{\delta(\lambda, 4\pi\mathbb{Z})}{\max(4\pi, |\lambda|)}\right)^{2L+1}\,.$$

*and, for all $\lambda \in \mathbb{R}$ and $A \subset \mathbb{R}$, $\delta(\lambda, A)$ denotes the distance of $\lambda$ to $A$ defined by $\delta(\lambda, A) = \inf\{|\lambda - x|\,,\ x \in A\}$.*

In equation (7), we adopt the convention that $\operatorname{atan}(\pm\infty) = \pm\pi/2$ so that $\alpha_L$ is well defined on $\mathbb{R}$.

Theorem 2 quantifies the quality of the analytic approximation. Observe that the function $U_L(\cdot)$ only depends on the parameter $L$. The higher $L$, the better the analytic approximation.

## 3 Moments approximations of the wavelet coefficients

Let $\{\mathbf{W}_{j,k},\, j \geqslant 0,\, k \in \mathbb{Z}\}$ denote the wavelet coefficients of the process $\mathbf{X}$ associated to the wavelet pair $(\varphi, \psi)$. At a given resolution $j \geqslant 0$, for $k \in \mathbb{Z}$, we define the dilated and translated functions $\psi_{j,k}(\cdot) = 2^{-j/2}\psi(2^{-j}\cdot - k)$. The wavelet coefficients of the process $\mathbf{X}$ are defined by

$$\mathbf{W}_{j,k} = \int_{\mathbb{R}} \widetilde{\mathbf{X}}(t)\psi_{j,k}(t)dt \quad j \geqslant 0, k \in \mathbb{Z},$$

where $\widetilde{\mathbf{X}}(t) = \sum_{k\in\mathbb{Z}} \mathbf{X}(k)\varphi(t-k)$. Given any $j \geqslant 0$ and any $k \in \mathbb{Z}$, $\mathbf{W}_{j,k}$ is a $p$-dimensional vector $\mathbf{W}_{j,k} = \begin{pmatrix} W_{j,k}(1) & W_{j,k}(2) & \dots & W_{j,k}(p)^T \end{pmatrix}$ where $W_{j,k}(a) = \int_{\mathbb{R}} \widetilde{X_a}(t)\psi_{j,k}(t)dt$, $a = 1, \dots, p$. Throughout the paper, we adopt the convention that large values of the scale index $j$ correspond to coarse scales (low frequencies).

We will consider the behavior of $\mathrm{Cov}(\mathbf{W}_{j,k})$, defined as follows

$$\mathrm{Cov}(\mathbf{W}_{j,k}) = \mathbb{E}\left[\mathbf{W}_{j,k}\overline{\mathbf{W}_{j,k}}^T\right] = \int_{-\pi}^{\pi} \boldsymbol{f}(\lambda)\, |\widehat{\tau}_j(\lambda)|^2 \, \mathrm{d}\lambda\,,$$

with $\widehat{\tau}_j(\lambda) = \int_{-\infty}^{\infty} \sum_{\ell\in\mathbb{Z}} \varphi(t+\ell)\, \mathrm{e}^{-\mathrm{i}\,\lambda\,\ell} 2^{-j/2}\psi(2^{-j}t)\mathrm{d}t$.

In practice, a finite number of observations of the process $\mathbf{X}$ are available, $\mathbf{X}(1), \mathbf{X}(2), \dots \mathbf{X}(N)$. Since the wavelets have a compact support, only a finite number of coefficients are non null at each scale $j$. More precisely for every $j \geqslant 0$, let $n_j$ denote the number of coefficients $\mathbf{W}_{j,k}$ evaluated using all the observations. For every $k < 0$ and $k > n_j$, the coefficients $\mathbf{W}_{j,k}$ are set to zero because of finite length of observations. The number $n_j$ is such that, when $j$ goes to infinity, $2^j n_j$ is equivalent to $T_\varphi + N$, where $T_\varphi$ denotes the length of the support of the function $\varphi(.)$. With CFW-PR(M,L) filters, $T_\varphi$ is equal to $2(M+L)$, while with CFW-C(M,L) filters, $T_\varphi = M + L + 1$. In the following, we will assume that $M$ is fixed and finite and that $L$ may go to infinity. If $L\,N^{-1} \to 0$, then $2^j n_j$ is equivalent to $2^{-j}N$. It is the case in particular when $2^{-j}L$ is bounded and $2^{-j}N \to \infty$. Then, the behavior of $n_j$ is similar to the framework of Moulines et al. [2008] and Achard and Gannaz [2016].

### 3.1 Motivations

In this section results obtained in Gannaz et al. [2017] are summarized. We begin with a bivariate ARFIMA(0,**d**,0) process defined as

$$X_\ell(k) = (1-\mathbb{L})^{-d_\ell} u_\ell(k), \quad \ell = 1,2,\ k \in \mathbb{Z}, \tag{8}$$

where $\mathbb{L}$ is a lag operator and $\begin{pmatrix} u_1(k) \\ u_2(k) \end{pmatrix}$ *i.i.d.* with distribution $\mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \mathbf{\Omega}\right)$, where $\mathbf{\Omega} = \begin{pmatrix} 1 & 0.8 \\ 0.8 & 1 \end{pmatrix}$. The spectral density of $(X_1, X_2)$ satisfies (M-1) with $\Theta_{\ell,m} = \Omega_{\ell,m} \mathrm{e}^{\mathrm{i}\,\phi_{\ell,m}}$, $\phi_{\ell,m} = \pi(d_1 - d_2)/2$. Let **d** be equal to $(0.2, 1.2)$. The phase is equal to $\pi/2$ and, hence, $\Theta_{\ell,m}$ is imaginary. Let us now illustrate the impossibility using real wavelets decomposition o infer $\Theta_{\ell,m}$.

We simulate $\mathbf{X}(1), \ldots, \mathbf{X}(2^J)$, with $J = 12$. For each scale $j \geqslant 0$, we evaluate the sample wavelet covariances as $\widehat{\mathbf{\Sigma}}(j) = \frac{1}{n_j} \sum_{k=0}^{n_j-1} W_{j,k}(1)\overline{W_{j,k}(2)} - \left(\frac{1}{n_j} \sum_{k=0}^{n_j-1} W_{j,k}(1)\right)\left(\frac{1}{n_j} \sum_{k=0}^{n_j-1} \overline{W_{j,k}(2)}\right)$, and the wavelet sample correlations as $\widehat{\Sigma}_{1,2}(j)/\sqrt{\widehat{\Sigma}_{1,1}(j)\widehat{\Sigma}_{2,2}(j)}$. Note that by correlations instead of covariance we did not really have to take care of the normalization factors.

Figure 1 displays the behavior of sample wavelet correlations with respect to the scale $j$ on 100 realizations of $(\mathbf{X}(1), \ldots, \mathbf{X}(2^J))$. First observe that for real wavelets (left colmn), $\Theta_{1,2}$ is equal to 0 for scale 8 which confirms the impossibility to identify $\Theta_{1,2}$. In addition, plots displayed in middle and right columns confirm that the imaginary part of the sample wavelet coefficient correlations does not vanish for CFW-PR(M,L) and CFW-C(M,L) filters. The average sample correlation seems to converge to $\Omega_{1,2}\mathrm{e}^{\mathrm{i}\phi_{1,2}}/\sqrt{\Omega_{1,1}\Omega_{2,2}}$ when the frequency decreases.

### 3.2 Theoretical results

We will now develop the theory of the behavior of $\mathrm{Cov}(\mathbf{W}_{j,k})$. This result consists in the extension of Proposition 3 of Achard and Gannaz [2016] to quasi-analytic wavelets. The results are obtained hereafter only for CFW-C(M,L) filters. Indeed, the same results are more difficult to obtain for CFW-PR(M,L) filters because no explicit expression of $\widehat{q}_{L,M}$ satisfying (6) is available.

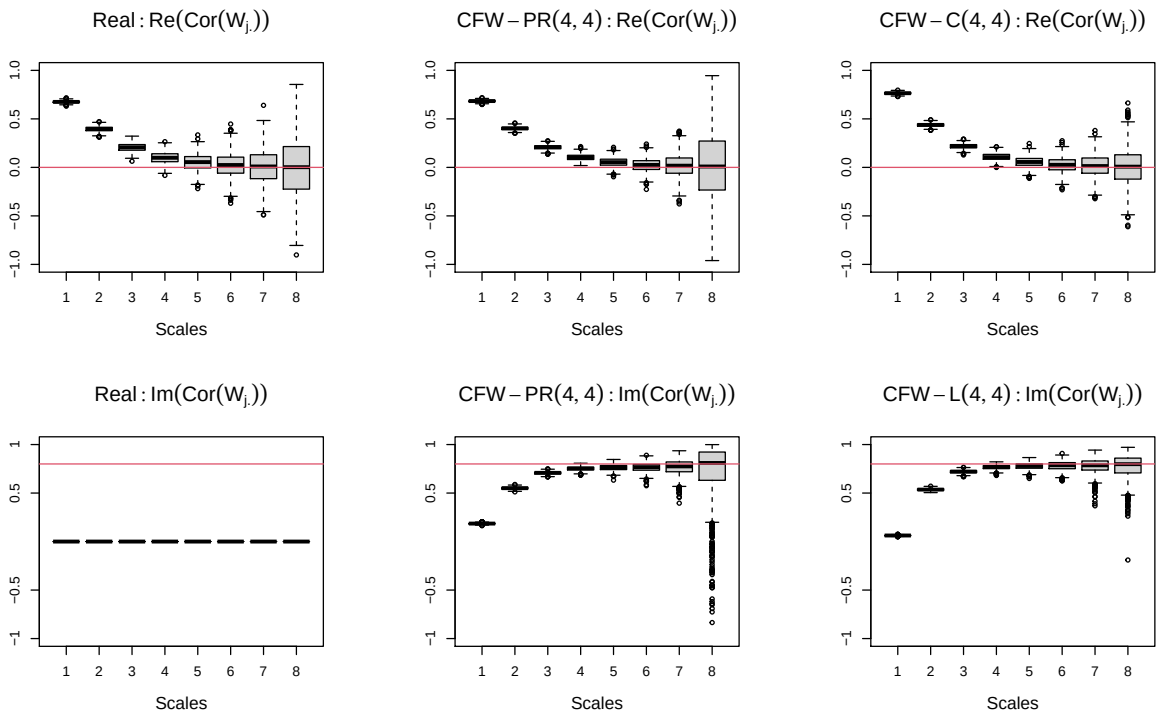


Figure 1: Boxplots of sample correlations between wavelet coefficients at different scales for the bivariate process defined in (8). First row gives the real part of the correlations and second row gives the imaginary part. Each column corresponds to a different wavelet filters, from left to right: Daubechies' real wavelets with M=4, CFW-PR(4,4) and CFW-C(4,4). Horizontal red lines correspond to the real value, that is, $\Omega_{1,2}\cos(\phi_{1,2})/\sqrt{\Omega_{1,1}\Omega_{2,2}}$ for the real part and $\Omega_{1,2}\sin(\phi_{1,2})/\sqrt{\Omega_{1,1}\Omega_{2,2}}$ for the imaginary part.

Our basic assumption on the regularity of the spectral density is the following.

(C-a) $-M + \beta/2 + 1/2 < d_\ell < M \quad$ for all $\quad \ell = 1, \ldots, p,\ M \geqslant 2$

Parameter $M$ is the number of vanishing moments and it also corresponds to the regularity of CFW-C(M,L) filters. Parameters $(d_\ell)_{\ell=1,\ldots p}$ and $\beta$ characterize the dependence in the spectral domain (M-1)-(M-2).

Let us first prove the following approximation using the regularity of the filters.

**Proposition 3.** *Let* $\mathbf{X}$ *be a p-multivariate long range dependent process with long memory parameters* $d_1, \ldots, d_p$ *with normalized spectral density* $f(\cdot)$ *satisfying (M-1) with short-range behavior (M-2). Consider* $\{W_{j,k}(\ell),\ (j,k) \in \mathbb{Z},\ \ell = 1, \ldots, p\}$ *the wavelet coefficients obtained with CFW-C(M,L) filters. Suppose that (C-a) hold. Then we have, for all* $j \geqslant 0$, $k \in \mathbb{Z}$,

$$\left| \mathrm{Cov}(W_{j,k}(\ell), W_{j,k}(m)) - 2^{j(d_\ell + d_m)} \Omega_{\ell,m} \int_{-\infty}^{\infty} \mathrm{e}^{\mathrm{sign}(\lambda)\, \phi_{\ell,m}}\ |\lambda|^{-d_\ell - d_m} \left| \widehat{\psi}(\lambda) \right|^2\ \mathrm{d}\lambda \right| \leqslant C_1\, 2^{j\,(d_\ell + d_m - \beta)}\ ,$$

*where* $C_1$ *is a constant only depending on* $M, L$ *and* $C_f, \beta, \|\mathbf{\Omega}\|, \{d_\ell, \ell = 1, \ldots, p\}$.

The proof is given in Appendix.

The result follows from the fact that CFW-C(L,M) satisfy the assumptions (W1)–(W4) described in Moulines et al. [2008] and Achard and Gannaz [2016] (see Appendix). Note that it does not depend on the quasi-analytic property.

The use of the Proposition 3 in inference needs the evaluation of the integral depending of $|\widehat{\psi}(\lambda)|^2$. With real wavelets, the approximation is given in Proposition 3 of Achard and Gannaz [2016]. Since $|\widehat{\psi}(\lambda)|^2$ is a real and symmetric function, the imaginary part of the integral is null. Consequently, a cosine term with the phase appears in the approximation of the covariance. That is, we would obtain in this framework an approximation of the form

$$\left| \mathrm{Cov}(W_{j,k}(\ell), W_{j,k}(m)) - 2^{j(d_\ell + d_m)}\, \Omega_{\ell,m} \cos(\phi_{\ell,m})\, 2 \int_0^{\infty} |\lambda|^{-\delta}\ \left| \widehat{\psi}(\lambda) \right|^2\ \mathrm{d}\lambda \right| \leqslant C\ \|\mathbf{\Omega}\|\ 2^{j\,(d_\ell + d_m - \beta)}.$$

It is straightforward that parameters $\{\Omega_{\ell,m}, \phi_{\ell,m}\}$ are not identifiable. Estimation can be derived in the case of a parametric phase, typically $\phi_{\ell,m} = \frac{\pi}{2}(d_\ell - d_m)$ (see Achard and Gannaz [2016]).

In the case of quasi-analytic wavelets, the imaginary part no longer vanishes. The control of quasi-analyticity, given by Theorem 2, leads to the following result.

**Proposition 4.** *Let* $\mathbf{X}$ *be a p-multivariate long range dependent process with long memory parameters* $d_1, \ldots, d_p$ *with normalized spectral density satisfying (M-1)–(M-2). Consider* $\{W_{j,k}(\ell),\ (j,k) \in \mathbb{Z},\ \ell = 1, \ldots, p\}$ *the wavelet coefficients obtained with CFW-C(M,L) filters. Suppose that (C-a) hold and that* $L$ *goes to infinity, with* $L2^{-2j} \to 0$ *when* $j$ *goes to infinity. Then, for all*

$(\ell, m) \in \{1, \dots, p\}^2$,

$$\left| 2^{-j(d_\ell + d_m)} \operatorname{Cov}(W_{j,k}(\ell), W_{j,k}(m)) - \Theta_{\ell,m}\, 4 \int_0^\infty |\lambda|^{-d_\ell - d_m} \left| \widehat{\psi}_h(\lambda) \right|^2 \mathrm{d}\lambda \right| \leqslant C_2 \left( 2^{-j\beta} + L 2^{-2j} + L^{-(2M - d_\ell - d_m + 1)} \right), \quad (9)$$

*where $C_2$ is a constant only depending on $M$ and $C_f, \beta, \|\mathbf{\Omega}\|, \{d_\ell, \ell = 1, \dots, p\}$.*

The proof is given in Appendix.

Convergence (9) can be written as follows: when $2^{-j\beta} + L 2^{-2j} + L^{-1} \to 0$, for all $(\ell, m) \in \{1, \dots, p\}^2$,

$$\lim_{j \to \infty} 2^{-j\,(d_\ell + d_m)} \operatorname{Cov}(W_{j,k}(\ell), W_{j,k}(m)) = G_{\ell,m}, \quad (10)$$

$$\text{with } G_{\ell,m} = \Theta_{\ell,m}\, K(d_\ell + d_m) \;\text{ and }\; K(\delta) = 4 \int_0^\infty |\lambda|^{-\delta} \left| \widehat{\psi}_h(\lambda) \right|^2 \mathrm{d}\lambda\,.$$

Common-factor wavelets, as stated by Proposition 4, have the ability of recovering simultaneously the magnitude and the phase. Observe that for real wavelets the rate of convergence of the limit above is $2^{j\beta}$. The rate with CFW-C(M,L) depends on $L$.

The specificity of CFW-C(M,L) filters is that the quality of the analyticity approximation is based only on parameter $L$, as written in Proposition 4. Nevertheless, if we want to have an approximation with the same quality than the one obtained with real wavelets, the choice of $L$ is more constrained. This tradeoff is due to the fact that the greater $L$, the better analyticity approximation, but the larger the length of the wavelets support. In practice, due to numerical instability, choosing high values (*i.e.* $\geqslant 8$) is not tractable. As the simulations developed in Section 5 show, nevertheless, the results have a good quality even with a smaller value of $L$.

### 3.3 Quality of approximation

To evaluate empirically the accuracy of the approximation, let us compare the empirical covariances of the example of Section 3.1 to the approximation of Proposition 3. Figure 2 displays the sample covariance of wavelet coefficients, respectively with real Daubechies filters with $M = 4$, CFW-PR(4,4) and CFW-C(4,4) filters. Observe that the covariance term is complex, and only the

magnitude is displayed in Figure 2.

Figure 2 shows the difference between our theoretical findings given in Proposition 3 and the simulations for both CFW-PR(M,L) and CFW-(M,L). To better evaluate the quality of the approximation with CFW-C(M,L) filters, let us display the same figure without the first scale. Figure 3 shows that indeed the approximation is improving when the scale $j$ increases. Nevertheless, the difference between results obtained with simulations at first scales (corresponding to the highest frequencies) and the approximation given in Proposition 3 is bigger using CFW-C(4,4) filters in comparison to Daubechies and CFW-PR(4,4) filters. Hence, the lowest scale used in estimation may be taken higher with CFW-C(M,L) filters. This choice may reduce the bias but increase the variance.

# 4 Estimation

Let $j_1 \geqslant j_0 \geqslant 1$ be respectively the upper and the lower resolution levels used in the estimation procedure. The estimation is based on the vectors of wavelet coefficients $\{\mathbf{W}_{j,k},\, j_0 \leqslant j \leqslant j_1,\, k \in \mathbb{Z}\}$. The total number of non-zero coefficients used for estimation is then $n = \sum_{j=j_0}^{j_1} n_j$. Note that if $2^{-j}L$ is bounded and $2^{-j}N \to \infty$, then $n_j$ is equivalent to $2^{-j}N$ when $j$ goes to infinity.

We will make the following assumption:

(C-b) $-M + \beta/2 + 1/2 < d_\ell < M - 1/2 \quad$ for all $\quad \ell = 1, \ldots, p,\ M \geqslant 2$ and $0 < \beta < 2$.

This condition is slightly stronger than (C-a), but it is still not restrictive. It allows to have simpler expressions in the conditions for convergence.

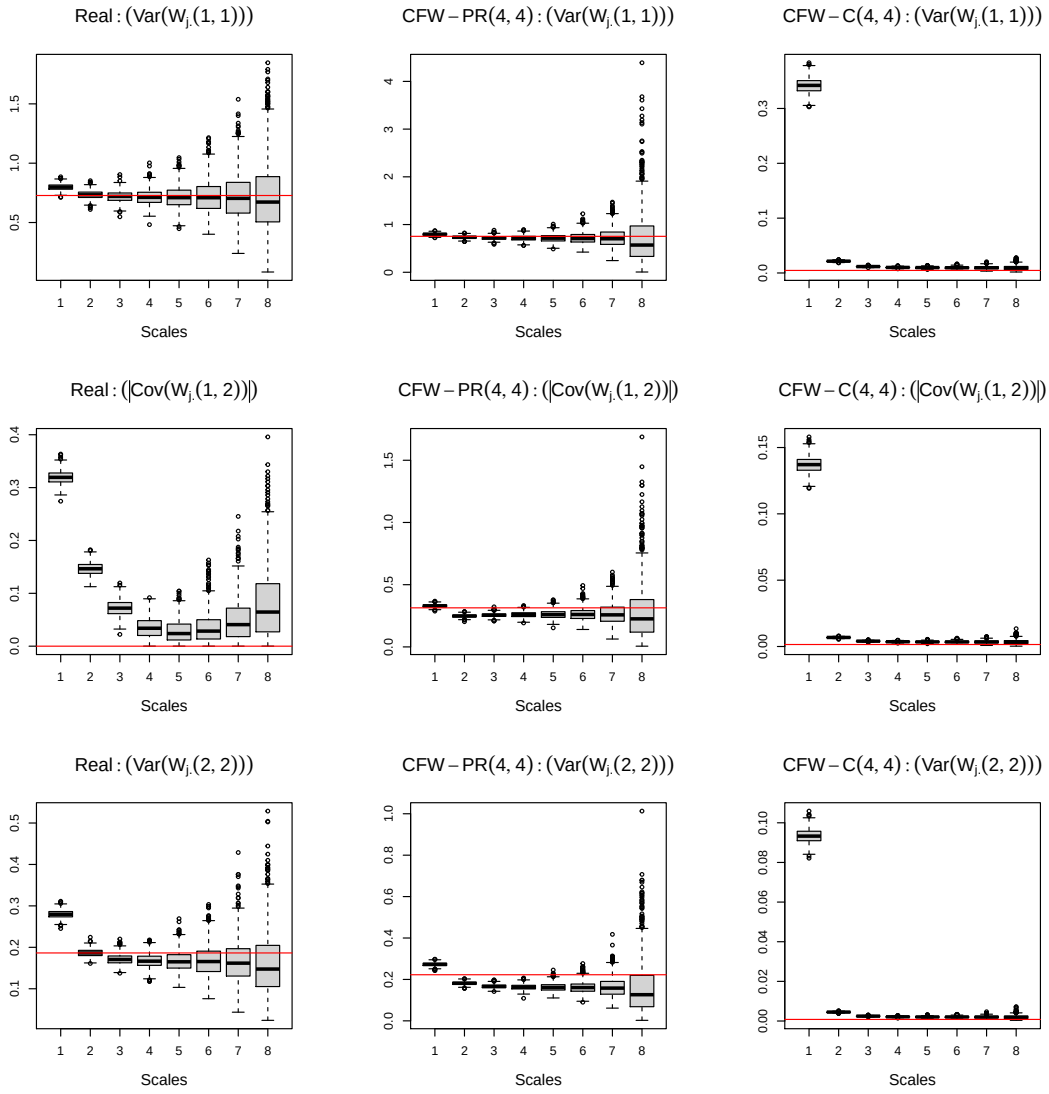


Figure 2: Boxplots of sample correlations between wavelet coefficients at different scales for the bivariate process defined in (8). The first row gives the variance of the first component, the second row gives the magnitude of the covariance and the third row gives the variance of the second component. Each column corresponds to a different wavelet filters, form left to right: Daubechies' real wavelets with M=4, CFW-PR(4,4) and CFW-C(4,4). Horizontal red lines correspond to the approximation given by Proposition 3.

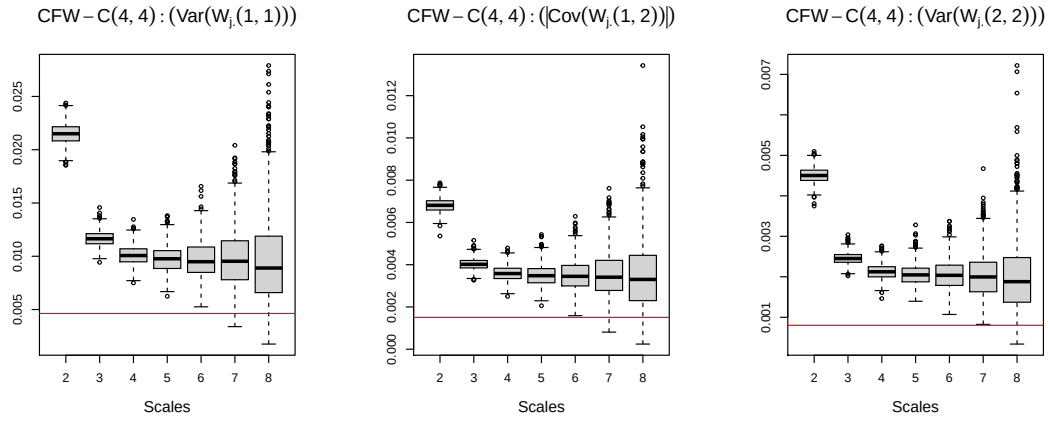


Figure 3: Boxplots of sample correlations between CFW-C(4,4) coefficients at different scales for the bivariate process defined in (8). From left to right, panels correspond respectively to the variance of the first component, the magnitude of the covariance and the variance of the second component. Horizontal red lines correspond to the approximation given by Proposition 3.

### 4.1 Estimation procedure

Based on approximation (10), the objective function $\mathcal{L}(\cdot)$ is defined by the wavelet Whittle approximation of the negative log-likelihood (see Achard and Gannaz [2016])

$$\mathcal{L}(\mathbf{G},\mathbf{d}) = \frac{1}{n}\sum_{j=j_0}^{j_1}\left[n_j \log\det\left(\mathbf{\Gamma}_j(\mathbf{d})\,\mathbf{G}\,\mathbf{\Gamma}_j(\mathbf{d})\right) + \sum_{k=0}^{n_j}\overline{\mathbf{W}_{j,k}}^{T}\left(\mathbf{\Gamma}_j(\mathbf{d})\,\mathbf{G}\,\mathbf{\Gamma}_j(\mathbf{d})\right)^{-1}\mathbf{W}_{j,k}\right]$$

where $\mathbf{\Gamma}_j(\mathbf{d})$ is the diagonal matrix with diagonal entries $2^{-j\,d_1},\dots,2^{-j\,d_p}$, and $\mathbf{G}$ is the matrix with elements $G_{\ell,m} = \Theta_{\ell,m}K(d_\ell + d_m)$, $1 \leqslant, \ell, m \leqslant p$. We can rewrite $\mathcal{L}(\cdot)$ as

$$\mathcal{L}(\mathbf{G},\mathbf{d}) = \frac{1}{n}\sum_{j=j_0}^{j_1}\left[n_j \log\det\left(\mathbf{\Gamma}_j(\mathbf{d})\,\mathbf{G}\,\mathbf{\Gamma}_j(\mathbf{d})\right) + \mathrm{trace}\left(\left(\mathbf{\Gamma}_j(\mathbf{d})\,\mathbf{G}\,\mathbf{\Gamma}_j(\mathbf{d})\right)^{-1}\mathbf{I}(j)\right)\right], \tag{11}$$

where $\mathbf{I}(j) = \sum_{k=0}^{n_j}\mathbf{W}_{j,k}\overline{\mathbf{W}_{j,k}}^{T}$ denotes the (non-normalized) empirical scalogram at scale $j$.

Note that when $\mathbf{G}$ is an Hermitian positive definite matrix, for all $j \geqslant 0$ and for all $\mathbf{d} \in (-0.5,\infty)^p$, $\det(\mathbf{\Gamma}_j(\mathbf{d})\,\mathbf{G}\,\mathbf{\Gamma}_j(\mathbf{d}))$ is real and strictly positive and $\mathrm{trace}\left((\mathbf{\Gamma}_j(\mathbf{d})\,\mathbf{G}\,\mathbf{\Gamma}_j(\mathbf{d}))^{-1}\mathbf{I}_j\right)$ is real. The objective function $\mathcal{L}(\mathbf{G},\mathbf{d})$ is hence well-defined for $\mathbf{G}$ in the set of Hermitian matrices and for all

$\mathbf{d} \in \mathbb{R}^p$, and takes its values in $\mathbb{R}$.

Differentiating expression (11) with respect to the matrix $\mathbf{G}$ yields

$$\frac{\partial \mathcal{L}}{\partial \mathbf{G}}(\mathbf{G}, \mathbf{d}) = \frac{1}{n} \sum_{j=j_0}^{j_1} \left[ n_j \mathbf{G}^{-1} - \mathbf{G}^{-1} \mathbf{\Gamma}_j(\mathbf{d})^{-1} \mathbf{I}(j) \mathbf{\Gamma}_j(\mathbf{d})^{-1} \mathbf{G}^{-1} \right]^T,$$

where the exponent $T$ denotes the transpose operator. Some keys for complex matrix differentiation can be found in Hjorungnes and Gesbert [2007]. Hence, the minimum for fixed $\mathbf{d}$ is attained at

$$\widehat{\mathbf{G}}(\mathbf{d}) = \frac{1}{n} \sum_{j=j_0}^{j_1} \mathbf{\Gamma}_j(\mathbf{d})^{-1} \mathbf{I}(j) \mathbf{\Gamma}_j(\mathbf{d})^{-1}.$$

In Shimotsu [2007] the resulting objective function only depends of $\mathbf{d}$ since the phases are parametric while in Baek et al. [2020] the authors consider a general form of phases. In both Shimotsu [2007] and Baek et al. [2020], with Fourier-based approach, a real matrix $\mathbf{G}(\mathbf{d})$ and complex valued matrices $\mathbf{\Gamma}_j(\mathbf{d})$, including the phases $\left(\phi_{\ell,m}\right)_{\ell,m=,,\ldots,p}$ are considered. $\mathbf{G}(\mathbf{d})$ and $\mathbf{\Gamma}_j(\mathbf{d})$ are estimated in a second step, together with parameter $\mathbf{d}$, minimizing the objective function obtained when replacing $\mathbf{G}$ by $\widehat{\mathbf{G}}(\mathbf{d})$ in (11). Yet, our procedure enables to estimate the magnitude of the correlation even when the phase is equal to $\pi/2$, with imaginary terms in $\mathbf{G}$.

Replacing $\mathbf{G}$ by $\widehat{\mathbf{G}}(\mathbf{d})$, the objective function is defined by

$$R(\mathbf{d}) := \mathcal{L}(\widehat{\mathbf{G}}(\mathbf{d}), \mathbf{d}) - p = \log\det(\widehat{\mathbf{G}}(\mathbf{d})) - \frac{1}{n} \sum_{j=j_0}^{j_1} n_j \log(\det\left(\mathbf{\Gamma}_j(\mathbf{d}) \mathbf{\Gamma}_j(\mathbf{d})\right).$$

Since $\mathbf{\Gamma}_j(\mathbf{d}) = \mathrm{Diag}\left(2^{-j\mathbf{d}}\right)$, we obtain

$$R(\mathbf{d}) = \log\det(\widehat{\mathbf{G}}(\mathbf{d})) + 2\log(2) \left( \frac{1}{n} \sum_{j=j_0}^{j_1} j n_j \right) \left( \sum_{\ell=1}^{p} d_\ell \right).$$

The vector of the long-memory parameters $\mathbf{d}$ is estimated by $\widehat{\mathbf{d}} = \mathrm{argmin}_{\mathbf{d}}\, R(\mathbf{d})$.

In a second step of estimation we define $\widehat{\mathbf{G}}(\widehat{\mathbf{d}})$, estimator of $\mathbf{G}$. And we recover an estimation of $\mathbf{\Theta}$ by

$$\widehat{\Theta}_{\ell,m} = \widehat{G}_{\ell,m}(\widehat{\mathbf{d}}) \,/\, K(\widehat{d}_\ell + \widehat{d}_m).$$

### 4.2 Asymptotic convergence

Following Moulines et al. [2008] and Achard and Gannaz [2016], we introduce an additional condition on the variance of the scalogram $\{\mathbf{I}(j)\}$. Examples of linear processes satisfying this condition can be found in Proposition 4 of Achard and Gannaz [2016].

**Condition (C)**

$$\text{For all } \ell, m = 1, \dots, p, \quad \sup_n \sup_{j \geqslant 0} \frac{|\mathrm{Var}\left(I_{\ell,m}(j)\right)|}{n_j 2^{2j(d_\ell + d_m)}} < \infty\,.$$

Let $\mathbf{d}^0$, $\mathbf{G}^0$ and $\mathbf{\Theta}^0$ denote the true values of the parameters. The consistency of the estimators can be established as in Achard and Gannaz [2016].

**Theorem 5.** *Suppose that (C-b) and assumptions of Proposition 4 hold. Assume that Condition (C) is satisfied. Let $j_0$ and $j_1$ be such that $2^{-j_0\beta} + N^{-1/2}2^{j_0/2} \to 0$ and $j_0 < j_1 \leqslant j_N$ with $j_N = \max\{j, n_j \geqslant 1\}$.*

*Let $j_0$ and $j_1$ satisfy $\log(N)^2(2^{-j_0\beta} + N^{-1/2}2^{j_0/2}) \to 0$ and $j_0 < j_1 \leqslant j_N$.*

*Consider CFW-C(M,L) filters with $M \geqslant 2$ and $L2^{-j_0} + \log(N)^3\, L^{-1}) \to 0$.*

*Then , $\forall (\ell, m) \in \{1, \dots, p\}^2$,*

$$\begin{aligned}
\widehat{\mathbf{d}} - \mathbf{d}^0 &= O_{\mathbb{P}}(L2^{-2j_0} + \log(N)L^{-1} + 2^{-j_0\beta} + N^{-1/2}2^{j_0/2}),\\
\widehat{G}_{\ell,m}(\widehat{\mathbf{d}}) - G_{\ell,m}(\mathbf{d}^0) &= O_{\mathbb{P}}(\log(N)(L2^{-2j_0} + \log(N)L^{-1} + 2^{-j_0\beta} + N^{-1/2}2^{j_0/2})),\\
\widehat{\Theta}_{\ell,m} - \Theta^0_{\ell,m} &= O_{\mathbb{P}}(\log(N)(L2^{-2j_0} + \log(N)L^{-1} + 2^{-j_0\beta} + N^{-1/2}2^{j_0/2})).
\end{aligned}$$

*Taking $2^{j_0} = N^{1/(1+2\beta)}$ and $N\, 2^{j_0\beta} \leqslant L \leqslant 2^{j_0}$,*

$$\widehat{\mathbf{d}} - \mathbf{d}^0 = O_{\mathbb{P}}(N^{-\beta/(1+2\beta)}).$$

Elements of proof are given in Appendix.

The condition on the scales used in estimation, $\log(N)^2(2^{-j_0\beta} + N^{-1/2}2^{j_0/2}) \to 0$, states that

highest frequencies should be removed. As explained in Achard and Gannaz [2019], the number of scales that should be removed depends on the short-range dependence through $\beta$. In practice, $j_0$ can be chosen applying a bootstrap procedure on the time series, see Achard and Gannaz [2019].

Next, the parameters $M$ and $L$ in CFW-C(M,L) filters are subject to the conditions (C-b) and $L2^{-j_0} + \log(N)^3 L^{-1}) \to 0$. Condition (C-b) depends only on $M$. It is not restrictive and very similar to the one given in Achard and Gannaz [2016] with real filters. The parameter $L$ quantifies the quality of the analytic approximation of CFW-C(M,L) filters. The condition $\log(N)^3 L^{-1} \to 0$ ensures that $L$ is high enough so that this approximation is satisfactory. Alternatively, $L$ should not be too high, condition $L\,2^{-j_0} \to 0$ ensures that the size of the support of the wavelets remains reasonable. As discussed in Section 5, in practice, the choice of $L$ is not critical, but this condition influences the choice of $j_0$. It must be higher than the usual choice for real filters. This also appears in the discussion in Section 3.3, where it can be seen that the behavior of the wavelet coefficients at first scales differs from the other scales.

### 4.3 Asymptotic normality

A useful result in estimation is the asymptotic normality. For real wavelet-based local Whittle estimation, in multivariate context, it has been studied by Gannaz [2020]. The proof of the latter can be extended to common-factor wavelets.

Let us introduce an additional assumption on the process $\mathbf{X}$.

(M-3) There exists a sequence $\{\mathbf{A}^{(\mathbf{D})}(u)\}_{u\in\mathbb{Z}}$ in $\mathbb{R}^{p\times p}$ such that $\sum_{u\in\mathbb{Z}} \max_{a,b=1,\dots,p} |A^{(\mathbf{D})}_{a,b}(u)|^2 < \infty$ and

$$\forall t \in \mathbb{Z},\ \ \left(\mathbb{L}^{D_a} X_a(t)\right)_{a=1,\dots,p} = \sum_{u\in\mathbb{Z}} \mathbf{A}^{(\mathbf{D})}(t+u)\varepsilon(u)$$

with $\varepsilon(t)$ weak white noise process, in $\mathbb{R}^p$. Let $\mathcal{F}_{t-1}$ denote the $\sigma$-field of events generated by $\{\varepsilon(s),\, s \leqslant t-1\}$. Assume that $\varepsilon$ satisfies $\mathbb{E}[\varepsilon(t)|\mathcal{F}_{t-1}] = 0$, $\mathbb{E}[\varepsilon_a(t)\varepsilon_b(t)|\mathcal{F}_{t-1}] = \mathbb{1}_{a=b}$ and $\mathbb{E}[\varepsilon_a(t)\varepsilon_b(t)\varepsilon_c(t)\varepsilon_d(t)|\mathcal{F}_{t-1}] = \mu_{a,b,c,d}$ with $|\mu_{a,b,c,d}| \leqslant \mu_\infty < \infty$, for all $a,b,c,d = 1,\dots,p$. For all $(a,b) \in \{1,\dots,p\}^2$, for all $\lambda \in \mathbb{R}$, the sequence $(2^{-j\,d_a}|A^{(\mathbf{D})*}_{a,b}(2^{-j}\lambda)|)_{j\geqslant 0}$ is convergent as $j$ goes to infinity.

For $u \geqslant 0$, $(\delta_1, \delta_2) \in (-\alpha, M)^2$, define $\widetilde{I}_u(\delta_1, \delta_2)$

$$\widetilde{I}_u(\delta_1, \delta_2) = \frac{2\pi}{K(\delta_1)K(\delta_2)} \int_{-\pi}^{\pi} \overline{\widetilde{D}_{u,\infty}(\lambda; \delta_1)} \widetilde{D}_{u,\infty}(\lambda; \delta_2) \, \mathrm{d}\lambda \,,$$

where $D_{u,\infty}(\lambda; \delta)$ is an approximation of the cross-spectral density between wavelet coefficients $\{\mathbf{W}(j,k),\, k \in \mathbb{Z}\}$ and $\{\mathbf{W}(j+u, 2^u k + \tau),\, \tau = 0, \ldots, 2^u - 1,\, k \in \mathbb{Z}\}$,

$$D_{u,\tau}(\lambda; \delta) = \sum_{t \in \mathbb{Z}} |\lambda + 2t\pi|^{-\delta} \overline{\widehat{\psi}(\lambda + 2t\pi)} \, 2^{u/2} \widehat{\psi}(2^u(\lambda + 2t\pi)) \, \mathrm{e}^{-\mathrm{i}2^u \tau(\lambda + 2t\pi)} \,,$$

$$\widetilde{D}_{u,\infty}(\lambda; \delta) = \sum_{\tau=0}^{2^{-u}-1} D_{u,\tau}(\lambda; \delta) \,.$$

We introduce

$$\mathcal{I}^d_\Delta(\delta_1, \delta_2) = \frac{2}{\kappa_\Delta} \widetilde{I}_0(\delta_1, \delta_2)$$
$$+ \frac{2}{\kappa^2_\Delta} \sum_{u=1}^{\Delta} (2^{u\delta_1} + 2^{u\delta_2}) \, 2^{-u} \frac{2 - 2^{-\Delta+u}}{2 - 2^{-\Delta}} ((u + \eta_{\Delta-u} - \eta_\Delta)(\eta_{\Delta-u} - \eta_\Delta) + \kappa_{\Delta-u}) \, \widetilde{I}_u(\delta_1, \delta_2)$$
$$\text{if } \Delta < \infty,$$

$$\mathcal{I}^d_\infty(\delta_1, \delta_2) = \widetilde{I}_0(\delta_1, \delta_2) + \sum_{u=1}^{\infty} (2^{u\delta_1} + 2^{u\delta_2}) \, 2^{-u} \, \widetilde{I}_u(\delta_1, \delta_2) \,, \quad \text{if } \Delta = \infty.$$

Define also

$$\mathbf{G} \,.\, \mathcal{I}^{\mathbf{d}} \,.\, \mathbf{G}(\Delta) = \mathrm{Diag}\left(\mathrm{vec}\left(\mathbf{G}^0\right)\right) \left(\mathcal{I}^d_\Delta(d^0_a + d^0_b, d^0_{a'} + d^0_{b'})_{(a,b),(a',b') \in \{1,\ldots,p^2\}}\right) \mathrm{Diag}\left(\mathrm{vec}\left(\mathbf{G}^0\right)\right) \,. \quad (12)$$

Additionally, let us denote

$$\mathcal{I}^G_\Delta(\delta_1, \delta_2) = \widetilde{I}_0(\delta_1, \delta_2) + \sum_{u=1}^{\Delta} (2^{u\delta_1} + 2^{u\delta_2}) 2^{-u} \, \frac{2 - 2^{-\Delta+u}}{2 - 2^{-\Delta}} \, \widetilde{I}_u(\delta_1, \delta_2) \qquad \text{if } \Delta < \infty,$$

$$\mathcal{I}^G_\infty(\delta_1, \delta_2) = \widetilde{I}_0(\delta_1, \delta_2) + \sum_{u=1}^{\infty} (2^{u\delta_1} + 2^{u\delta_2}) 2^{-u} \, \widetilde{I}_u(\delta_1, \delta_2) \qquad \text{if } \Delta = \infty.$$

Let us also define

$$\mathbf{G}\,.\,\mathcal{I}^{\mathbf{G}}\,.\,\mathbf{G}(\Delta) = \mathrm{Diag}\left(\mathrm{vec}\big(\mathbf{G}^0\big)\right)\left(\mathcal{I}^G_\Delta(d^0_a + d^0_b, d^0_{a'} + d^0_{b'})_{(a,b),(a',b')\in\{1,\dots,p^2\}}\right)\mathrm{Diag}\left(\mathrm{vec}\big(\mathbf{G}^0\big)\right). \quad (13)$$

The asymptotic normality of the estimator of the long-memory parameters is established by our next theorem.

**Theorem 6.** *Suppose that conditions of Theorem 5 are satisfied and that assumption (M-3) hold. Let $j_0 < j_1 \leqslant j_N$ with $j_N = \max\{j, n_j \geqslant 1\}$ such that*

$$j_1 - j_0 \to \Delta \in \{1, \dots, \infty\},\ \log(N)^2(N2^{-j_0(1+2\beta)} + N^{-1/2}2^{j_0/2}) \to 0.$$

*Define $n = \sum_{j=j_0}^{j_1} n_j$.*

*Consider CFW-C(M,L) filters with $M \geqslant 2$ and $\log(N)^2\, N^{1/2}2^{-j_0/2}(L2^{-j_0} + L^{-1}) \to 0$.*

*Then,*

- *$\sqrt{n}(\widehat{\mathbf{d}} - \mathbf{d}^0)$ converges in distribution to a centered Gaussian distribution with a variance equal to*

$$\mathbf{V}^{(\mathbf{d})}(\Delta) = \frac{1}{2\,\log(2)^2}(\mathbf{G}^{0-1} \circ \mathbf{G}^0 + \mathbf{I}_p)^{-1}\,\boldsymbol{\Upsilon}(\Delta)\,(\mathbf{G}^{0-1} \circ \mathbf{G}^0 + \mathbf{I}_p)^{-1},$$

  *where $\mathbf{I}_p$ is the identity matrix in $\mathbb{R}^{p\times p}$ and with entry $(a, a')$ of $\boldsymbol{\Upsilon}^{(\Delta)}$, for $(a, a') \in \{1, \dots, p\}^2$, given by*

$$\Upsilon_{a,a'}(\Delta) = \sum_{b,b'=1,\dots,p} (G^{0-1})_{a,b}(G^{0-1})_{a',b'}\left(\mathbf{G}\,.\,\mathcal{I}^{\mathbf{d}}\,.\,\mathbf{G}_{(a,a'),(b,b')}(\Delta) + \mathbf{G}\,.\,\mathcal{I}^{\mathbf{d}}\,.\,\mathbf{G}_{(a,b'),(a',b)}(\Delta)\right)$$

  *where quantities $\mathbf{G}\,.\,\mathcal{I}^{\mathbf{d}}\,.\,\mathbf{G}(\Delta)$ are defined by (12).*

- *$\mathrm{vec}\left(\sqrt{n}\left(\widehat{\mathbf{G}}(\widehat{\mathbf{d}}) - \mathbf{G}^0\right)\right)$ converges in distribution to a centered Gaussian distribution with a variance equal to $\mathbf{V}^{\mathbf{G}(\Delta)}$, with*

$$V^{(\mathbf{G})}_{(a,b),(a',b')}(\Delta) = \mathbf{G}\,.\,\mathcal{I}^{\mathbf{G}}\,.\,\mathbf{G}_{(a,a'),(b,b')}(\Delta) + \mathbf{G}\,.\,\mathcal{I}^{\mathbf{G}}\,.\,\mathbf{G}_{(a,b'),(a',b)}(\Delta)$$

  *where quantities $\mathbf{G}\,.\,\mathcal{I}^{\mathbf{G}}\,.\,\mathbf{G}(\Delta)$ are defined by (13).*

The proof is very similar to the one of Gannaz [2020]. Some points are detailed in Appendix.

As detailed in Gannaz [2020], these results allow to build hypothesis tests on the long-memory parameters and on the long-run covariance. The conditions on the hyperparameters are very similar to the ones obtained for the consistency, with only an additional $\log(N)$ term.

# 5 Simulation study

In this section, we verify the accuracy of the covariance approximation given in Proposition 4 and the consistency of the parameters estimates provided in Proposition 5 on simulated data. We consider 1000 Monte-Carlo simulations of bivariate long-memory processes $\mathbf{X}$ observed at $\mathbf{X}(1), \dots, \mathbf{X}(N)$ with $N = 2^{12}$. For each process, we compute the wavelet coefficients using CFW-PR(4,4) and CFW-PR(4,4) filters. We also compare the quality of estimation of parameters $\mathbf{d}$ to the one given by real wavelets, namely Daubechies' wavelets with 4 vanishing moments. We choose here not to compare with Fourier-based procedure since we want to consider non-stationary cases. For a comparison between Fourier-based and wavelet-based local Whittle estimations in such a framework, we refer the reader to Achard and Gannaz [2019].

The estimated parameters are $\mathbf{d} = (d_1, d_2)$, the magnitude of the long-run covariance $\mathbf{\Omega}$, the phase $\phi = \phi_{1,2}$ and the long-run correlation $\rho = \frac{\Omega_{1,2}}{\sqrt{\Omega_{1,1}\Omega_{2,2}}}$. For each parameter, we will evaluate the quality of estimation by the bias, the standard deviation (std) and the Root Mean Squared Error, $\text{RMSE} = \sqrt{\text{bias}^2 + \text{std}^2}$.

Two models are considered: models admitting a linear representation called ARFIMA and multivariate fractional Brownian motions (mFBM).

## 5.1 ARFIMA models

We first provide an estimation example on linear time series. Let $\xi$ be a $p$-dimensional white noise with $\mathbb{E}[\xi(t) \mid \mathcal{F}_{t-1}] = 0$ and $\mathbb{E}[\xi(t)\xi(t)^T \mid \mathcal{F}_{t-1}] = \Sigma$ with $\Sigma$ positive definite, where $\mathcal{F}_{t-1}$ is the $\sigma$-field generated by $\{\xi(s),\, s < t\}$. The spectral density of $\xi$ satisfies $f_\xi(\lambda) = \mathbf{\Sigma}$.

Let $(\mathbf{A}_k)_{k\in\mathbb{N}}$ be a sequence in $\mathbb{R}^{p\times p}$ with $\mathbf{A}_0$ the identity matrix and $\sum_{k=0}^{\infty}\|\mathbf{A}_k\|^2<\infty$. Let $\mathbf{A}(\cdot)$ be the discrete Fourier transform of the sequence, $\mathbf{A}(\lambda)=\sum_{k=0}^{\infty}\mathbf{A}_k e^{ik\lambda}$. We assume that $|\mathbf{A}(\mathbb{L})|$ has all its roots outside the unit circle so that $\mathbf{A}(\cdot)^{-1}$ is defined and smooth on $\mathbb{R}$. We also define $(B_k)_{k\in\mathbb{N}}$ to be a sequence in $\mathbb{R}^{p\times p}$ with $B_0$ the identity matrix and $\sum_{k=0}^{\infty}\|B_k\|^2<\infty$. Let $\mathbf{B}(\cdot)$ be the discrete Fourier transform of the sequence, $\mathbf{B}(\lambda)=\sum_{k=0}^{\infty}\mathbf{B}_k e^{ik\lambda}$.

Let us define the process $\mathbf{X}$

$$\mathbf{A}(\mathbb{L})\,\mathrm{Diag}\left((\mathbb{1}-\mathbb{L})^{\mathbf{d}}\right)\,\mathbf{X}(t)=\mathbf{B}(\mathbb{L})\xi(t). \tag{14}$$

The spectral density satisfies

$$\boldsymbol{f}(\lambda)=(1-e^{-i\lambda})^{-\mathbf{d}}\mathbf{A}(e^{-i\lambda})^{-1}\mathbf{B}(e^{-i\lambda})f_\xi(\lambda)\mathbf{B}(e^{i\lambda})^T\mathbf{A}(e^{i\lambda})^{T^{-1}}(1-e^{i\lambda})^{-\mathbf{d}}.$$

In particular

$$f_{\ell,m}(\lambda)\sim_{\lambda\to 0^+} G_{\ell,m}e^{-i\pi/2(d_\ell-d_m)}\lambda^{-(d_\ell+d_m)}\ ,$$

with $\mathbf{G}=\mathbf{A}(0)^{-1}\mathbf{B}(0)f_\xi(\lambda)\mathbf{B}(0)^T\mathbf{A}(0)^{T^{-1}}=\mathbf{A}(0)^{-1}\mathbf{B}(0)\mathbf{\Sigma}\mathbf{B}(0)^T\mathbf{A}(0)^{T^{-1}}$ a real valued matrix. Condition (M-2) is satisfied with $\beta=\min_\ell(d_\ell)$. In this case $\boldsymbol{f}(0^+)=\boldsymbol{f}(0^-)$.

This corresponds to Model A of Lobato [1997]. Note that this model satisfies the definition of LRD processes of Kechagias and Pipiras [2014] when $f_\xi(\lambda)\sim_{\lambda\to 0^+}\mathbf{\Sigma}$, this is satisfied when $\xi$ is a white-noise process.

Following (14), we have simulated $\{\mathbf{X}(1),\dots\mathbf{X}(N)\}$ in (14) with $N=2^{12}$, null $\mathbf{A}_k$ and $\mathbf{B}_k$ for $k\geqslant 0$. That is, there is no short-range terms in the model. We consider three sets of values for $\mathbf{d}$, $\mathbf{d}\in\{(0.2,0.2),\ (0.2,0.4),\ (0.2,0.8)\}$. Matrix $\mathbf{\Sigma}$ is equal to $\begin{pmatrix}1 & \rho\\ \rho & 1\end{pmatrix}$, with $\rho=0.8$. The phase is equal to $\pi(d_1-d_2)/2$ which is respectively equal to $0,\pi/10,3\pi/10$. Simulations were done using R package *multiwave* [Achard and Gannaz, 2015].

*Remark.* The objective of this simulation part is to compare estimations based respectively on real and complex wavelet filters. Hence, only simulations with null $\mathbf{A}_k$ and $\mathbf{B}_k$ are considered. We refer to Achard and Gannaz [2019] for results with non null AR and MA parts, with real wavelet filters.

Figure 4 displays the boxplots of the correlations between the wavelet coefficients obtained by

CFW-PR(4,4) filter at different scales. It illustrates that the approximation of Proposition 4 is accurate, especially for the high scales (lowest frequencies), even if it has not been established theoretically for such filters. The figure shows that the approximation of Proposition 4 is slightly more accurate for the real part of wavelet correlations than for the imaginary part.

(a) $\mathbf{d} = (0.2, 0.2)$ (b) $\mathbf{d} = (0.2, 0.4)$ (c) $\mathbf{d} = (0.2, 0.8)$

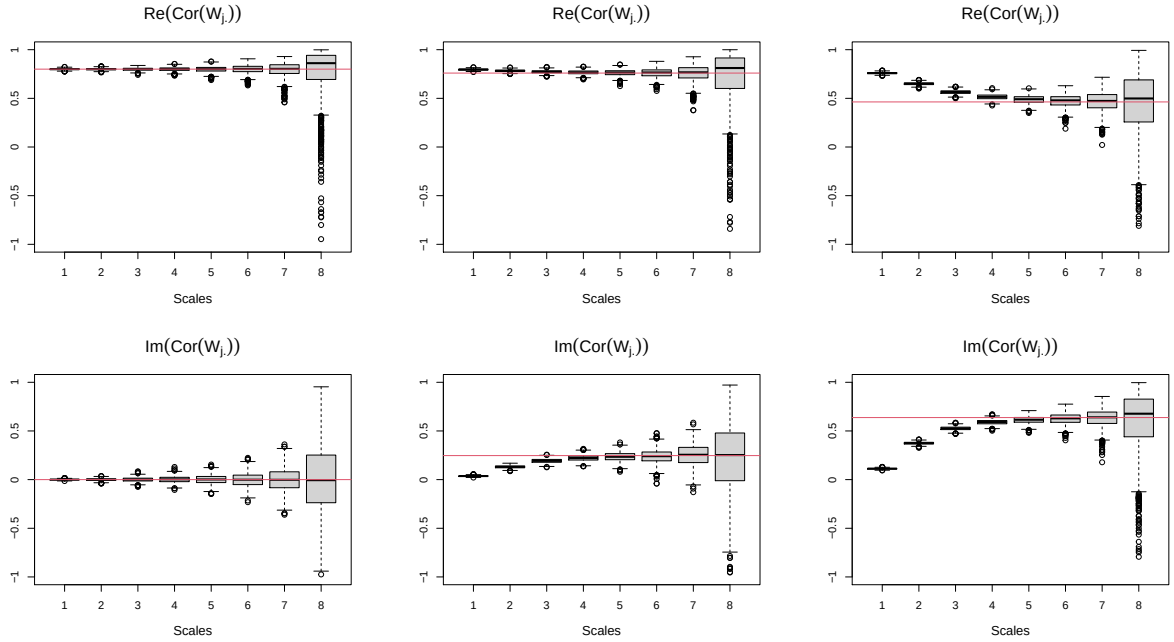


Figure 4: Boxplots of correlation between CFW-PR(4,4) coefficients at different scales for ARFIMA processes. First row gives the real part of the correlations and second row gives the imaginary part. Each column corresponds to a given value of parameter $\mathbf{d}$. Horizontal red lines correspond to the approximation given by Proposition 4, $\rho\cos(\phi)r_K$ for the real part and $\rho\sin(\phi)r_K$ for the imaginary part, with $r_K = K(d_1+d_2)/\sqrt{K(2\,d_1)K(2\,d_2)}$.

Results for the estimation of the long-memory parameters $\mathbf{d}$ are displayed respectively in Table 1 for CFW-PR(4,4) filter and in Table 2 for CFW-C(4,4) filter. Based on Achard and Gannaz [2019], the only hyperparameter that needs to be chosen is the minimal scale $j_0$. The example in Section 3.3 illustrates that the behaviors of Daubechies' wavelets and CFW-PR(M,L) are very similar with respect to the scales. Hence, for both of them, we consider $j_0 = 1$ when both components of the time series are stationary, and $j_0 = 2$ when a component is not stationary. These choices are motivated by previous studies done in Achard and Gannaz [2019].

Table 1 shows that the estimation of $\mathbf{d}$ with CFW-PR(M,L) filters is good and similar to the Daubechies' real wavelet-based estimation.

Concerning CFW-C(M,L) estimation, it has been shown in Section 3.3 that the approximation of Proposition 3 is not accurate for the first scales. Therefore, results are displayed with $j_0 = 4$ in each case, to reduce the bias of the long-run covariance estimation. Since less scales are available in the procedure, the variance increases and the quality of estimation is lower than the one based on CFW-PR(M,L). Table 2 illustrates that the quality of estimation with CFW-C(M,L) filters present a good accuracy, however lower than the one with CFW-PR(M,L) filters.

| $\mathbf{d}$ | bias | std | RMSE | PR/Real |
|---|---|---|---|---|
| 0.2 | -0.0066 | 0.0165 | 0.0178 | 1.0375 |
| 0.2 | -0.0069 | 0.0156 | 0.0171 | 1.0228 |
| 0.2 | -0.0080 | 0.0165 | 0.0183 | 1.1537 |
| 0.4 | -0.0138 | 0.0159 | 0.0211 | 1.1888 |
| 0.2 | -0.0094 | 0.0255 | 0.0272 | 0.9072 |
| 0.8 | -0.0145 | 0.0270 | 0.0307 | 1.3319 |

Table 1: Results for the estimation of long-memory parameters $\mathbf{d}$ with CFW-PR(4,4) filters using ARFIMA processes. $j_0 = 1$ for $\mathbf{d} \in \{(0.2, 0.2), (0.2, 0.4)\}$ and $j_0 = 2$ for $\mathbf{d} \in \{(0.2, 0.8)\}$. PR/Real denotes the ratio between the RMSE given by CFW-PR(4,4) filter and the RMSE given by Daubechies' real filter.

| $\mathbf{d}$ | bias | std | RMSE | C/PR |
|---|---|---|---|---|
| 0.2 | -0.0152 | 0.0401 | 0.0429 | 2.4155 |
| 0.2 | -0.0145 | 0.0392 | 0.0418 | 2.4504 |
| 0.2 | -0.0158 | 0.0394 | 0.0425 | 2.3229 |
| 0.4 | -0.0150 | 0.0385 | 0.0413 | 1.9626 |
| 0.2 | -0.0166 | 0.0397 | 0.0430 | 1.5837 |
| 0.8 | -0.0163 | 0.0390 | 0.0422 | 1.3762 |

Table 2: Results for the estimation of long-memory parameters $\mathbf{d}$ with CFW-C(4,4) filters using ARFIMA processes. $j_0 = 4$. C/PR denotes the ratio between the RMSE given by CFW-C(4,4) filter and the RMSE given by CFW-PR(4,4) filter.

Table 3 and Table 4 give the results for the covariance structure estimation, respectively for CFW-

PR(4,4) filter and for CFW-C(4,4) filter. The quality of estimation using CFW-PR(4,4) improves significantly the results obtained with real wavelets. This is not the case for CFW-C(4,4) filters where the quality of estimation deteriorates.

With CFW-PR(M,L) filters, the results for the phase parameter $\phi$ are less satisfactory. A bias term can be observed when the phase $\phi$ increases. This term has for example an order of $\pi/10$ when estimating the phase of $3\pi/10$ corresponding to the case $\mathbf{d} = (0.2, 0.8)$. Yet our results seem comparable to the one obtained in Baek et al. [2020]. Estimating the phase is challenging. Interestingly, the quality of the estimation based on CFW-C(M,L) filters is much stable with respect to the phase values. Indeed, even when the phase value increases, the bias remains constant.

| **d** | | bias | std | RMSE | PR/Real |
|---|---|---|---|---|---|
| ( 0.2 , 0.2 ) | $\Omega_{1,1}$ | 0.0087 | 0.0227 | 0.0243 | 0.5089 |
| | $\Omega_{1,2}$ | 0.0076 | 0.0204 | 0.0218 | 0.5509 |
| | $\Omega_{2,2}$ | 0.0092 | 0.0229 | 0.0247 | 0.5184 |
| | correlation | 4e-04 | 0.0057 | 0.0057 | 0.9702 |
| | phase | 1e-04 | 0.0081 | 0.0081 | . |
| ( 0.2 , 0.4 ) | $\Omega_{1,1}$ | 0.0082 | 0.0229 | 0.0243 | 0.4919 |
| | $\Omega_{1,2}$ | 0.0028 | 0.0206 | 0.0208 | 0.4883 |
| | $\Omega_{2,2}$ | 0.0146 | 0.0231 | 0.0273 | 0.3142 |
| | correlation | -0.0063 | 0.0058 | 0.0085 | 0.6079 |
| | phase | 0.1879 | 0.0083 | 0.1880 | . |
| ( 0.2 , 0.8 ) | $\Omega_{1,1}$ | -0.0184 | 0.0365 | 0.0409 | 0.8883 |
| | $\Omega_{1,2}$ | -0.086 | 0.0292 | 0.0908 | 0.4163 |
| | $\Omega_{2,2}$ | -0.1145 | 0.0352 | 0.1198 | 2.1877 |
| | correlation | -0.0342 | 0.0083 | 0.0351 | 0.1788 |
| | phase | 0.2790 | 0.017 | 0.2795 | . |

Table 3: Results for the estimation of matrices $\mathbf{\Theta}$ with CFW-PR(4,4) filters on ARFIMA processes. $j_0 = 1$ for $\mathbf{d} \in \{(0.2, 0.2), (0.2, 0.4)\}$ and $j_0 = 2$ for $\mathbf{d} \in \{(0.2, 0.8)\}$.

We can be observe that estimation with CFW-C(4,4) filters has a lower bias and higher variance than estimation with CFW-PR(4,4) filters. The higher $j_0$, the lower the bias, but the higher the variance. For the estimation of $\mathbf{d}$, the main difficulty is to control the variance depending on the choice of $j_0$. Since the bias for the phase estimation is critical, higher $j_0$ for CFW-PR(4,4) filters can be considered. Moreover, as illustrated in Figure 4, the quality of the approximation of the imaginary part of the correlation is not accurate for the three highest scales. Hence, to ensure a

| **d** | | bias | std | RMSE | C/PR |
|---|---|---|---|---|---|
| ( 0.2 , 0.2 ) | $\Omega_{1,1}$ | 0.0488 | 0.175 | 0.1817 | 7.5165 |
| | $\Omega_{1,2}$ | 0.0367 | 0.1252 | 0.1305 | 6.0138 |
| | $\Omega_{2,2}$ | 0.0451 | 0.1685 | 0.1744 | 7.1674 |
| | correlation | 0.0012 | 0.0173 | 0.0173 | 2.9994 |
| | phase | 8e-04 | 0.0367 | 0.0367 | 4.5335 |
| ( 0.2 , 0.4 ) | $\Omega_{1,1}$ | 0.0469 | 0.1726 | 0.1789 | 7.3547 |
| | $\Omega_{1,2}$ | 0.0124 | 0.1198 | 0.1205 | 5.7810 |
| | $\Omega_{2,2}$ | -0.0086 | 0.1628 | 0.1630 | 5.9735 |
| | correlation | -3e-04 | 0.0172 | 0.0172 | 2.0196 |
| | phase | 0.0225 | 0.0364 | 0.0428 | 0.2278 |
| ( 0.2 , 0.8 ) | $\Omega_{1,1}$ | 0.0476 | 0.1715 | 0.178 | 4.3507 |
| | $\Omega_{1,2}$ | -0.0266 | 0.1156 | 0.1186 | 1.3064 |
| | $\Omega_{2,2}$ | -0.0973 | 0.1499 | 0.1787 | 1.4916 |
| | correlation | -0.0026 | 0.0175 | 0.0177 | 0.5036 |
| | phase | 0.0645 | 0.0357 | 0.0737 | 0.2637 |

Table 4: Results for the estimation of matrices $\mathbf{\Theta}$ with CFW-C(4,4) filters on ARFIMA processes. $j_0 = 4$. C/PR denotes the ratio between the RMSE given by CFW-C(4,4) filter and the RMSE given by CFW-PR(4,4) filter.

good approximation and consequently a good estimation for the correlation and the phase, it seems more appropriate to remove the three first scales and to consider $j_0 = 4$ also for CFW-PR(M,L).

Table 5 and Table 6 display the results when considering $j_0 = 4$, with CFW-PR(M,L), respectively for the long-run dependence parameter $\mathbf{d}$ and for the long-run correlation and the phase. Results for the long-run covariance are omitted, for simplicity. Table 5 shows that even if the bias of $\widehat{\mathbf{d}}$ increases, it has a similar order of magnitude. Interestingly, Table 6 highlights that the quality of estimation of CFW-C(4,4) and CFW-PR(4,4) filters are then very similar. In particular, the bias of the phase and of the correlation decrease, as illustrated in Table 3. With $j_0 = 4$, the RMSE is significantly lower than Baek et al. [2020]'s Fourier-based local Whittle estimator. This approach, hence, improves significantly the estimation of the phase with respect to others procedures. Observe that it is nevertheless sensitive to the choice of $j_0$.

| **d** | bias | std | RMSE | PR/Real |
|---|---|---|---|---|
| 0.2 | -0.002 | 0.0412 | 0.0412 | 1.1056 |
| 0.2 | -0.0037 | 0.0419 | 0.0421 | 1.1177 |
| 0.2 | -0.0029 | 0.041 | 0.0411 | 1.0477 |
| 0.4 | -0.0035 | 0.0418 | 0.0419 | 1.056 |
| 0.2 | -0.0054 | 0.0409 | 0.0412 | 0.9472 |
| 0.8 | -0.0078 | 0.0423 | 0.0430 | 0.9049 |

Table 5: Results for the estimation of long-memory parameters $\mathbf{d}$ with CFW-PR(4,4) filters on ARFIMA processes. For real filter, $j_0 = 1$ for $\mathbf{d} \in \{(0.2, 0.2), (0.2, 0.4)\}$ and $j_0 = 2$ for $\mathbf{d} \in \{(0.2, 0.8)\}$. For CFW-PR(4,4) filter, $j_0 = 4$. PR/Real denotes the ratio between the RMSE given by CFW-PR(4,4) filter and the RMSE given by Daubechies' real filter.

| **d** | | bias | std | RMSE |
|---|---|---|---|---|
| ( 0.2 , 0.2 ) | correlation | 8e-04 | 0.0166 | 0.0166 |
| | phase | 0 | 0.0340 | 0.0340 |
| ( 0.2 , 0.4 ) | correlation | 2e-04 | 0.0164 | 0.0164 |
| | phase | 0.0241 | 0.0345 | 0.0421 |
| ( 0.2 , 0.8 ) | correlation | -0.0033 | 0.0169 | 0.0172 |
| | phase | 0.0654 | 0.0341 | 0.0738 |

Table 6: Results for the estimation of matrices $\mathbf{\Theta}$ with CFW-PR(4,4) filters on ARFIMA processes with $j_0 = 4$.

### 5.2 Multivariate fractional Brownian motions

We now consider a multivariate fractional Brownian motion (mFBM). Since mFBM are not stationary, Fourier-based estimation is not available (without a differentiation). A specificity of mFBM is that it does not have a linear representation, even if it can be seen as the limit process of a linear representation, see Amblard et al. [2013].

The $p$-multivariate fractional Brownian motion $(\mathbf{X}(t))_{t\in\mathbb{R}}$ of long-memory parameter $\mathbf{d}$, for any $\mathbf{d} \in (0.5, 1.5)^p$ is a process satisfying the three following properties:

- $\mathbf{X}(t)$ is Gaussian for any $t \in \mathbb{R}$;
- $\mathbf{X}$ is self-similar with parameter $\mathbf{d} - 1/2$, *i.e.* for every $t \in \mathbb{R}$ and $a > 0$, $(X_1(at), \dots, X_p(at))$ has the same distribution as $(a^{d_1 - 1/2} X_1(t), \dots, a^{d_p - 1/2} X_p(t))$;
- the increments are stationary.

Another usual parametrization is the one with Hurst parameters, equal to $\mathbf{d} - 1/2$.

We introduce the following quantities, for $1 \leqslant \ell, m \leqslant p$:

$$\begin{aligned} \sigma_\ell &= \mathbb{E}[X_\ell(1)^2]^{1/2} \\ r_{\ell,m} &= r_{m,\ell} = \mathrm{Cor}(X_\ell(1), X_m(-1)) \\ \eta_{\ell,m} &= -\eta_{m,\ell} = (\mathrm{Cor}(X_\ell(1), X_m(-1)) - \mathrm{Cor}(X_\ell(-1), X_m(1)))/c_{\ell,m} \\ &\text{with } c_{\ell,m} = \begin{cases} 2(1 - 2^{d_\ell + d_m - 1}) & \text{if } d_\ell + d_m \neq 1, \\ 2\log(2) & \text{if } d_\ell + d_m = 1, \end{cases} \end{aligned}$$

where $\mathrm{Cor}(X_1, X_2)$ denotes the Pearson correlation between variables $X_1$ and $X_2$. The quantities $(\eta_{\ell,m})_{\ell,m=1,\dots,p}$ measure the dissymmetry of the process. A mFBM is time reversible if the distribution of $\mathbf{X}(-t)$ is equal to the distribution of $\mathbf{X}(t)$ for every $t$. Didier and Pipiras [2011] established that zero-mean multivariate Gaussian stationary processes $\mathbf{X}$ is equivalent to $\mathbb{E}[\mathbf{X}_\ell(t)\mathbf{X}_m(s)] = \mathbb{E}[\mathbf{X}_\ell(s)\mathbf{X}_m(t)]$ for all $(s, t)$, which corresponds to the definition of time reversibility used in Kechagias and Pipiras [2014]. A mFBM is time-reversible if and only if $\eta_{\ell,m} = 0$ for all $(\ell, m)$.

Coeurjolly et al. [2013] characterize the spectral behaviour of the increments of a mFBM. If $f^{(1,1)}_{\ell,m}$ denotes the cross-spectral density of $(\mathbb{L}X_\ell, \mathbb{L}X_m)$, then

$$f^{(1,1)}_{\ell,m}(\lambda) = 2\,\Omega_{\ell,m} \frac{1 - \cos(\lambda)}{|\lambda|^{d_\ell + d_m}} e^{i\phi_{\ell,m}},$$

with

$$\Omega_{\ell,m} = \begin{cases} \sigma_\ell \sigma_m \Gamma(d_\ell + d_m) \left( r_{\ell,m}^2 \cos^2(\frac{\pi}{2}(d_\ell + d_m)) + \eta_{\ell,m}^2 \sin^2(\frac{\pi}{2}(d_\ell + d_m)) \right)^{1/2} & \text{if } d_\ell + d_m \neq 2 \\ \sigma_\ell \sigma_m \Gamma(d_\ell + d_m) \left( r_{\ell,m}^2 + \eta_{\ell,m}^2 \frac{\pi^2}{4} \right)^{1/2} & \text{if } d_\ell + d_m = 2 \end{cases}$$

$$\phi_{\ell,m} = \begin{cases} \text{atan}\left( \frac{\eta_{\ell,m}}{r_{\ell,m}} \tan(\frac{\pi}{2}(d_\ell + d_m)) \right) & \text{if } d_\ell + d_m \neq 2 \\ \text{atan}\left( \frac{\eta_{\ell,m}}{r_{\ell,m}} \frac{\pi}{2} \right) & \text{if } d_\ell + d_m = 2. \end{cases}$$

Let $\mathbf{\Theta}$ be given by $\mathbf{\Theta} = (\Omega_{\ell,m} e^{i\phi_{\ell,m}})_{\ell,m=1,\ldots,p}$. When $\lambda$ tends to 0, the spectral density $f_{\ell,m}^{(1,1)}(\lambda)$ is equivalent to $\Theta_{\ell,m}|\lambda|^{-(d_\ell+d_m-2)}$. Thus, assumption (M-1) holds. Assumption (M-2) is satisfied for any $0 < \beta < 2$. We can verify easily that time-reversibility is still equivalent to $\phi_{\ell,m} = 0$ in this setting.

Note that the set of parameters $\{d_\ell, \sigma_\ell, r_{\ell,m}, \eta_{\ell,m}, \ell, m = 1, \ldots, p\}$ is not identifiable. Indeed, for $0 < a < 1$, $\{d_\ell, \sigma_\ell, r_{\ell,m}, \eta_{\ell,m}, \ell, m = 1, \ldots, p\}$ and $\{d_\ell, \sqrt{a}\,\sigma_\ell, r_{\ell,m}/a, \eta_{\ell,m}/a, \ell, m = 1, \ldots, p\}$ lead to the same expressions of $f_{\ell,m}^{(1,1)}(\cdot)$. It thus seems reasonable to parameterize the fractional Brownian motion by $\{d_\ell, \Theta_{\ell,m},\ \ell, m = 1, \ldots, p\}$.

We consider two mFBM, both with parameters $\sigma_1 = \sigma_2 = 1$ and $\mathbf{d} = (1, 1.2)$.

**Case 1.** $\eta_{1,2} = 0.9$, $r_{1,2} = 0.6$.

The phase $\phi_{1,2}$ is approximately equal to $\pi/7$ and $\mathbf{\Omega} \simeq \begin{pmatrix} 1.000 & 0.699 \\ 0.699. & 1.005 \end{pmatrix}$, giving a long-run correlation $\rho \simeq 0.70$.

**Case 2.** $\eta_{1,2} = -0.6$, $r_{1,2} = 0.2$.

The phase $\phi_{1,2}$ is approximately equal to $-\pi/4$ and $\mathbf{\Omega} \simeq \begin{pmatrix} 1.000 & 0.293 \\ 0.293 & 1.005 \end{pmatrix}$, giving a long-run correlation $\rho \simeq 0.29$.

Simulations were done using R functions provided by J-F Coeurjolly at https://sites.google.com/site/homepagejfc/software.

Figure 5 represents the boxplots of CFW-PR(4,4) wavelet correlations at different scales in Case 1 and in Case 2. The good behavior of approximation is observed except for highest frequencies.

Identical observations are obtained for CFW-C(4,4) filters, hence the figure is not displayed here.

We now consider the local Whittle estimation of the parameters. Based on the discussion of Section 5.1, and on Figure 5, we fix $j_0 = 4$. Table 7 and Table 8 highlight the good behavior of the estimation of long-memory parameters $\mathbf{d}$ respectively for CFW-PR(4,4) and CFW-C(4,4) filters. Again, when considering $j_0 = 4$ for both of them, estimation procedures are equivalent for both common-factor wavelets. By comparison to the real wavelet-based estimation (with $j_0 = 2$ as suggested by Achard and Gannaz [2019]), the RMSE is increased. This is mainly due to the choice of the hyperparameter $j_0$.

| | $\mathbf{d}$ | bias | std | RMSE | ratio PR/Real |
|---|---|---|---|---|---|
| Case 1 | 1 | -0.0065 | 0.0464 | 0.0469 | 1.5577 |
| | 1.2 | -0.0059 | 0.0475 | 0.0478 | 2.0886 |
| Case 2 | 1 | -0.0051 | 0.0510 | 0.0513 | 1.6812 |
| | 1.2 | -0.0035 | 0.0515 | 0.0516 | 1.9884 |

Table 7: Results for the estimation of long-memory parameters $\mathbf{d}$ with CFW-PR(4,4) filter on mFBMs. Hyperparameter $j_0$ satisfies $j_0 = 4$ for CFW-PR(4,4) and $j_0 = 2$ for real filters. PR/Real denotes the ratio between the RMSE given by CFW-PR(4,4) filter and the RMSE given by Daubechies' real filter.

| | $\mathbf{d}$ | bias | std | RMSE | ratio C/PR |
|---|---|---|---|---|---|
| Case 1 | 1 | -0.0155 | 0.0409 | 0.0437 | 0.9323 |
| | 1.2 | -0.0133 | 0.0402 | 0.0423 | 0.8849 |
| Case 2 | 1 | -0.0177 | 0.0448 | 0.0482 | 0.9397 |
| | 1.2 | -0.0116 | 0.0473 | 0.0487 | 0.9441 |

Table 8: Results for the estimation of long-memory parameters $\mathbf{d}$ with CFW-C(4,4) filter on mFBMs. Hyperparameter $j_0$ satisfies $j_0 = 4$. C/PR denotes the ratio between the RMSE given by CFW-C(4,4) filter and the RMSE given by CFW-PR(4,4) filter.

Table 9 and Table 10 give the results obtained for the estimation of the covariance structure, that is, $\mathbf{\Omega}$, $\rho$ and $\phi$. It is not possible to compare our results with alternative non parametric procedures because real wavelet-based estimation cannot estimate the long-run covariance or correlation but only its real part, and Fourier-based estimations are not valid for non-stationary time series.

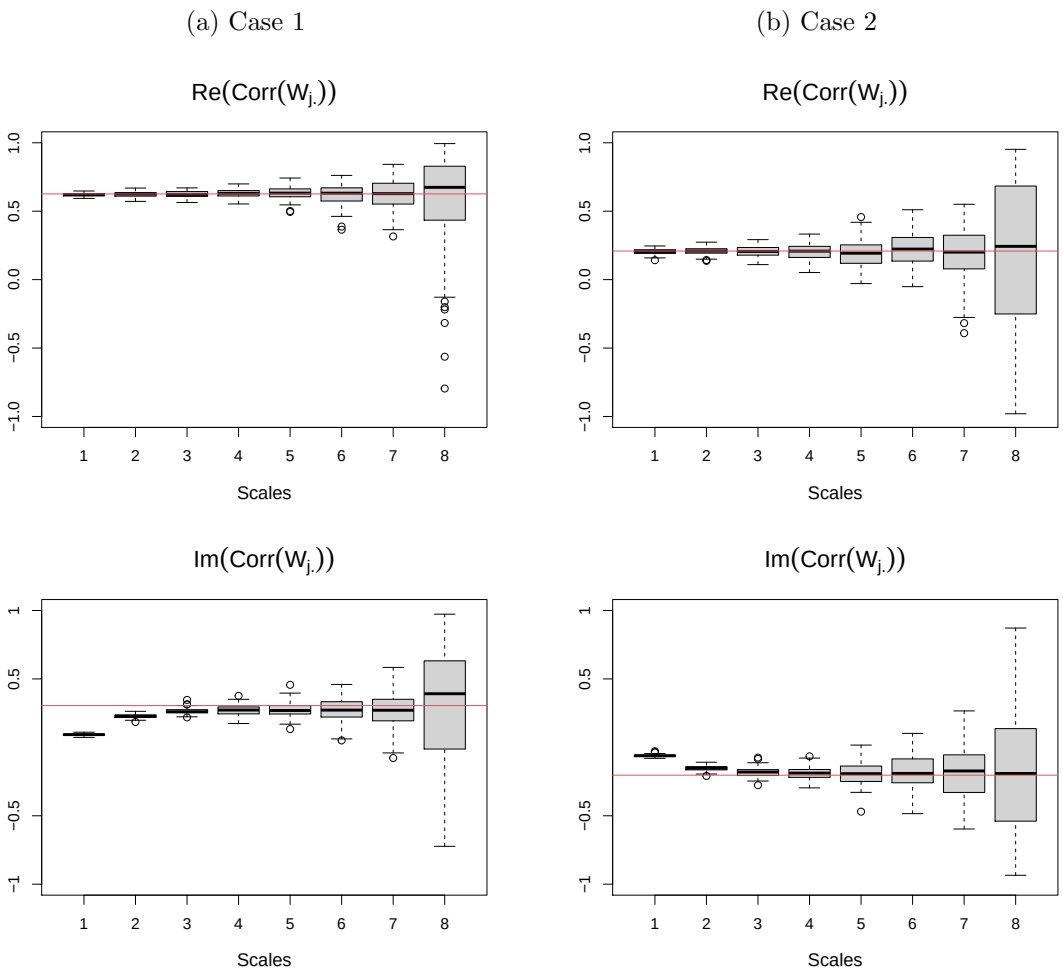


Figure 5: Boxplots of correlation between CFW-PR(4,4) coefficients at different scales for the simulated mFBM in Case 1 (left column–(a)) and in Case 2 (right column–(b)). First row gives the real part of the correlations and second row gives the imaginary part. Horizontal red lines correspond to the approximation given by Proposition 4, that is, $\rho\cos(\phi)r_K$ for the real part and $\rho\sin(\phi)r_K$ for the imaginary part, with $r_K = K(d_1+d_2)/\sqrt{K(2\,d_1)K(2\,d_2)}$.

The results of CFW-PR(4,4) and CFW-C(4,4) are similar. We observe a high bias and high standard deviation for the estimation of $\mathbf{\Omega}$. On the other hand, we observe good quality for the estimation of $\rho$ and of $\phi$ are .

| | | bias | std | RMSE |
|---|---|---|---|---|
| Case 1 | $\Omega_{1,1}$ | -0.1999 | 0.1544 | 0.2526 |
| | $\Omega_{1,2}$ | -0.1592 | 0.0925 | 0.1841 |
| | $\Omega_{2,2}$ | -0.2471 | 0.1504 | 0.2892 |
| | correlation | 0.0971 | 0.0243 | 0.1001 |
| | phase | 0.0039 | 0.0526 | 0.0528 |
| Case 2 | $\Omega_{1,1}$ | -0.2048 | 0.1654 | 0.2633 |
| | $\Omega_{1,2}$ | -0.0647 | 0.0501 | 0.0818 |
| | $\Omega_{2,2}$ | -0.2508 | 0.1588 | 0.2969 |
| | correlation | 0.0967 | 0.0434 | 0.1059 |
| | phase | -0.0087 | 0.1549 | 0.1551 |

Table 9: Results for the estimation of matrices $\mathbf{\Theta}$ with CFW-PR(4,4) filter on mFBMs. Hyperparameter $j_0$ satisfies $j_0 = 4$.

| | | bias | std | RMSE | ratio C/PR |
|---|---|---|---|---|---|
| Case 1 | $\Omega_{1,1}$ | -0.1548 | 0.1501 | 0.2156 | 0.8537 |
| | $\Omega_{1,2}$ | -0.1305 | 0.0865 | 0.1565 | 0.8501 |
| | $\Omega_{2,2}$ | -0.2088 | 0.1365 | 0.2495 | 0.8625 |
| | correlation | 0.0965 | 0.0245 | 0.0995 | 0.9940 |
| | phase | 0.0043 | 0.0493 | 0.0495 | 0.9376 |
| Case 2 | $\Omega_{1,1}$ | -0.1427 | 0.1632 | 0.2168 | 0.8235 |
| | $\Omega_{1,2}$ | -0.052 | 0.0527 | 0.074 | 0.9043 |
| | $\Omega_{2,2}$ | -0.2107 | 0.1596 | 0.2643 | 0.8905 |
| | correlation | 0.0936 | 0.045 | 0.1039 | 0.9805 |
| | phase | -0.0061 | 0.156 | 0.1561 | 1.0064 |

Table 10: Results for the estimation of matrices $\mathbf{\Theta}$ with CFW-C(4,4) filter on mFBMs. Hyperparameter $j_0$ satisfies $j_0 = 4$. C/PR denotes the ratio between the RMSE given by CFW-C(4,4) filter and the RMSE given by CFW-PR(4,4) filter.

# 6 Application on a neuroscience dataset

We applied our framework on fMRI data acquired on rats. We use functional Magnetic Resonance images (fMRI) of both dead and live rats. Our aim is to estimate the brain connectivity, that is, the

significant correlations between brain regions where fMRI signals are recorded. For this data set, we know that for dead rats the recordings are only noise, as no legitimate functional activity should be detected. Thus, the estimated graphs should be empty. We also expect non-empty graphs for live rats under anesthetic, as brain activity keeps on during anesthetic. The dataset are freely available at https://10.5281/zenodo.2452871 [Becq et al., 2020a,b].

### 6.1 Description of the dataset

Functional Magnetic Resonance Images (fMRI) were acquired for both dead and live rats (complete description is available in Becq et al. [2020b]). 25 rats were scanned and identified in 4 different groups: DEAD, ETO_L, ISO_W and MED_L. The first group contain dead rats and the three last groups correspond to different anesthetics. The duration of the scanning was 30 minutes with a time repetition of 0.5 second so that $N = 3600$ time points were available at the end of experience. After preprocessing as explained in Becq et al. [2020b], $p = 51$ time series for each rat were extracted. Each time series captures the functioning of a given region of the rat brain based on an anatomical atlas.

For each rat, we compute the estimators of

- the vector of long-memory parameters, $\widehat{\mathbf{d}}$,
- the magnitude of the correlations, $\widehat{\rho} = \{\widehat{\rho}_{\ell,m},\ 1 \leqslant \ell < m \leqslant p\}$ with $\widehat{\rho}_{\ell,m} = \dfrac{\widehat{\Omega}_{\ell,m}}{\sqrt{\widehat{\Omega}_{\ell,\ell}\widehat{\Omega}_{m,m}}}$,
- the phases, $\widehat{\phi} = \{\widehat{\phi}_{\ell,m},\ 1 \leqslant \ell < m \leqslant p\}$.

Estimation was done with CFW-PR(4,4) filters. Densities of the estimators are represented on the figures using `R` default kernel-based estimation.

### 6.2 Results and group comparisons

Figure 6 shows the empirical distribution of the estimated empirical estimators $\widehat{d}$. As expected, the long-memory parameters for dead rats are close to zero. The distributions are centered around zero, with a Gaussian-like shape. For rats under anesthetics, densities are not centered around zero and the variance between brain regions are higher than what is observed for dead rats. Long-memories for rats under anesthetic ISO_W are higher than under other anesthetics.

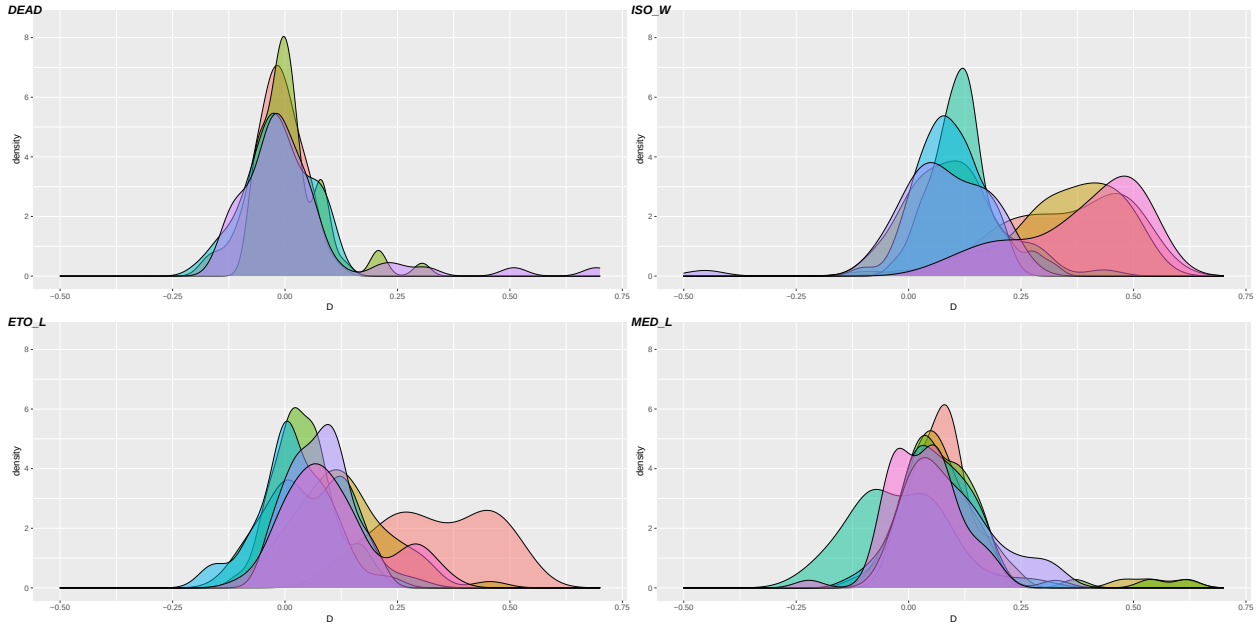


Figure 6: Plot of the empirical distribution of the long memory parameters $\widehat{\mathbf{d}}$ obtained for the 4 groups of rats. Each color corresponds to a rat.

Distributions of magnitude and phases of the estimated correlations, $\rho$ and $\phi$, for each rats are displayed respectively in Figure 7 and Figure 8. First, as expected, the magnitudes obtained for the dead rats seem significantly different from live rats. For dead rats, distributions have a small support, that is, only 9 on the 5100 values (0.18%) satisfy $\widehat{\rho} > 0.3$. Note also that no major differences are observed between rats. Next, ISO_W and ETO_L present quite similar distributions, with possibly high magnitudes. By contrast, correlations for MED_L anesthetic are lower. These results tend to show that MED_L anesthetic is stronger than the other anesthetics, leading to less connections between brain regions.

The phase parameter can be interpreted as an asymmetry in the properties at large lags among the components of the signals for each brain regions (a null phase is equivalent to time-reversibility). Distributions displayed in Figure 8 correspond to the empirical densities of the upper triangular matrices of phases, $\{\phi_{\ell,m},\, 1 \leqslant \ell < m \leqslant p\}$. This explains why the distributions are not symmetric.

For the dead rats, we observe mainly uniform distributions. For the live rats, Figure 8 shows that the distributions have heavy tails. The tails are heavier for MED_L anesthetic than for other

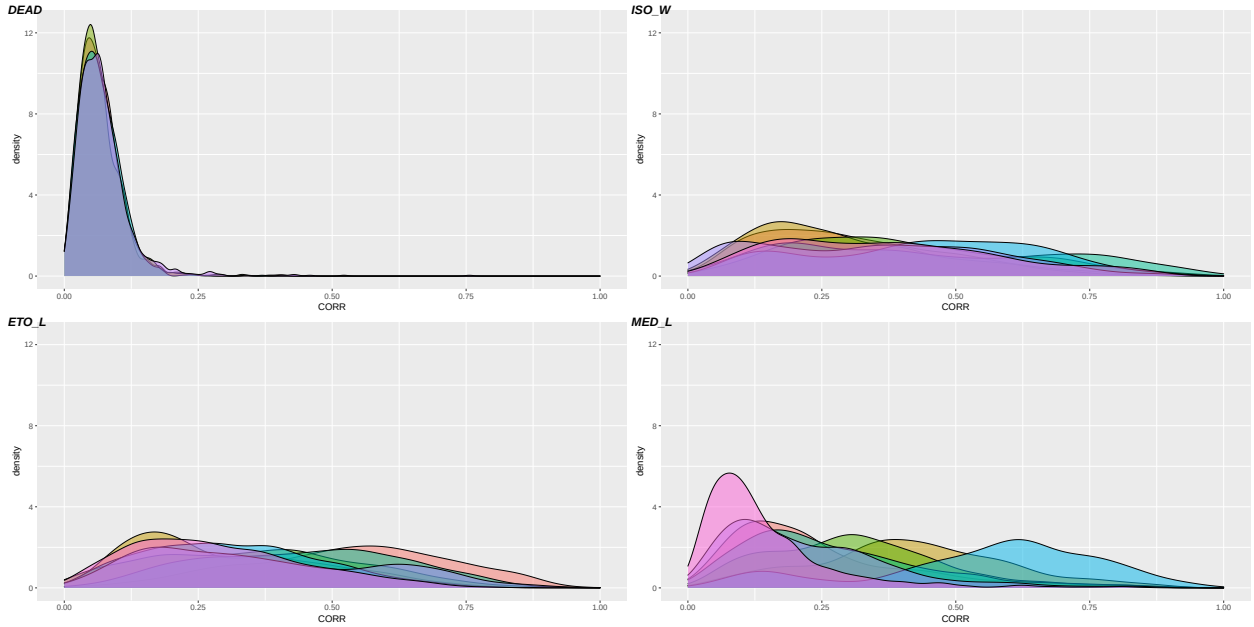


Figure 7: Plot of the empirical distribution of the correlation magnitudes $\widehat{\rho}$ obtained for the 4 groups of rats. Each color corresponds to a rat.

anesthetics. This can be explained by the fact that the phase is non-informative when the magnitude is close to zero. To illustrate this fact, Figure 9 shows the distributions of the estimated phases $\phi$ corresponding to magnitudes satisfying $\rho > 0.3$ (this choice is motivated by the observation on the support of dead rats' correlations above). Distributions then have smallest tails. It can be observed that the support of the phases are larger for live rats than for dead rats. Next, the 95%-quantiles of absolute values (*i.e.* $q$ such that 95% of absolute values of phases are lower than $q$) are respectively 2.95, 1.90, 1.89, 1.61 for Dead rats, ISO_W, ETO_L and MED_L. It seems that ISO_W has a higher support, meaning that shifts appear in the connections between brain region, with respect to other anesthetics. Yet, we have not tested whether the difference is significant.

### 6.3 Graphs with correlations and phases

We first compute the adjacency matrix obtained for each rat within each group. Edges correspond to a magnitude higher than 0.3. The value of the threshold is motivated by the observation of the supports obtained for dead rats. We then select the edges which are present in all the graphs of

the rats of the group. One graph is then obtained per group. For each group, we then compute the average of the estimated phase for each detected edge. Figure 10 illustrates the graphs obtained for the 4 different groups.

We colored each edge based on the average phase when it satisfies $|\phi_{\ell,m}| > 1.1|\phi^*_{\ell,m}|$ where $\phi^*_{\ell,m} = -\frac{\pi}{2}(d_\ell - d_m)$, $(\ell, m) \in \{1, \dots, p\}^2$. The value $\phi^*_{\ell,m}$ corresponds to the phase of causal linear representations with power-law coefficients [Kechagias and Pipiras, 2014] and to the ARFIMA modelling used in Achard and Gannaz [2016] with similar data. The more colored edges, the more the phase behavior differs from the previous modelling.

The DEAD group has indeed no edges. The MED_L group has less edges than the two other groups of anesthetic. It hence seems that MED_L anesthetic inhibits more the activity. Next ETO_L group and ISO_W group have a similar number of edges (respectively 133 and 145), but the phases differ. More than half of mean phases are outside the interval $[-1.1|\phi^*|, 1.1|\phi^*|]$ for ETO_L and ISO_W group, with similar proportions. This observation is interesting since it illustrates that the modelling of these data is complex. The introduction of a general phase enables to take into account this complexity. Concerning the physical interpretation, no easy conclusion can be given. As it was mentioned in Buxton [2013], the time scale of BOLD (Blood oxygenation level dependent) response is very small in comparison to the neuronal activity, the delay is observed to be equal to a few seconds. Considering the different time scales involved in the production of the BOLD response, we may hypothesize that lags are not the underlying phenomenon that produces phase differences in the fMRI signals. However, as stated in Buxton [2013], the time scale can vary in the same subject depending on the physiological baseline state, which is known to be modified under anesthetization.

# 7 Conclusion

This work was motivated by a neuroscience application, namely the inference of fractal connectivity from fMRI recordings. We studied the local Whittle estimators for multivariate time series presenting long-memory. Our modelling allows a complex covariance structure with phase components which can be interpreted as shifts in the coupling between time series. We introduced quasi-analytic wavelet filters to handle the possible non-stationarity of the real data application. The resulting procedures offer a consistent estimation of the main parameters of the model. Indeed, we estab-

lished that so called *Common-Factor wavelets* are an efficient tool for recovering the long-memory structure as well as the covariance structure, including magnitude and phase. A simulation study on linear processes and on multivariate Brownian motions illustrate the good performance of the proposed procedure. The real data application highlights the ability of the procedure to distinguish dead rats from live rats. We also show the differences between three anesthetics and the fact that one of them slows down more intensively the brain activity.

**Ackowledgments**

The authors are grateful to Marianne Clausel and Franǫis Roueff who contributed significantly to this work. We also thank Jean-François Coeurjolly for providing the code for the simulations of multivariate Brownian motions. We also would like to thank Emmanuel Barbier and Guillaume Becq for providing us the data of the resting state fMRI on the rats.

# A Expression of the CFW-C(M,L) filter

This section aims at giving the expression of the CFW-C(M,L) filters which are used to compute the wavelet coefficients. Let first recalls the expression of common-factor wavelets.

## A.1 Expression of the CFW-C(M,L) pair

Let us recall the expression of the low-pass filter $\widehat{h}^{(L)}$ and the high-pass filter $\widehat{h}^{(H)}(\lambda)$:

$$\widehat{h}^{(L)}(\lambda) = \sqrt{2}\left(\frac{1+e^{-i\lambda}}{2}\right)^{M} \widehat{q}_{L,M}(\lambda)\,\widehat{d}_{L}(\lambda) \quad \text{and} \quad \widehat{h}^{(H)}(\lambda) = \overline{\widehat{h}^{(L)}(\lambda+\pi)}e^{-i\lambda}\,,$$

All the same,

$$\widehat{g}^{(L)}(\lambda) = \sqrt{2}\left(\frac{1+e^{-i\lambda}}{2}\right)^{M} \widehat{q}_{L,M}(\lambda)\,\overline{\widehat{d}_{L}(\lambda)}e^{-i\lambda L} \quad \text{and} \quad \widehat{g}^{(H)}(\lambda) = \overline{\widehat{g}^{(L)}(\lambda+\pi)}e^{-i\lambda}.$$

Let us now explicit $\widehat{\varphi}(\cdot)$. We have $\widehat{\varphi}(\cdot) = \widehat{\varphi}_h(\cdot) + \mathrm{i}\widehat{\varphi}_g(\cdot)$ with

$$\widehat{\varphi}_h(\lambda) = 2^{-1/2}\prod_{j=1}^{\infty} 2^{-1/2}\widehat{h}^{(L)}(2^{-j}\lambda) = 2^{-1/2}\prod_{j=1}^{\infty}\big(\frac{1+\mathrm{e}^{-\mathrm{i}2^{-j}\lambda}}{2}\big)^M \widehat{q}_{L,M}(2^{-j}\lambda)\widehat{d}_L(2^{-j}\lambda),$$

$$\text{and}\quad \widehat{\varphi}_g(\lambda) = 2^{-1/2}\prod_{j=1}^{\infty}\big(\frac{1+\mathrm{e}^{-\mathrm{i}2^{-j}\lambda}}{2}\big)^M \widehat{q}_{L,M}(2^{-j}\lambda)\overline{\widehat{d}_L(2^{-j}\lambda)}\mathrm{e}^{-\mathrm{i}2^{-j}\lambda L}.$$

When $\widehat{q} = 1$, the expressions above become

$$\widehat{\varphi}_h(\lambda) = \prod_{j=1}^{\infty}\big(\frac{1+\mathrm{e}^{-\mathrm{i}2^{-j}\lambda}}{2}\big)^M \overline{\widehat{d}_L(2^{-j}\lambda)}, \tag{15}$$

$$\widehat{\varphi}_g(t) = \prod_{j=1}^{\infty}\big(\frac{1+\mathrm{e}^{-\mathrm{i}2^{-j}\lambda}}{2}\big)^M \overline{\widehat{d}_L(2^{-j}\lambda)}\mathrm{e}^{-\mathrm{i}2^{-j}\lambda L}. \tag{16}$$

Next, we can explicit $\widehat{\psi}(\cdot)$. We have $\widehat{\psi}(\cdot) = \widehat{\psi}_h(\cdot) + \mathrm{i}\widehat{\psi}_g(\cdot)$ with

$$\begin{aligned}
&\widehat{\psi}_h(\lambda)\\
&= 2^{-3/2}\widehat{h}^{(H)}(\lambda/2)\prod_{j=2}^{\infty} 2^{-1/2}\widehat{h}^{(L)}(2^{-j}\lambda)\\
&= 2^{-1}\Big(\frac{1-\mathrm{e}^{\mathrm{i}\lambda/2}}{2}\Big)^M \overline{\widehat{q}_{L,M}(\lambda/2+\pi)}\overline{\widehat{d}_L(\lambda/2+\pi)}\mathrm{e}^{-\mathrm{i}\lambda/2}\prod_{j=2}^{\infty}\big(\frac{1+\mathrm{e}^{-\mathrm{i}2^{-j}\lambda}}{2}\big)^M \widehat{q}_{L,M}(2^{-j}\lambda)\widehat{d}_L(2^{-j}\lambda),
\end{aligned}$$

and

$$\widehat{\psi}_g(\lambda) = -2^{-1/2}\Big(\frac{1-\mathrm{e}^{-\mathrm{i}\lambda/2}}{2}\Big)^M \overline{\widehat{q}_{L,M}(\lambda/2+\pi)}\,\widehat{d}_L(\lambda/2+\pi)\mathrm{e}^{+\mathrm{i}(\lambda/2+\pi)L}\mathrm{e}^{-\mathrm{i}\lambda/2}$$
$$\prod_{j=2}^{\infty}\big(\frac{1+\mathrm{e}^{-\mathrm{i}2^{-j}\lambda}}{2}\big)^M \widehat{q}_{L,M}(2^{-j}\lambda)\overline{\widehat{d}_L(2^{-j}\lambda)}\mathrm{e}^{-\mathrm{i}2^{-j}\lambda L}.$$

When $\widehat{q}=1$, the expressions above become

$$\widehat{\psi}_h(\lambda) = 2^{-1/2}\Big(\frac{1-\mathrm{e}^{-\mathrm{i}\lambda/2}}{2}\Big)^M \overline{\widehat{d}_L(\lambda/2+\pi)}\mathrm{e}^{-\mathrm{i}\lambda/2}\prod_{j=2}^{\infty}\Big(\frac{1+\mathrm{e}^{-\mathrm{i}2^{-j}\lambda}}{2}\Big)^M \widehat{d}_L(2^{-j}\lambda)$$

$$\widehat{\psi}_g(\lambda) = -2^{-1/2}\Big(\frac{1-\mathrm{e}^{-\mathrm{i}\lambda/2}}{2}\Big)^M \widehat{d}_L(\lambda/2+\pi)\mathrm{e}^{+\mathrm{i}(\lambda/2+\pi)L}\mathrm{e}^{-\mathrm{i}\lambda/2}\prod_{j=2}^{\infty}\Big(\frac{1+\mathrm{e}^{-\mathrm{i}2^{-j}\lambda}}{2}\Big)^M \overline{\widehat{d}_L(2^{-j}\lambda)}\mathrm{e}^{-\mathrm{i}2^{-j}\lambda L}.$$

We shall use the following equality

$$\prod_{\ell=1}^{\infty}\left(\frac{1+\mathrm{e}^{-\mathrm{i}2^{-\ell}\lambda}}{2}\right) = \frac{1-e^{-\mathrm{i}\lambda}}{-\mathrm{i}\lambda} = \mathrm{e}^{-\mathrm{i}\lambda/2}\frac{\sin(\lambda/2)}{\lambda/2}.$$

See *e.g.* [Mallat, 1999, page 245]. It yields

$$\widehat{\varphi}_h(\lambda) = 2^{-1/2}\Big(\frac{\sin(\lambda/2)}{\lambda/2}\Big)^M \prod_{j=1}^{\infty}\overline{\widehat{d}_L(2^{-j}\lambda)}, \tag{17}$$

$$\widehat{\varphi}_g(t) = 2^{-1/2}\Big(\frac{\sin(\lambda/2)}{\lambda/2}\Big)^M \prod_{j=1}^{\infty}\overline{\widehat{d}_L(2^{-j}\lambda)}\mathrm{e}^{-\mathrm{i}2^{-j}\lambda L}, \tag{18}$$

and

$$\widehat{\psi}_h(\lambda) = 2^{-1/2}\sin(\lambda/4)^M\Big(\frac{\sin(\lambda/4)}{\lambda/4}\Big)^M \overline{\widehat{d}_L(\lambda/2+\pi)}\mathrm{e}^{-\mathrm{i}\lambda}\prod_{j=2}^{\infty}\widehat{d}_L(2^{-j}\lambda), \tag{19}$$

$$\widehat{\psi}_g(\lambda) = 2^{-1/2}\sin(\lambda/4)^M\Big(\frac{\sin(\lambda/4)}{\lambda/4}\Big)^M \widehat{d}_L(\lambda/2+\pi)\mathrm{e}^{+\mathrm{i}(\lambda+\pi)L}\mathrm{e}^{-\mathrm{i}\lambda}\prod_{j=2}^{\infty}\overline{\widehat{d}_L(2^{-j}\lambda)}\mathrm{e}^{-\mathrm{i}2^{-j}\lambda L}. \tag{20}$$

## A.2 A first property

The function $\widehat{d}_L(\lambda)$ satisfies

$$\widehat{d}_L(\lambda) = \mathrm{e}^{\mathrm{i}\,\lambda(-L/2+1/4)}\left[\cos(\lambda/4)^{2L+1} + \mathrm{i}\,(-1)^{L+1}\sin(\lambda/4)^{2L+1}\right].$$

We deduce the following lemma.

**Lemma 7.**

$$\sup_{\lambda\in\mathbb{R}}\left|\widehat{d}_L(\lambda)\right| = 1.$$

The proof is straightforward and it is thus omitted.

Using the fact that, for all $\lambda \in \mathbb{R}$, $\left|\frac{\sin(\lambda)}{\lambda}\right| \leqslant 1$, a direct consequence is the following result.

**Lemma 8.**

$$\sup_{\lambda\in\mathbb{R}}|\widehat{\varphi}(\lambda)| \leqslant C_\infty, \tag{21}$$

*with* $C_\infty = 1$.

### A.3 Expression of the wavelet filters

Recall that

$$W_{j,k}(\ell) = \int_{-\infty}^{\infty}\sum_{s\in\mathbb{Z}} X_\ell(s)\varphi(t+s)\psi_{j,k}(t)\mathrm{d}t.$$

Hence,

$$W_{j,k}(\ell) = \sum_{s\in\mathbb{Z}}\tau_j(2^j k - s)X_\ell(s), \quad j \geqslant 0, k \in \mathbb{Z}\,,$$

$$\text{with } \widehat{\tau}_j(\lambda) = \sum_{s\in\mathbb{Z}}\tau_j(2^j k - s)\mathrm{e}^{\mathrm{i}\lambda s} = \int_{-\infty}^{\infty}\sum_{\ell\in\mathbb{Z}}\varphi(t+\ell)\,\mathrm{e}^{-\mathrm{i}\lambda\ell}2^{-j/2}\psi(2^{-j}t)\mathrm{d}t.$$

## B Properties of CFW-C(M,L) filters

Let us introduce the following properties.

(W1) *Finite support.* $\varphi$ and $\psi$ have finite support.

(W2) *Vanishing moments.* There exist $M \geqslant 0$ and $C_m > 0$ such that for all $j \geqslant 0$ and $\lambda \in \mathbb{R}$,

$$\left|\widehat{\psi}(\lambda)\right| \leqslant C_m\ |\lambda|^M\ ,$$

with $C_m$ positive constant possibly depending on $M$.

(W3) *Uniform smoothness.* There exist $\alpha > 1$ and $C_s > 0$ such that for all $\lambda \in \mathbb{R}$,

$$\left|\widehat{\psi}(\lambda)\right| \leqslant \frac{C_s}{(1+|\lambda|)^{\alpha}},$$

with $\alpha$ and $C_s$ depending on $L$ and $M$.

(W4) *Scaling function.* There exist $C_\varphi$ depending on $M$ such that, for all $\lambda \in (-\pi, \pi)$, for all $k \in \mathbb{Z} \setminus \{0\}$,

$$|\widehat{\varphi}(\lambda + 2k\,\pi)| \leqslant C_\varphi \ |\lambda|^M \ .$$

Properties (W1), (W2), (W3) and (W4) correspond respectively to (W1), (W2), (W3) and (W4) in the context of real-wavelets. We can establish that they are satisfied by CFW-C(L,M) wavelets.

**Proposition 9.** *When $M \geqslant 2$, and $L \geqslant 1$, CFW-C(M,L) wavelets satisfy (W2), (W3) and (W4), with $\alpha = M$ and constants $C_m = 1$, $C_s = 2 \cdot 5^M$ and $C_\varphi = 2$.*

The proof is given in Section B.1.

A remarkable property of $CFW - C(M, L)$ filters is that the regularity of the wavelets is only determined by the parameter $M$, since all parameters and constants in the proposition above only depend on $M$. All the same, the quasi-analyticity only depends on the parameter $L$, through Theorem 2.

With these assumptions, we can establish some properties about wavelet filters. At a given scale $j \geqslant 0$, for any $k \in \mathbb{Z}$, wavelet coefficients $W_{j,k}(\ell)$ of a process $X_\ell(\cdot)$ can be decomposed as

$$W_{j,k}(\ell) = \sum_{s \in \mathbb{Z}} \tau_j(2^j k - s) X_\ell(s), \quad j \geqslant 0, k \in \mathbb{Z}\,.$$

See Section A.3. We recover [Moulines et al., 2007, Proposition 3]. More precisely, we can establish the following results.

**Proposition 10.** *Suppose (W1), (W2), (W3), and (W4) and inequality (21) hold. Then, for all*

$j \geqslant 0$, *for all* $\lambda \in \mathbb{R}$,

$$\left|\widehat{\tau}_j(\lambda) - 2^{j/2}\widehat{\varphi}(\lambda)\overline{\widehat{\psi}(2^j\lambda)}\right| \leqslant C_\tau 2^{j(1/2-\alpha)}|\lambda|^M, \tag{22}$$

$$\left|\widehat{\varphi}(\lambda)\overline{\widehat{\psi}(2^j\lambda)}\right| \leqslant C_\infty C_m \left|2^j\lambda\right|^M, \tag{23}$$

$$\left|\widehat{\varphi}(\lambda)\overline{\widehat{\psi}(2^j\lambda)}\right| \leqslant \frac{C_\infty C_s}{(1+2^j|\lambda|)^\alpha}, \tag{24}$$

$$\left|2^{-j/2}\widehat{\tau}_j(2^{-j}\lambda)\right| \leqslant C_{m\tau}\ |\lambda|^M, \tag{25}$$

*and, for all* $j, j' \geqslant 0$, *for all* $|2^{-j}\lambda| \leqslant \pi$,

$$\left|2^{-j/2}\widehat{\tau}_j(2^{-j}\lambda)\right| \leqslant C_{s\tau}\,(1+|\lambda|)^{-\alpha}, \tag{26}$$

$$\left|2^{-j/2}\widehat{\tau}_j(2^{-j}\lambda)\right| \leqslant C_{ms\tau}\ |\lambda|^M\,(1+|\lambda|)^{-\alpha-M}, \tag{27}$$

*with* $C_\tau$, $C_{m\tau}$, $C_{s\tau}$, *and* $C_{ms\tau}$ *positive constants only depending on* $\alpha, M$, $C_m$, $C_s$ *and* $C_\varphi$.

The proof is given in Section B.2.

The following property corresponds to (79) in [Moulines et al., 2007, Proposition 3]. In the real wavelets context, it is a consequence of (W1) to (W4) but we prove it separately here to explicit the constants.

**Proposition 11.** *Consider CFW-C(M,L) wavelets, with* $M \geqslant 2$. *For all* $j \geqslant 1$,

$$\sup_{|\lambda|\leqslant\pi}\left|\left|2^{-j/2}\widehat{\tau}_j(2^{-j}\lambda)\right|^2 - \left|\widehat{\psi}(\lambda)\right|^2\right| \leqslant C_a\, 2^{-\gamma j}\,|\lambda|^{2M}\ , \tag{28}$$

*with* $\gamma = 2$ *and* $C_a = 2\,(M+L+1)$.

The proof is given in Section B.3.

## B.1 Proof of Proposition 9

We first establish that CFW-C(M,L) filters satisfy properties (W1)–(W4).

### B.1.1 Property (W1)

By definition, the filters $h^{(L)}$ and $h^{(H)}$, $g^{(L)}$ and $g^{(H)}$ have length $M+L+1$. Hence, $\varphi$ and $\psi$ have finite support.

### B.1.2 Property (W2)

Let $\lambda \in \mathbb{R}$. Recall that $\widehat{\psi}_h(\lambda)$ and $\widehat{\psi}_g(\lambda)$ are given respectively by (19) and (20). Observe that $|\sin(\lambda/4)|^M \leqslant 4^{-M} |\lambda|^M$. Since $\sup |\widehat{d}_L| = 1$, assumption (W2) follows with a constant $C_m = 1$, for all $M \geqslant 1$.

### B.1.3 Property (W3)

Let $\lambda \in \mathbb{R}$. Recall that $\widehat{\psi}_h(\lambda)$ is given by (19). Since $\sup_{\lambda\in\mathbb{R}} \left|\widehat{d}_L(\lambda)\right| \leqslant 1$, we obtain

$$\left|\widehat{\psi}_h(\lambda)\right| \leqslant |\sin(\lambda/4)|^M \left|\frac{\sin(\lambda/4)}{\lambda/4}\right|^M.$$

Since $\frac{\sin(x/4)}{|x/4|}(1+|x|) = \left|\frac{\sin(x/4)}{x/4}\right| + 4\,|\sin(x/4)| \leqslant 5$ for any $x \in \mathbb{R}\setminus\{0\}$, it follows that

$$\left|\widehat{\psi}_h(\lambda)\right| \leqslant |\sin(\lambda/4)|^M \left(\frac{5}{1+|\lambda|}\right)^M.$$

Consequently, $\left|\widehat{\psi}_h(\lambda)\right| \leqslant \left(\frac{5}{1+|\lambda|}\right)^M$. A similar result can be proved for filter $\widehat{\psi}_g$. By triangular inequality, we get

$$\left|\widehat{\psi}(\lambda)\right| \leqslant \frac{C_s}{(1+|\lambda|)^\alpha},$$

with $\alpha = M$ and a constant $C_s$ equal to $2 \cdot 5^M$.

### B.1.4 Property (W4)

Let $\lambda = \omega + 2k\pi$, $|\omega| \leqslant \pi$, $k \in \mathbb{Z}$, $k \neq 0$. We distinguish two cases:

- If $k$ is odd,
  Using the expressions of $\widehat{\varphi}_h(\lambda)$ and $\widehat{\varphi}_g(\lambda)$ given respectively by (15) and (16),

$$\widehat{\varphi}_h(\lambda) = \big(\cos(\lambda/4)\big)^M \overline{\widehat{d}_L(\lambda/2)} \prod_{j=2}^{\infty} \big(\frac{1+\mathrm{e}^{-\mathrm{i}2^{-j}\lambda}}{2}\big)^M \overline{\widehat{d}_L(2^{-j}\lambda)},$$

$$\widehat{\varphi}_g(\lambda) = \big(\cos(\lambda/4)\big)^M \overline{\widehat{d}_L(\lambda/2)} \mathrm{e}^{-\mathrm{i}\lambda L/2} \prod_{j=2}^{\infty} \big(\frac{1+\mathrm{e}^{-\mathrm{i}2^{-j}\lambda}}{2}\big)^M \overline{\widehat{d}_L(2^{-j}\lambda)} \mathrm{e}^{-\mathrm{i}2^{-j}\lambda L}.$$

  Observe that

$$\big|\cos(\lambda/4)\big|^M = |\cos(\omega/4 + k\pi/2)|^M = |\sin(\omega/4)|^M \leqslant |\omega|^M/4^M.$$

  With Lemma 7, we obtain that

$$|\widehat{\varphi}_h(\lambda)| \leqslant |\omega|^M/4^M \quad \text{and} \quad |\widehat{\varphi}_g(\lambda)| \leqslant |\omega|^M/4^M.$$

  Hence,

$$|\widehat{\varphi}(\lambda)| \leqslant |\omega|^M \cdot 2/4^M.$$

- If $k$ is even, $k \geqslant 2$,
  Let us use the expressions of $\widehat{\varphi}_h(\lambda)$ and $\widehat{\varphi}_g(\lambda)$ given respectively by (17) and (18). We have,

$$\Big|\frac{\sin(\lambda/2)}{\lambda/2}\Big|^M = \Big|\frac{\sin(\omega/2)}{\omega/2}\Big| \Big|\frac{\omega}{\lambda}\Big|^M \leqslant \Big|\frac{\omega}{\lambda}\Big|^M \leqslant \Big|\frac{\omega}{\pi}\Big|^M.$$

  Using Lemma 7,

$$|\widehat{\varphi}_h(\lambda)| \leqslant |\omega|^M/\pi^M \quad \text{and} \quad |\widehat{\varphi}_g(\lambda)| \leqslant |\omega|^M/\pi^M.$$

  Hence,

$$|\widehat{\varphi}(\lambda)| \leqslant |\omega|^M \cdot 2/\pi^M.$$

We deduce that (W4) follows with $C_\varphi = 1$ when $M \geqslant 1$.

## B.2 Proof of Proposition 10

This section aims at recovering similar results than those given in [Moulines et al., 2007, Proposition 3] with explicit constants.

### B.2.1 Proof of inequality (22)

Observe that $t \mapsto \sum_{k\in\mathbb{Z}} \widehat{\varphi}(\lambda+2k\pi)\mathrm{e}^{\mathrm{i}\,t\,(\lambda+2k\pi)}$ is $2\pi$-periodic, integrable, and that its $\ell$-th Fourier coefficient is $2\pi\varphi(t-\ell)$,

$$\frac{1}{2\pi}\int_{-\infty}^{\infty}\Big(\sum_{k\in\mathbb{Z}}\widehat{\varphi}(\lambda+2k\pi)\mathrm{e}^{\mathrm{i}\,t\,(\lambda+2k\pi)}\Big)\mathrm{e}^{-\mathrm{i}\ell\lambda}\mathrm{d}\lambda = \frac{1}{2\pi}\int_{-\infty}^{\infty}\widehat{\varphi}(\lambda)\mathrm{e}^{\mathrm{i}(t-\ell)\lambda}\mathrm{d}\lambda = \varphi(t-\ell).$$

It follows that

$$\sum_{\ell\in\mathbb{Z}}\varphi(t-\ell)\mathrm{e}^{-\mathrm{i}\,t\,\lambda} = \sum_{k\in\mathbb{Z}}\widehat{\varphi}(\lambda+2k\pi)\mathrm{e}^{\mathrm{i}\,t\,(\lambda+2k\pi)}.$$

Hence, as in [Moulines et al., 2007, p180],

$$\begin{aligned}
\widehat{\tau}_j(\lambda) &= \int_{-\infty}^{\infty}\Big(\sum_{k\in\mathbb{Z}}\widehat{\varphi}(\lambda+2k\pi)\mathrm{e}^{\mathrm{i}\,t\,(\lambda+2k\pi)}\Big)2^{-j/2}\psi(2^{-j}t)\mathrm{d}t\\
&= \sum_{k\in\mathbb{Z}}\widehat{\varphi}(\lambda+2k\pi)\int_{-\infty}^{\infty}2^{-j/2}\psi(2^{-j}t)\mathrm{e}^{\mathrm{i}\,t\,(\lambda+2k\pi)}\mathrm{d}t\\
&= 2^{j/2}\sum_{k\in\mathbb{Z}}\widehat{\varphi}(\lambda+2k\pi)\overline{\widehat{\psi}(2^j(\lambda+2k\pi))}.
\end{aligned}$$

We obtain

$$\left|\widehat{\tau}_j(\lambda) - 2^{j/2}\widehat{\varphi}(\lambda)\overline{\widehat{\psi}(2^{-j}\lambda)}\right| = 2^{j/2}\left|\sum_{k\in\mathbb{Z},k\neq 0}\widehat{\varphi}(\lambda+2k\pi)\overline{\widehat{\psi}(2^j(\lambda+2k\pi))}\right|. \tag{29}$$

Property (W3) yields, for all $\lambda\in(-\pi,\pi)$,

$$\left|\widehat{\psi}(2^j(\lambda+2k\pi))\right| \leqslant \frac{C_s}{|2^j(\lambda+2k\pi)|^\alpha} \leqslant \frac{C_s}{(2^j\pi\,(2\,|k|-1))^\alpha}.$$

Using the inequality above and (W4) in (29), we get

$$\left|\widehat{\tau}_j(\lambda) - 2^{j/2}\widehat{\varphi}(\lambda)\overline{\widehat{\psi}(2^j\lambda)}\right| \leqslant C_\tau 2^{j(1/2-\alpha)}|\lambda|^M,$$

with $C_\tau = C_s C_\varphi 2\zeta(\alpha)$, with $\zeta(\cdot)$ the Riemann's zeta-function. Since $\alpha > 1$, $C_\tau < \infty$. If $\alpha \geqslant 2$, $\zeta(\alpha) \leqslant \pi^2/6 < 2$, and, hence, one can take $C_\tau = 4\, C_s C_\varphi$.

### B.2.2 Proof of inequality (23) and inequality (24)

Using (W2), (W3) and inequality (21), inequalities (23) and (24) are straightforward.

### B.2.3 Proof of inequality (25) and inequality (26)

With inequalities (22), (23) and (24), we get

$$\begin{aligned} &|\widehat{\tau}_j(\lambda)| \leqslant C_\infty C_m \left|2^j\lambda\right|^M + C_\tau 2^{j(1/2-\alpha)}|\lambda|^M \\ \text{and} \quad &|\widehat{\tau}_j(\lambda)| \leqslant \frac{2^{j/2}}{(1+|2^j\lambda|)^\alpha}\Big(C_\infty C_s 2^{-j/2} + C_\tau 2^{-j\alpha}(1+2^j|\lambda|)^\alpha|\lambda|^M\Big). \end{aligned}$$

It follows that

$$\begin{aligned} &|\widehat{\tau}_j(\lambda)| \leqslant C_{m\tau}\, \left|2^j\lambda\right|^M \\ \text{and} \quad &2^{-j/2}\left|\widehat{\tau}_j(2^{-j}\lambda)\right| \leqslant C_{s\tau}\frac{1}{(1+|\lambda|)^\alpha} \quad \text{when } |\lambda| \leqslant \pi, \end{aligned}$$

with $C_{m\tau} = C_\infty\, C_m + C_\tau$, and $C_{s\tau} = C_\infty\, C_s + C_\tau(1+\pi)^{\alpha+M}$.

### B.2.4 Proof of inequality (27)

With (25), inequality (27) is straightforward with $C_{ms\tau} = C_{m\tau}(1+\pi)^{\alpha+M}$.

### B.3 Proof of Proposition 11

Let $|\lambda| \leqslant \pi$. Inequalities (22) and (23) imply that

$$\begin{aligned}
&\left| \left|2^{-j/2}\widehat{\tau}(2^{-j}\lambda)\right|^2 - \left|\widehat{\varphi}(2^{-j}\lambda)\overline{\widehat{\psi}(\lambda)}\right|^2 \right| \\
&\leqslant \left|2^{-j/2}\widehat{\tau}(2^{-j}\lambda) - \widehat{\varphi}(2^{-j}\lambda)\overline{\widehat{\psi}(\lambda)}\right| \left( \left|2^{-j/2}\widehat{\tau}(2^{-j}\lambda)\right| + \left|\widehat{\varphi}(2^{-j}\lambda)\overline{\widehat{\psi}(\lambda)}\right| \right) \\
&\leqslant C_\tau 2^{-j\alpha} \left|2^{-j}\lambda\right|^M \left(C_{m\tau} + C_\infty C_m\right) |\lambda|^M \\
&\leqslant C_{a1}\, 2^{-\gamma_1 j}\, |\lambda|^{2M} \ ,
\end{aligned}$$

with $C_{a1} = C_\tau \left(C_{m\tau} + C_m C_\infty\right)$ and $\gamma_1 = M + \alpha$.

Next,

$$\left| \left|\widehat{\varphi}(2^{-j}\lambda)\overline{\widehat{\psi}(\lambda)}\right|^2 - \left|\widehat{\psi}(\lambda)\right|^2 \right| \leqslant \left|\widehat{\psi}(\lambda)\right|^2 \left| \left|\widehat{\varphi}(2^{-j}\lambda)\right|^2 - 1 \right| \leqslant C_m^2 \, |\lambda|^{2M} \left| \left|\widehat{\varphi}(2^{-j}\lambda)\right|^2 - 1 \right| .$$

The control of the right-hand side is obtained with the following result.

**Lemma 12.** *There exists a constant $C_Z = (M + L + 1)$ such that for all $j \in \mathbb{N}$, for all $\left|2^{-j}\lambda\right| < \pi$,*

$$\begin{aligned}
\left| 2\left|\widehat{\varphi}_h(2^{-j}\lambda)\right|^2 - 1 \right| &\leqslant C_Z \, \left|2^{-j}\lambda\right|^2 , \\
\left| 2\left|\widehat{\varphi}_g(2^{-j}\lambda)\right|^2 - 1 \right| &\leqslant C_Z \, \left|2^{-j}\lambda\right|^2 .
\end{aligned}$$

*Proof.* The proof is only derived for $\widehat{\varphi}_h(\cdot)$. It is similar for $\widehat{\varphi}_g(\cdot)$. Recall that

$$2^{1/2}\widehat{\varphi}_h(2^{-j}\lambda) = \left(\frac{\sin(\lambda/2^{j+1})}{\lambda/2^{j+1}}\right)^M \prod_{\ell=j+1}^{\infty} \overline{\widehat{d}_L(2^{-\ell}\lambda)}.$$

The triangular inequality implies that

$$\begin{aligned}\left|2\left|\widehat{\varphi}_h(2^{-j}\lambda)\right|^2-1\right| &\leqslant \left|\Big(\frac{\sin(\lambda/2^{j+1})}{\lambda/2^{j+1}}\Big)^{2M}-1\right|\prod_{\ell=j+1}^{\infty}\left|\widehat{d}_L(2^{-\ell}\lambda)\right|^2+\left|\prod_{\ell=j+1}^{\infty}\left|\widehat{d}_L(2^{-\ell}\lambda)\right|^2-1\right| \\ &\leqslant 2M\,\left|\left|\frac{\sin(\lambda/2^{j+1})}{\lambda/2^{j+1}}\right|-1\right|+\sum_{\ell=j+1}^{\infty}\left|\left|\widehat{d}_L(2^{-\ell}\lambda)\right|^2-1\right|,\end{aligned}$$

where we have used the equality $(x^K-1)=(x-1)\sum_{m=0}^{K-1}x^m$ for all $x\in\mathbb{R}$, $K\in\mathbb{N}$, and the fact that for all $x\in\mathbb{R}\setminus\{0\}$, $|\sin(x)/x|\leqslant 1$ and $\left|\widehat{d}_L(x)\right|\leqslant 1$.

Taylor inequality states that for all $x\in\mathbb{R}\setminus\{0\}$, $|\sin(x)-x|\leqslant |x|^3/6$. Additionally, for all $\left|2^{-\ell}\lambda/4\right|\leqslant 1$,

$$\begin{aligned}\left|\left|\widehat{d}_L(2^{-\ell}\lambda)\right|^2-1\right| &= \left|\cos(2^{-\ell}\lambda/4)^{2(2L+1)}-1+\sin(2^{-\ell}\lambda/4)^{2(2L+1)}\right| \\ &\leqslant (2L+1)\sin\big(2^{-\ell}\lambda/4\big)^2+\big(2^{-\ell}\lambda/4\big)^{2(2L+1)} \\ &\leqslant (2L+1)\big(2^{-\ell}\lambda/4\big)^2+\big(2^{-\ell}\lambda/4\big)^2.\end{aligned}$$

We get

$$\begin{aligned}\left|2\left|\widehat{\varphi}_h(2^{-j}\lambda)\right|^2-1\right| &\leqslant \frac{M}{3}\,\left|2^{-j-1}\lambda\right|^2+(2L+1)\left|2^{-j}\lambda\right|^2\sum_{\ell=3}^{\infty}2^{-2\ell}+\left|2^{-j}\lambda\right|^2\sum_{\ell=3}^{\infty}2^{-2\ell} \\ &\leqslant \big(\frac{M}{6}+\frac{L}{2}+\frac{1}{2}\big)\,\left|2^{-j}\lambda\right|^2.\end{aligned}$$

Hence,

$$\left|2\left|\widehat{\varphi}_h(2^{-j}\lambda)\right|^2-1\right|\leqslant C_Z\,\left|2^{-j}\lambda\right|^2,$$

with $C_Z=(M+L+1)/2$. □

We deduce that

$$\left|\left|\widehat{\varphi}(2^{-j}\lambda)\overline{\widehat{\psi}(\lambda)}\right|^2-\left|\widehat{\psi}(\lambda)\right|^2\right|\leqslant C_Z\left|2^{-j}\lambda\right|^2\left|\widehat{\psi}(\lambda)\right|^2\leqslant C_{a2}\,2^{-\gamma_2 j}\,|\lambda|^{2M}\,,$$

with $C_{a2}=C_m^2\,C_Z\pi^2$ and $\gamma_2=2$, when $|\lambda|\leqslant\pi$.

We obtain inequality (28) with $C_a = \max\{C_{a1}, C_{a2}\}$, $\gamma = \min\{\gamma_1, \gamma_2\}$. Observe that for all $M \geqslant 2$ $\gamma_1 > 2$, and, hence, $\gamma = 2$ and we can take $C_a = 2\,(M + L + 1) \geqslant C_{a2}$.

# C Asymptotic behavior of the wavelet covariance

This section deals with the proofs of the results of Section 3. We will prove stronger results than Proposition 3 and Proposition 4 which are stated below. To better highligth the role of the number of vanishing moments $M$ and the regularity $\alpha$, we keep these parameters, even if, for CFW-PR(M,L) and CFW-C(,L) filters, we have $\alpha = M$ by Proposition 9. Hence, we formulate here the assumption on the parameters on both $\alpha$ and $M$.

(C-c) $-\alpha + \beta/2 + 1/2 < d_\ell < M \quad$ for all $\quad \ell = 1, \ldots, p$, $M \geqslant 2$ and $0 < \beta < 2$.

Assumption (C-c) is equivalent to assumption (C-a).

Proposition 3 follows from the following proposition.

**Proposition 13.** *Let* $\mathbf{X}$ *be a* $p$*-multivariate long range dependent process with long memory parameters* $d_1, \ldots, d_p$ *with normalized spectral density* $f(\cdot)$ *satisfying (M-1) with short-range behavior (M-2). Consider* $\{W_{j,k}(\ell), \ell = 1, \ldots, p, j \geqslant 0, k \in \mathbb{Z}\}$ *the wavelet coefficients of* $\mathbf{X}$ *obtained with CFW-C(M,L) filters,* $M, L \geqslant 2$*. Then we have, for all* $j \geqslant 0$*,* $k \in \mathbb{Z}$*,*

$$\left| 2^{-j\,(d_\ell + d_m)} \mathrm{Cov}(W_{j,k}(\ell), W_{j,k}(m)) - \Omega_{\ell,m} \int_{-\infty}^{\infty} |\lambda|^{-d_\ell - d_m}\, \mathrm{e}^{\mathrm{sign}(\lambda)\,\phi_{\ell,m}} \left| \widehat{\psi}(\lambda) \right|^2 \mathrm{d}\lambda \right| \leqslant C_1' \max\{2^{-j\beta},\, L\, 2^{-2\,j}\}.$$

*where* $C_1'$ *is a constant only depending on* $M$ *and* $C_f, \beta, \|\mathbf{\Omega}\|, \{d_\ell, \ell = 1, \ldots, p\}$*.*

Next, Proposition 4 is a consequence of the following proposition.

**Proposition 14.** *Let* $\mathbf{X}$ *be a* $p$*-multivariate long range dependent process with long memory parameters* $d_1, \ldots, d_p$ *with normalized spectral density satisfying (M-1)–(M-2). Consider* $\{W_{j,k}(\ell),\ (j,k) \in \mathbb{Z},\ \ell = 1, \ldots, p\}$ *the wavelet coefficients obtained with CFW-C(M,L) filters,* $M, L \geqslant 2$*. Suppose that*

*(C-c) hold and that $L$ goes to infinity and $L2^{-2j} \to 0$ when $j$ goes to infinity. Then, for all $j \geqslant 0$,*

$$\left| 2^{-j\,(d_\ell + d_m)} \operatorname{Cov}(W_{j,k}(\ell), W_{j,k}(m)) - 4\,\Theta_{\ell,m} \int_0^\infty |\lambda|^{-d_\ell - d_m} \left| \widehat{\psi}_h(\lambda) \right|^2 \mathrm{d}\lambda \right| \leqslant C_2 \, \min\{2^{-j\beta}, L2^{-2j}, L^{-(2M - d_\ell - d_m + 1)}\} \,,$$

*where $C_2$ is a constant only depending on $M$ and $C_f, \beta, \|\mathbf{\Omega}\|, \{d_\ell, \ell = 1, \ldots, p\}$.*

## C.1 Proof of Proposition 13

Let $j \geqslant 0$, $k \in \mathbb{Z}$. The quantity $\operatorname{Cov}(\mathbf{W}_{j,k})$ can be decomposed as

$$\operatorname{Cov}(\mathbf{W}_{j,k}) = \mathbf{A}_j^{(+)} + \mathbf{A}_j^{(-)} \,, \quad \text{with} \quad \mathbf{A}_j^{(+)} = \int_0^{\pi\, 2^j} \boldsymbol{f}(2^{-j}\lambda)\, 2^{-j} \left| \widehat{\tau}_j(2^{-j}\lambda) \right|^2 \mathrm{d}\lambda \,,$$
$$\mathbf{A}_j^{(-)} = \int_{-\pi\, 2^j}^0 \boldsymbol{f}(2^{-j}\lambda)\, 2^{-j} \left| \widehat{\tau}_j(2^{-j}\lambda) \right|^2 \mathrm{d}\lambda \,.$$

Also recall that $\mathbf{\Gamma}_j(\mathbf{d})$ is the diagonal matrix with diagonal entries $2^{-j\,d_1}, \ldots, 2^{-j\,d_p}$.

We now sum up the main points for the convergence of $\operatorname{Cov}(\mathbf{W}_{j,k})$.

1. Behavior of $\mathbf{A}_j^{(+)}$.
   We introduce

$$\mathbf{B}_j^{(+)} = \int_0^{\pi\, 2^j} \mathbf{\Gamma}_j(\mathbf{d})^{-1} \mathbf{\Lambda}(\lambda) \mathbf{\Theta} \mathbf{\Lambda}(\lambda) \mathbf{\Gamma}_j(\mathbf{d})^{-1}\, 2^{-j} \left| \widehat{\tau}_j(2^{-j}\lambda) \right|^2 \mathrm{d}\lambda,$$
$$\mathbf{I}_j^{(+)\mathrm{inf}} = \int_0^{\pi\, 2^j} \mathbf{\Gamma}_j(\mathbf{d})^{-1} \mathbf{\Lambda}(\lambda) \mathbf{\Theta} \mathbf{\Lambda}(\lambda) \mathbf{\Gamma}_j(\mathbf{d})^{-1} \left| \widehat{\psi}(\lambda) \right|^2 \mathrm{d}\lambda,$$
$$\mathbf{I}_j^{(+)\mathrm{sup}} = \int_{\pi\, 2^j}^\infty \mathbf{\Gamma}_j(\mathbf{d})^{-1} \mathbf{\Lambda}(\lambda) \mathbf{\Theta} \mathbf{\Lambda}(\lambda) \mathbf{\Gamma}_j(\mathbf{d})^{-1} \left| \widehat{\psi}(\lambda) \right|^2 \mathrm{d}\lambda,$$
$$\mathbf{I}^{(+)} = \mathbf{I}_j^{(+)\mathrm{inf}} + \mathbf{I}_j^{(+)\mathrm{sup}} = \int_0^\infty \mathbf{\Gamma}_j(\mathbf{d})^{-1} \mathbf{\Lambda}(\lambda) \mathbf{\Theta} \mathbf{\Lambda}(\lambda) \mathbf{\Gamma}_j(\mathbf{d})^{-1} \left| \widehat{\psi}(\lambda) \right|^2 \mathrm{d}\lambda.$$

   The steps of the convergence are:

   (a) $2^{-j(d_\ell + d_m - \beta)} \left| \mathbf{A}_j^{(+)} - \mathbf{B}_j^{(+)} \right|$ is bounded using the regularity of the spectral density $\boldsymbol{f}^S(\cdot)$

at the origin, that is, (M-2), together with inequality (25).

(b) $2^{-j(d_\ell+d_m-\beta)}\left|\mathbf{B}_j^{(+)}-\mathbf{I}_j^{(+)\inf}\right|$ is bounded using the convergence of the filter $\tau_j$ to $\widehat{\psi}(.)$, through property (28). We shall need inequality (25) to control the integral around zero and inequality (26) to control the upper part.

(c) $2^{-j(d_\ell+d_m-\beta)}\left|\mathbf{I}_j^{(+)\sup}\right|$ is bounded using the regularity of $\widehat{\psi}(.)$, that is, using (W3).

All together, we shall obtain the convergence of $\mathbf{A}_j^{(+)}$ to $\mathbf{I}^{(+)}$, which gives the property.

2. Behavior of $\mathbf{A}_j^{(-)}$
We can apply the same arguments as for $\mathbf{A}_j^{(+)}$ and obtain the convergence of $\mathbf{A}_j^{(-)}$ to $\mathbf{I}^{(-)}$, with
$$\mathbf{I}^{(-)}=\int_{-\infty}^{0}\mathbf{\Gamma}_j(\mathbf{d})^{-1}\mathbf{\Lambda}(\lambda)\mathbf{\Theta}\mathbf{\Lambda}(\lambda)\mathbf{\Gamma}_j(\mathbf{d})^{-1}\ \left|\widehat{\psi}(\lambda)\right|^2\ \mathrm{d}\lambda.$$

In the following, $(\ell,m)\in\{1,\dots,p\}^2$ will denote two arbitrary indexes.

### C.1.1 Spectral approximation, $\left|\mathbf{A}_j^{(+)}-\mathbf{B}_j^{(+)}\right|$

First notice that $\mathbf{\Gamma}_j(\mathbf{d})^{-1}\mathbf{\Lambda}(2^j\lambda)=\mathbf{\Lambda}(\lambda)$. Hence,
$$\begin{aligned}\left|\mathbf{A}_j^{(+)}-\mathbf{B}_j^{(+)}\right| &\leqslant \int_0^\pi \left|\boldsymbol{f}(\lambda)\ -\mathbf{\Lambda}(\lambda)\mathbf{\Theta}\mathbf{\Lambda}(\lambda)\right|\left|\widehat{\tau}_j(\lambda)\right|^2\ \mathrm{d}\lambda\\ &\leqslant \int_0^\pi \left|\mathbf{\Lambda}(\lambda)\mathbf{\Theta}\mathbf{\Lambda}(\lambda)\right|\circ\left|\boldsymbol{f}^S(\lambda)-1\right|\left|\widehat{\tau}_j(\lambda)\right|^2\ \mathrm{d}\lambda.\end{aligned}$$

Property (M-2) gives
$$\left(\left|\mathbf{\Lambda}(\lambda)\mathbf{\Theta}\mathbf{\Lambda}(\lambda)\right|\circ\left|\boldsymbol{f}^S(\lambda)-1\right|\right)_{\ell,m}\leqslant C_f\ \|\mathbf{\Omega}\|\ |\lambda|^{-d_\ell-d_m+\beta}\,.$$

With a change of variable,
$$\left|\mathbf{A}_j^{(+)}-\mathbf{B}_j^{(+)}\right|_{\ell,m}\leqslant C_f\ \|\mathbf{\Omega}\|\ 2^{j(d_\ell+d_m-\beta)}\int_0^{2^j\pi}|\lambda|^{-d_\ell-d_m+\beta}\left|2^{-j/2}\widehat{\tau}_j(2^{-j}\lambda)\right|^2\ \mathrm{d}\lambda.$$

We split the integral in two parts. First, with (25),

$$\int_0^1 |\lambda|^{-d_\ell - d_m + \beta} \left| 2^{-j/2} \widehat{\tau}_j(2^{-j}\lambda) \right|^2 \mathrm{d}\lambda \leqslant C_{m\tau}^2 \int_0^1 |\lambda|^{-d_\ell - d_m + \beta + 2M} \mathrm{d}\lambda.$$

As the parameters satisfy (C-c), the integral is bounded by a constant depending on $(d_\ell, d_m, \beta, M)$. The bound is independent on $L$ since the constant $C_{m\tau}$ does not depend on $L$.

Next, using the regularity given by (26),

$$\int_1^{2^j \pi} |\lambda|^{-d_\ell - d_m + \beta} \left| 2^{-j/2} \widehat{\tau}_j(2^{-j}\lambda) \right|^2 \mathrm{d}\lambda \leqslant C_{s\tau}^2 \int_1^{\infty} \frac{|\lambda|^{-d_\ell - d_m + \beta}}{(1+|\lambda|)^{2\alpha}} \mathrm{d}\lambda.$$

Property (C-c) ensures that the right hand side is bounded by a constant depending on $d_\ell, d_m$, $\beta$, $\alpha$, $M$, and not depending on $L$.

### C.1.2 Asymptotic of the filters, $\left| \mathbf{B}_j^{(+)} - \mathbf{I}_j^{(+)\mathbf{inf}} \right|$

This step uses the convergence of the filter $\widehat{\tau}_j$ to $\widehat{\psi}(.)$, through property (28). First,

$$\mathbf{\Gamma}_j(\mathbf{d})\, 2^{j\beta} \left| \mathbf{B}_j^{(+)} - \mathbf{I}_j^{(+)\mathrm{inf}} \right| \mathbf{\Gamma}_j(\mathbf{d}) \;\leqslant\; 2^{j\beta} \int_0^{2^j\pi} \mathbf{\Lambda}(\lambda) \mathbf{\Theta} \mathbf{\Lambda}(\lambda) \; \left| \left| 2^{-j/2} \widehat{\tau}_j(2^{-j}\lambda) \right|^2 - \left| \widehat{\psi}(\lambda) \right|^2 \right| \mathrm{d}\lambda.$$

Using (W3) and (26),

$$\begin{aligned} \left| \left| 2^{-j/2} \widehat{\tau}_j(2^{-j}\lambda) \right|^2 - \left| \widehat{\psi}(\lambda) \right|^2 \right| &\leqslant \left| 2^{-j/2} \widehat{\tau}_j(2^{-j}\lambda) \right|^2 + \left| \widehat{\psi}(\lambda) \right|^2 \\ &\leqslant (C_{s\tau}^2 + C_s^2)\, |\lambda|^{-2\alpha}. \end{aligned}$$

Thus,

$$\begin{aligned}
&\left(\int_1^{2^j\pi} \mathbf{\Lambda}(\lambda)\mathbf{\Theta}\mathbf{\Lambda}(\lambda)\, 2^{j\beta} \left|\left|2^{-j/2}\widehat{\tau}_j(2^{-j}\lambda)\right|^2 - \left|\widehat{\psi}(\lambda)\right|^2\right| \,\mathrm{d}\lambda\right)_{(\ell,m)} \\
&\leqslant \|\mathbf{\Omega}\|\, (C_{s\tau}^2 + C_s^2)\, 2^{j\beta} \int_1^{2^j\pi} |\lambda|^{-d_\ell - d_m - 2\alpha}\, \mathrm{d}\lambda \\
&\leqslant \|\mathbf{\Omega}\|\, (C_{s\tau}^2 + C_s^2)\, \pi^{-d_\ell - d_m - 2\alpha + 1}\, 2^{j(-d_\ell - d_m - 2\alpha + \beta + 1)},
\end{aligned}$$

where last inequality follows from the evaluation of the integral term. The right-hand side does not depend on $L$, and it tends to 0 when $j$ goes to infinity due to (C-c).

It remains to consider the integral on $(0, 1)$. Property (28) states that

$$\begin{aligned}
&\left(\int_0^1 \mathbf{\Lambda}(\lambda)\mathbf{\Theta}\mathbf{\Lambda}(\lambda)\, 2^{j\beta} \left|\left|2^{-j/2}\widehat{\tau}_j(2^{-j}\lambda)\right|^2 - \left|\widehat{\psi}(\lambda)\right|^2\right| \,\mathrm{d}\lambda\right)_{(\ell,m)} \\
&\qquad\qquad \leqslant \|\mathbf{\Omega}\|\, C_a\, 2^{j(\beta-\gamma)} \int_0^1 |\lambda|^{-d_\ell - d_m + 2M}\, \mathrm{d}\lambda.
\end{aligned}$$

The right-hand side tends to 0 when $j$ goes to infinity since $\beta < \gamma$ and $\max\{d_\ell, \ell = 1, \ldots, p\} < M + 1/2$.

When $M \geqslant 2$, $\gamma = 2$ and $C_a = 2\,(M + L + 1)$. With a fixed $M$, this term is, hence, bounded up to a constant by $\max\{1, L\, 2^{j(\beta-\gamma)}\} = \max\{1, L\, 2^{j(\beta-2)}\}$.

### C.1.3 Regularity of the filters, $\left|\mathbf{I}_j^{(+)\mathbf{sup}}\right|$

This step uses the regularity of $\widehat{\psi}$. Indeed, property (W3) entails that

$$\begin{aligned}
\left|\mathbf{I}_{j,\ell m}^{(+)\mathrm{sup}}\right| &\leqslant \|\mathbf{\Omega}\|\, 2^{j(d_\ell + d_m)} \int_{2^j\pi}^{\infty} |\lambda|^{-d_\ell - d_m} \left|\widehat{\psi}(\lambda)\right|^2 \,\mathrm{d}\lambda \\
&\leqslant C_s^2\, \|\mathbf{\Omega}\|\, 2^{j(d_\ell + d_m)} \int_{2^j\pi}^{\infty} \frac{|\lambda|^{-d_\ell - d_m}}{(1 + |\lambda|)^{2\alpha}}\, \mathrm{d}\lambda \\
&\leqslant C_s^2\, \|\mathbf{\Omega}\|\, \pi^{-\beta}\, 2^{j(d_\ell + d_m - \beta)} \int_{2^j\pi}^{\infty} |\lambda|^{-d_\ell - d_m + \beta - 2\alpha}\, \mathrm{d}\lambda,
\end{aligned}$$

where last inequality results from the fact that when $|\lambda| \geqslant 2^j\pi$, then $1 \leqslant |\lambda|^\beta 2^{-j\beta}\pi^\beta$. Property (C-c) thus implies that $2^{-j(d_\ell+d_m-\beta)}\left|\mathbf{I}_{j,\ell m}^{(+)\sup}\right|$ is bounded by a constant depending of $(d_\ell, d_m, \beta, \|\mathbf{\Omega}\|, M)$.

## C.2 Proof of Proposition 14

Recall that

$$\begin{aligned}
\widehat{\psi}(\lambda) = \widehat{\psi}_h(\lambda) + \mathrm{i}\,\widehat{\psi}_g(\lambda) &= \left(1 - \mathrm{e}^{\mathrm{i}\eta_L(\lambda)}\right)\widehat{\psi}_h(\lambda)\ , \\
\text{with}\quad \alpha_L(\lambda) &= 2(-1)^L \operatorname{atan}\left(\tan^{2L+1}(\lambda/4)\right)\ , \\
\eta_L(\lambda) &= -a_L(\lambda/2+\pi) + \sum_{j=1}^{\infty} a_L(2^{-j-1}\lambda)\ .
\end{aligned}$$

Theorem 2 states that, for all $\lambda \in \mathbb{R}$,

$$\left|\widehat{\psi}(\lambda) - 2\mathbb{1}_{\mathbb{R}_+}(\lambda)\,\widehat{\psi}_h(\lambda)\right| = U_L(\lambda)\left|\widehat{\psi}_h(\lambda)\right|\ ,$$

with

$$U_L(\lambda) \leqslant 2\sqrt{2}\left(\log_2\left(\frac{\max(4\pi,|\lambda|)}{2\pi}\right)+2\right)\left(1-\frac{\delta(\lambda,4\pi\mathbb{Z})}{\max(4\pi,|\lambda|)}\right)^{2L+1}\ .$$

We deduce from Theorem 2 the following result, which gives a similar inequality but in a form that can be more useful in future developments.

**Corollary 15.** *For all $\widehat{q}_{L,M}$ real polynomial of $(e^{-i\lambda})$, for all $|\lambda| \leqslant 2\pi$,*

$$\left|\left|\widehat{\psi}(\lambda)\right|^2 - 4\,\mathbb{1}_{\mathbb{R}_+}(\lambda)\,\left|\widehat{\psi}_h(\lambda)\right|^2\right| \leqslant 18\sqrt{2}\left(1-\frac{|\lambda|}{2\pi}\right)^{2L+1}\left|\widehat{\psi}_h(\lambda)\right|^2\ .$$

Let us introduce also

$$\begin{aligned}
\mathbf{I}_j^{\inf} &= \int_{-\pi 2^j}^{\pi 2^j} \mathbf{\Gamma}_j(\mathbf{d})^{-1}\mathbf{\Lambda}(\lambda)\mathbf{\Theta}\mathbf{\Lambda}(\lambda)\mathbf{\Gamma}_j(\mathbf{d})^{-1}\ \left|\widehat{\psi}(\lambda)\right|^2\ \mathrm{d}\lambda, \\
\mathbf{I}_j^{\inf,\text{analytic}} &= \int_{0}^{\pi 2^j} \mathbf{\Gamma}_j(\mathbf{d})^{-1}\mathbf{\Lambda}(\lambda)\mathbf{\Theta}\mathbf{\Lambda}(\lambda)\mathbf{\Gamma}_j(\mathbf{d})^{-1}\,4\,\left|\widehat{\psi}_h(\lambda)\right|^2\ \mathrm{d}\lambda, \\
\mathbf{I}_j^{\sup,\text{analytic}} &= \int_{\pi 2^j}^{\infty} \mathbf{\Gamma}_j(\mathbf{d})^{-1}\mathbf{\Lambda}(\lambda)\mathbf{\Theta}\mathbf{\Lambda}(\lambda)\mathbf{\Gamma}_j(\mathbf{d})^{-1}\,4\,\left|\widehat{\psi}_h(\lambda)\right|^2\ \mathrm{d}\lambda.
\end{aligned}$$

Following the proof of Proposition 13, the steps of the proof are the following:

1. $2^{-j(d_\ell+d_m-\beta)}\left|\mathrm{Cov}(\mathbf{W}_{j,k})-\mathbf{I}_j^{\mathrm{inf}}\right|$ is bounded, up to a constant, by $L2^{-2j}$. This result was already obtained in the proof of Proposition 13.
2. $2^{-j(d_\ell+d_m-\beta)}\left|\mathbf{I}_j^{\mathrm{inf}}-\mathbf{I}_j^{\mathrm{inf,analytic}}\right|$ is bounded up to a constant by $\max\{1, L^{-2M-d_\ell-d_m+1}2^{j\beta}\}$, using the quasi-analyticity property, stated in Corrolary 15.
3. $2^{-j(d_\ell+d_m-\beta)}\left|\mathbf{I}_j^{\mathrm{sup,analytic}}\right|$ is bounded. This result is straigthforward with step 1.(c) in the proof of Proposition 13, since $|\widehat{\psi}_h(\lambda)|\leqslant|\widehat{\psi}_h(\lambda)|$.

Hence, it only remains to prove step 2. That is, we want to establish that the quantity

$$2^{j\beta}\int_{-2^j\pi}^{2^j\pi}|\lambda|^{-d_\ell-d_m}\left|\left|\widehat{\psi}(\lambda)\right|^2-4\,\mathbb{1}_{\mathbb{R}_+}(\lambda)\,\left|\widehat{\psi}_h(\lambda)\right|^2\right|\mathrm{d}\lambda$$

is bounded up to a constant by $\max\{1, L^{-2M-d_\ell-d_m+1}2^{j\beta}\}$.

We decompose the integral on the sub-intervals $(-1,1)$, $(-2^j\pi,-1)$ and $(1,2^j\pi)$.

**On $(-1,1)$.**

With assumption (W2), Corollary 15 leads to

$$\begin{aligned}\int_{-1}^{1}|\lambda|^{-d_\ell-d_m}\left|\left|\widehat{\psi}(\lambda)\right|^2-4\,\mathbb{1}_{\mathbb{R}_+}(\lambda)\,\left|\widehat{\psi}_h(\lambda)\right|^2\right|\mathrm{d}\lambda&\leqslant 36\sqrt{2}\int_0^1\left(1-\frac{|\lambda|}{4\pi}\right)^{2L+1}|\lambda|^{-d_\ell-d_m+2M}\mathrm{d}\lambda,\\&\leqslant 36\sqrt{2}\,B(2M-d_\ell-d_m,2L+1),\end{aligned}$$

where $B$ is the Beta function. Using Stirling's approximation, for fixed $M$, $d_\ell$, $d_m$, when $L$ goes to infinity the right-hand side is equivalent to

$$36\sqrt{2}\,\Gamma(2M-d_\ell-d_m+1)\,(2L+2)^{-(2M-d_\ell-d_m+1)}\,.$$

Up to a constant, this term is bounded by $\max\{1, L^{-(2M-d_\ell-d_m+1)}\}$.

**On $(-2^j\pi, -1)$ and $(1, 2^j\pi)$.**

(W3) entails that, for all $\lambda \in \mathbb{R}$,

$$\left| \left|\widehat{\psi}(\lambda)\right|^2 - 4\,\mathbb{1}_{\mathbb{R}_+}(\lambda)\,\left|\widehat{\psi}_h(\lambda)\right|^2 \right| \leqslant (C_s^2 + 4\cdot C_s^2)\,|\lambda|^{-2\alpha}.$$

Thus,

$$2^{j\beta}\int_1^{2^j\pi} |\lambda|^{-d_\ell - d_m + \beta} \left| \left|\widehat{\psi}(\lambda)\right|^2 - 4\,\mathbb{1}_{\mathbb{R}_+}(\lambda)\,\left|\widehat{\psi}_h(\lambda)\right|^2 \right| \mathrm{d}\lambda \leqslant 5\,C_s^2 \int_1^{\infty} |\lambda|^{-d_\ell - d_m + \beta - 2\alpha}\,\mathrm{d}\lambda\,,$$

and the same inequality holds when considering the integral on $(-2^j\pi, -1)$. The integral on the right hand side is bounded when (C-c) holds, for any $L, M \geqslant 1$.

# D Asymptotic behavior of the estimators

We detail some points that are changed with the complex wavelet setting in the proofs of consistence and of asymptotic normality, with respect to the real wavelets setting.

First, recall that, for all $j \geqslant 0$, $n_j$ denotes the number of non zero wavelet coefficients $\{\mathbf{W}_{j,k},\, k \in \mathbb{Z}\}$. Under the assumptions that $2^{-j_0}L$ is bounded and that $2^{-j_0}N \to \infty$, the sequence $n_j$ is equivalent to $2^{-j}N$ when $j$ goes to infinity. These assumptions are made in both Theorem 5 and Theorem 6. Hence, $n_j$ behaves similarly to in Achard and Gannaz [2016] and in Gannaz [2020].

## D.1 Proof of Theorem 5

For complex wavelets the approximation of the wavelet covariance does not admit the same bound as for real wavelets. Hence, the study of the term

$$S^{(1)}_{\ell,m}(\mu) = \sum_{j=j_0}^{j_1} n_j \mu_j \left[ \frac{\mathrm{Cov}(W_{j,k}(\ell), W_{j,k}(m))}{2^{j(d^0_\ell + d^0_m)}} - G^0_{\ell,m} \right].$$

defined on page 499 of Achard and Gannaz [2016] is modified. Consequently, Proposition 8 and Proosition 9 of Achard and Gannaz [2016] do not hold anymore. They are replaced respectively by Proposition 16 and Proposition 17 below.

Let us take $\ell$ and $m$ in $1, \dots, p$, and define, for any sequence $\mu = \{\mu_j,\, j \geqslant 0\}$,

$$S_{\ell,m}(\mu) = \sum_{j,k} \mu_j \left( \frac{W_{j,k}(\ell) W_{j,k}(m)}{2^{j(d^0_\ell + d^0_m)}} - G^0_{\ell,m} \right) = \sum_{j=j_0}^{j_1} \mu_j \left( \frac{I_{\ell,m}(j)}{2^{j(d^0_\ell + d^0_m)}} - n_j G^0_{\ell,m} \right).$$

$S_{\ell,m}(\mu)$ is decomposed in two terms $S^{(0)}_{\ell,m}(\mu)$ and $S^{(1)}_{\ell,m}(\mu)$,

$$S^{(0)}_{\ell,m}(\mu) = \sum_{j=j_0}^{j_1} \mu_j \frac{1}{2^{j(d^0_\ell + d^0_m)}} \sum_k \left( W_{j,k}(\ell) W_{j,k}(m) - \mathrm{Cov}(W_{j,k}(\ell), W_{j,k}(m)) \right),$$
$$S^{(1)}_{\ell,m}(\mu) = \sum_{j=j_0}^{j_1} n_j \mu_j \left[ \frac{\mathrm{Cov}(W_{j,k}(\ell), W_{j,k}(m))}{2^{j(d^0_\ell + d^0_m)}} - G^0_{\ell,m} \right].$$

**Proposition 16.** *Assume that the sequences $\mu$ belong to the set $\{\{\mu_j\}_{j\geqslant 0},\ |\mu_j| \leqslant \frac{1}{n_j}\}$. Suppose that (C-b) holds. Under condition (C), $\sup_{\{\mu,\, |\mu_j| \leqslant \frac{1}{n_j}\}} S_{\ell,m}(\mu)$ is uniformly bounded by $2^{-j_0\beta} + L2^{-2j_0} + j_1 L^{-1} + N^{-1/2} 2^{j_1/2}$ up to a multiplicative constant, that is,*

$$\sup_{\mu \in \{(\mu_j)_{j\geqslant 0},\ |\mu_j| \leqslant \frac{1}{n_j}\}} \{S_{\ell,m}(\mu)\} = O_{\mathbb{P}}(2^{-j_0\beta} + L2^{-2j_0} + j_1 L^{-1} + N^{-1/2} 2^{j_1/2}).$$

*Proof.* From Proposition 14, there exists $C' > 0$ such that

$$|S^{(1)}_{\ell,m}(\mu)| \leqslant C' \sum_{j=j_0}^{j_1} \left(2^{-\beta j} + L2^{-j} + L^{-(2M - d_\ell - d_m + 1)}\right) n_j |\mu_j|.$$

Assumption(C-b) yields

$$|S^{(1)}_{\ell,m}(\mu)| \leqslant C' \sum_{j=j_0}^{j_1} \left(2^{-\beta j} + L2^{-j} + L^{-1}\right) n_j |\mu_j|, \tag{30}$$

with $C > 0$. Under the assumption $|\mu_j| \leqslant \frac{1}{n_j}$, we have the inequality $|S^{(1)}_{\ell,m}(\mu)| \leqslant C \sum_{j=j_0}^{j_1} \left(2^{-\beta j} +$

$L2^{-j} + L^{-1}\big)$. The right-hand bound is equivalent to $2^{-j_0\beta} + L2^{-2j_0} + j_1 L^{-(2M-d_\ell-d_m+1)}$ up to a constant.

The term $S^{(0)}$ is unchanged and the proposition follows as in the proof of Proposition 8 of Achard and Gannaz [2016]. □

**Proposition 17.** *Let $0 < j_0 \leqslant j_1 \leqslant j_N$. Suppose that (C-b) holds. Assume that the sequences $\mu$ belong to the set*

$$\mathcal{S}(q,\gamma,c) = \{\{\mu_j\}_{j\geqslant 0}, |\mu_j| \leqslant \frac{c}{n}|j-j_0+1|^q 2^{(j-j_0)\gamma},\ \forall j = j_0,\dots j_1\}$$

*with $0 \leqslant \gamma < 1$. Under condition (C), $\sup_{\mu\in\mathcal{S}(q,\gamma,c)} S_{\ell,m}(\mu)$ is uniformly bounded by $2^{-j_0\beta} + L2^{-2j_0} + \log(N)L^{-1} + H_\gamma(N^{-1/2}2^{j_0/2})$ up to a constant,*

$$\sup_{\mu\in\mathcal{S}(q,\gamma,c)} \{S_{\ell,m}(\mu)\} = O_{\mathbb{P}}(2^{-j_0\beta} + L2^{-2j_0} + j_1 L^{-1} + H_\gamma(N^{-1/2}2^{j_0/2}))$$

$$\text{with } H_\gamma(u) = \begin{cases} u & \text{if } \ 0 \leqslant \gamma < 1/2, \\ \log(1+u^{-2})^{q+1}\, u & \text{if } \ \gamma = 1/2, \\ \log(1+u^{-2})^{q}\, u^{2(1-\gamma)} & \text{if } \ 1/2 < \gamma < 1. \end{cases}$$

In particular, for any $0 \leqslant \gamma < 1$, under the assumptions $2^{-j_0\beta} + N^{-1/2}2^{j_0/2} \to 0$, and $L2^{-2j_0} + \log(N)L^{-1} \to 0$, we have $\sup_{\mu\in\mathcal{S}(q,\gamma,c)}\{S_{\ell,m}(\mu)\} = o_{\mathbb{P}}(1)$.

*Proof.* Under the assumptions of the proposition, one deduce from inequality (30) that,

$$\begin{aligned}\sup_{\mu\in\mathcal{S}(q,\gamma,c)} |S^{(1)}_{\ell,m}(\mu)| &\leqslant cC\frac{1}{n}\sum_{j=j_0}^{j_1} n_j\big(2^{-\beta j} + L2^{-j} + L^{-(2M-d_\ell-d_m+1)}\big)2^{\gamma(j-j_0))}(j-j_0+1)^q \\ &\leqslant cC\, 2^{-\beta j_0}\sum_{i=0}^{j_1-j_0} 2^{-(1+\beta-\gamma)i}(i+1)^q + cC\, L2^{-j_0}\sum_{i=0}^{j_1-j_0} 2^{-(2-\gamma)i}(i+1)^q \\ &\quad + cC\, j_1 L^{-(2M-d_\ell-d_m+1)}\sum_{i=0}^{j_1-j_0} 2^{-(1-\gamma)i}(i+1)^q.\end{aligned}$$

The sum on the right-hand side of the inequality tends to 0 under the assumptions of the proposition,

since $1-\gamma > 0$.

The term $S^{(0)}$ is unchanged and the proposition follows as in the proof of Proposition 9 of Achard and Gannaz [2016]. □

The rest of the proof is very similar to the real case and it is omitted. Remark that a key of the proof is Oppenheim's inequality, which holds for complex matrices, see Horn and Johnson [1990].

### D.2 Proof of Theorem 6

The properties of wavelet filters in the proofs of Gannaz [2020] are used through Proposition 31 of Gannaz [2020]. The inequalities (I1) and (I2) in Gannaz [2020] correspond respectively to (27) and (22) of Proposition 10. Inequality (I3) in Gannaz [2020] follows with the proof of Proposition 31 of Gannaz [2020]. The constants in (27) and (22) do not depend of $L$, which allows to use these inequalities as in the proofs of Gannaz [2020].

The other property of the wavelets used in the proofs of Gannaz [2020] is the convergence of $|\widehat{\varphi}(2^{-j}\lambda)|$ to 1 when $j$ goes to infinity (page 29 of Gannaz [2020]). When $L2^{-2j}$ goes to infinity, Lemma 12 yields

$$\lim_{j\to\infty} \left| |\widehat{\varphi}(2^{-j}\lambda)|^2 - 1 \right| = 0,$$

which is the desired result.

Finally, the approximation of the sample wavelet covariance is changed. It is sufficient, to use results of Proposition 4 (instead of Proposition 1 of Gannaz [2020]), to check that, for all $(\ell, m) \in \{1,\dots,p\}^2$,

$$\lim_{j\to\infty} \sqrt{n_j} \left| 2^{-j\,(d_\ell + d_m - \beta)} \mathrm{Cov}(W_{j,k}(\ell), W_{j,k}(m)) - G_{\ell,m} \right| = 0.$$

Hence, based on Proposition 4, it is sufficient to have

$$\lim_{j\to\infty} N^{1/2} 2^{-j/2} \left( L2^{-2j} + L^{-(2M - d_\ell - d_m + 1)} \right) = 0,$$

since when $j$ goes to infinity, $n_j$ is equivalent to $N2^{-j} + 1$.

The rest of the proof does not present major changes. It is thus omitted.

## References


P. Abry and D. Veitch. Wavelet analysis of long-range-dependent traffic. *Information Theory, IEEE Transactions on*, 44(1):2–15, 1998.

S. Achard and I. Gannaz. *multiwave: Estimation of multivariate long-memory models parameters*, 2015. R package.

S. Achard and I. Gannaz. Multivariate wavelet Whittle estimation in long-range dependence. *Journal of Time Series Analysis*, 37:476–512, 2016.

S. Achard and I. Gannaz. Wavelet-based and Fourier-based multivariate Whittle estimation: multiwave. *Journal of Statistical Software, Articles*, 89(6):1–31, 2019.

S. Achard, M. Clausel, I. Gannaz, and F. Roueff. New results on approximate Hilbert pairs of wavelet filters with common factors. *Applied and Computational Harmonic Analysis*, 49(3): 1025–1045, 2020.

P.-O. Amblard, J.-F. Coeurjolly, F. Lavancier, and A. Philippe. Basic properties of the multivariate fractional brownian motion. In *Séminaires et congrès*, volume 28, pages 65–87, 2013.

C. Baek, S. Kechagias, and V. Pipiras. Asymptotics of bivariate local Whittle estimators with applications to fractal connectivity. *Journal of Statistical Planning and Inference*, 205:245–268, 2020.

G. G. J.-P. C. Becq, E. Barbier, and S. Achard. Brain networks of rats under anesthesia using resting-state fmri: comparison with dead rats, random noise and generative models of networks. *Journal of Neural Engineering*, 2020a. URL http://iopscience.iop.org/10.1088/1741-2552/ab9fec.

G. J.-P. Becq, T. Habet, N. Collomb, M. Faucher, C. Delon-Martin, V. Coizet, S. Achard, and E. L. Barbier. Functional connectivity is preserved but reorganized across several anesthetic regimes. *NeuroImage*, 219:116945, 2020b. doi: https://doi.org/10.1016/j.neuroimage.2020.116945.

R. B. Buxton. The physics of functional magnetic resonance imaging (fmri). *Reports on Progress in Physics*, 76(9):096601, 2013.

J.-F. Coeurjolly, P.-O. Amblard, and S. Achard. Wavelet analysis of the multivariate fractional Brownian motion. *ESAIM: Probability and Statistics*, 17:592–604, 2013.

G. Didier and V. Pipiras. Integral representations and properties of operator fractional Brownian motions. *Bernoulli*, 17(1):1–33, 2011.

I. Gannaz. Asymptotic normality of wavelet covariances and of multivariate wavelet whittle estimators. *arXiv preprint arXiv:2012.09436*, 2020.

I. Gannaz, S. Achard, M. Clausel, and F. Roueff. Analytic wavelets for multivariate time series analysis. In *Wavelets and Sparsity XVII*, volume 10394, page 103941X. International Society for Optics and Photonics, 2017.

R. Gençay, F. Selçuk, and B. J. Whitcher. *An introduction to wavelets and other filtering methods in finance and economics*. Academic Press, 2001.

J. Geweke and S. Porter-Hudak. The estimation and application of long memory time series models. *Journal of Time Series Analysis*, 4(4):221–238, 1983.

A. Hjorungnes and D. Gesbert. Complex-valued matrix differentiation: Techniques and key results. *IEEE Transactions on Signal Processing*, 55(6):2740–2746, 2007.

R. A. Horn and C. R. Johnson. *Matrix analysis*. Cambridge university press, 1990.

S. Kechagias and V. Pipiras. Definitions and representations of multivariate long-range dependent time series. *Journal of Time Series Analysis*, 2014. doi: 10.1111/jtsa.12086.

H. Künsch. Statistical aspects of self-similar processes. In *Proceedings of the 1st World Congress of the Bernoulli Society, Vol. 1 (Tashkent, 1986)*, pages 67–74, Utrecht, 1987. VNU Sci. Press.

I. N. Lobato. Consistency of the averaged cross-periodogram in long memory series. *Journal of Time Series Analysis*, 18(2):137–155, 1997.

I. N. Lobato. A semiparametric two-step estimator in a multivariate long memory model. *Journal of Econometrics*, 90(1):129–153, 1999.

S. Mallat. *A wavelet tour on signal processing*. Academic Press, 2 edition, 1999.

E. Moulines, F. Roueff, and M. S. Taqqu. On the spectral density of the wavelet coefficients of long-memory time series with application to the log-regression estimation of the memory parameter. *Journal of Time Series Analysis*, 28(2):155–187, 2007.

E. Moulines, F. Roueff, and M. S. Taqqu. A wavelet Whittle estimator of the memory parameter of a nonstationary Gaussian time series. *The Annals of Statistics*, pages 1925–1956, 2008.

P. M. Robinson. Semiparametric analysis of long-memory time series. *The Annals of Statistics*, pages 515–539, 1994.

P. M. Robinson. Log-periodogram regression of time series with long range dependence. *The Annals of Statistics*, pages 1048–1072, 1995a.

P. M. Robinson. Gaussian semiparametric estimation of long range dependence. *The Annals of Statistics*, 23(5):1630–1661, 1995b.

P. M. Robinson. Multiple local Whittle estimation in stationary systems. *The Annals of Statistics*, 36(05):2508–2530, 2008.

R. J. Sela and C. M. Hurvich. The averaged periodogram estimator for a power law in coherency. *Journal of Time Series Analysis*, 33(2):340–363, 2012.

I. W. Selesnick. Hilbert transform pairs of wavelet bases. *Signal Processing Letters, IEEE*, 8(6): 170–173, 2001.

I. W. Selesnick. The design of approximate Hilbert transform pairs of wavelet bases. *Signal Processing, IEEE Transactions on*, 50(5):1144–1152, 2002.

K. Shimotsu. Gaussian semiparametric estimation of multivariate fractionally integrated processes. *Journal of Econometrics*, 137(2):277–310, 2007.

J.-P. Thiran. Recursive digital filters with maximally flat group delay. *Circuit Theory, IEEE Transactions on*, 18(6):659–664, 1971.

B. Whitcher and M. J. Jensen. Wavelet estimation of a local long memory parameter. *Exploration Geophysics*, 31(1/2):94–103, 2000.

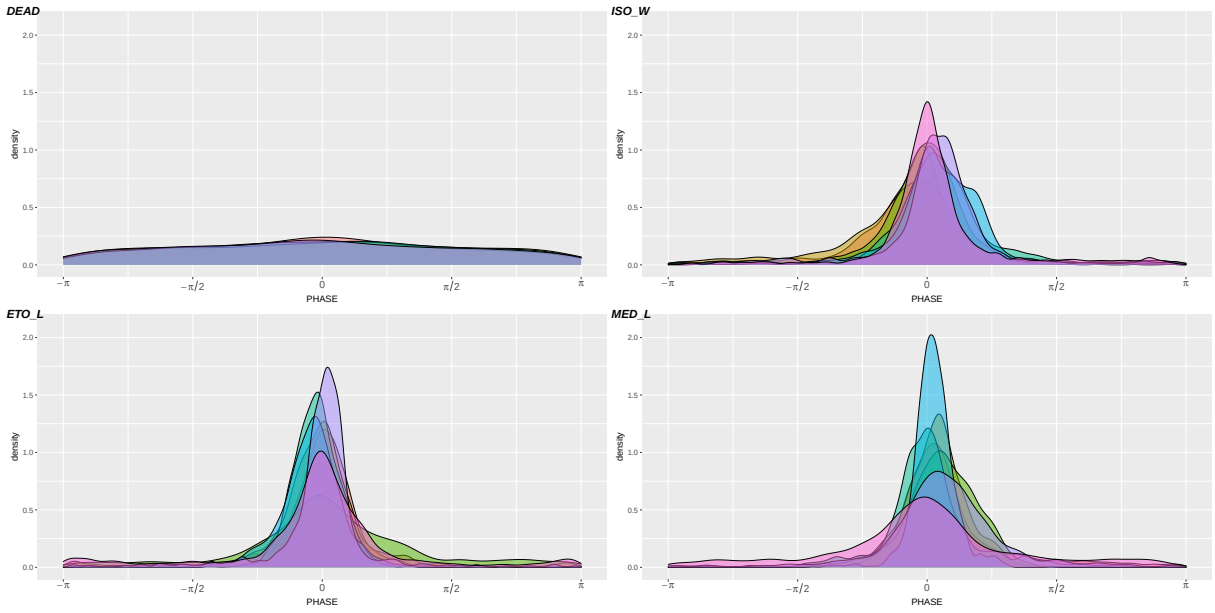


Figure 8: Plot of the empirical distribution of the phases $\widehat{\phi}$ obtained for the 4 groups of rats without thresholding the correlations. Each color corresponds to a rat.

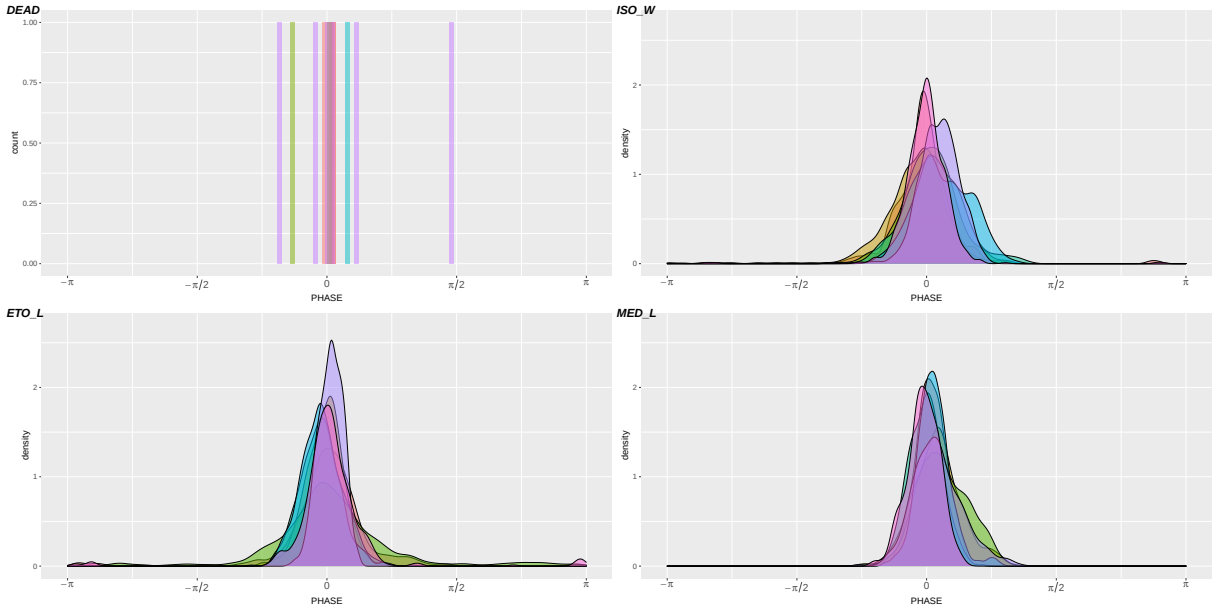


Figure 9: Plot of the empirical distribution of the phases obtained for the 4 groups of rats after first thresholding the correlations. Only phases associated to correlations with a magnitude higher than 0.3 are considered. Each color corresponds to a rat. For dead rats, a bar plot is provided rather than a density plot due to the low number of values.

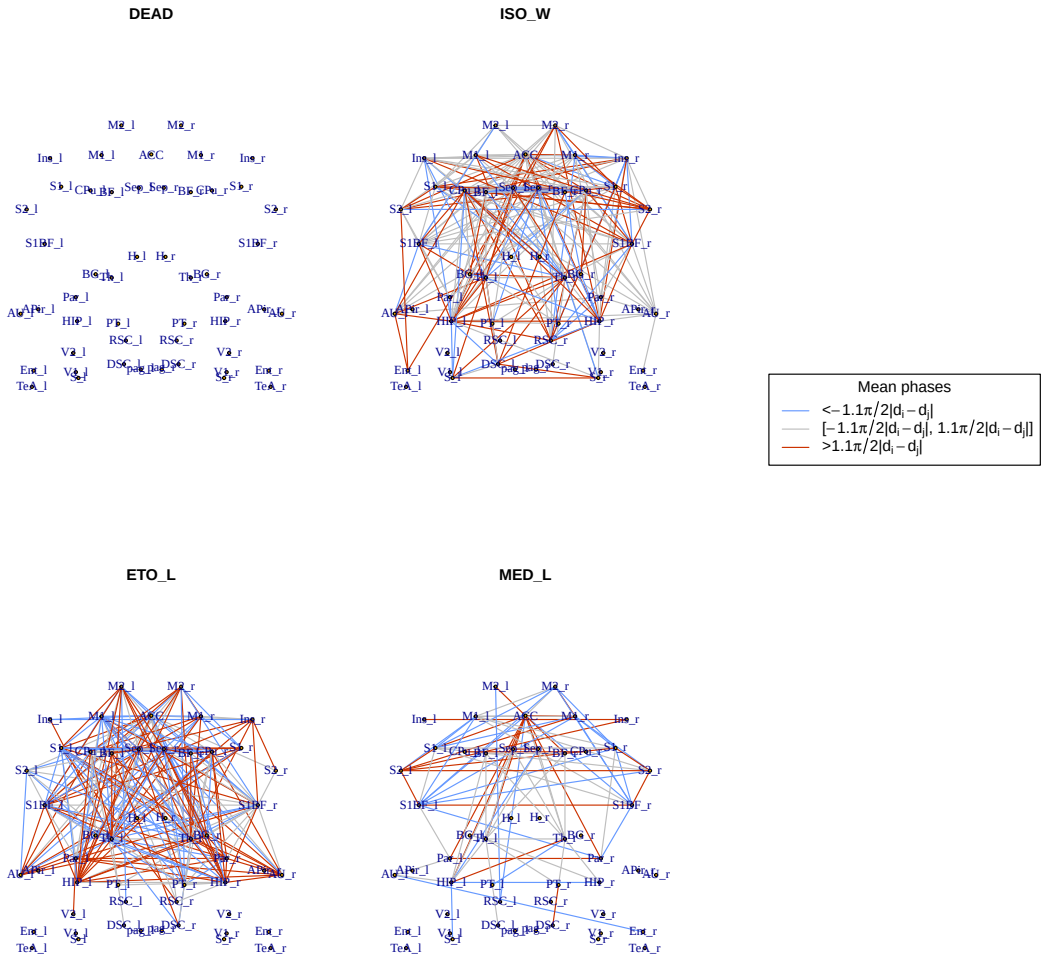


Figure 10: Plot of the average graphs with correlations and phases obtained for 4 groups of rats: DEAD, ISO_W, ETO_L and MED_L. Only edges corresponding to an average correlation's magnitude higher than 0.3 are displayed. Red edges correspond to positive average phases higher than $1.1|\phi^*|$, blue edges correspond to negative average phases lower than -$1.1|\phi^*|$, and grey edges to average phases between -$1.1|\phi^*|$ and $1.1|\phi^*|$. The quantities $\phi^*$ are equal to $\phi^*_{\ell,m} = -\frac{\pi}{2}(d_\ell - d_m)$.