



**HAL**  
open science

# Local Whittle estimation with (quasi-)analytic wavelets

Sophie Achard, Irène Gannaz

► **To cite this version:**

Sophie Achard, Irène Gannaz. Local Whittle estimation with (quasi-)analytic wavelets. *Journal of Time Series Analysis*, 2024, 45 (3), pp.421-443. 10.1111/jtsa.12719 . hal-03272326v3

**HAL Id: hal-03272326**

**<https://hal.science/hal-03272326v3>**

Submitted on 4 Aug 2023

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

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



Distributed under a Creative Commons Attribution 4.0 International License

# Local Whittle estimation with (quasi-)analytic wavelets

Sophie Achard

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

Irène Gannaz

Univ. Grenoble Alpes, CNRS, Grenoble INP\*, G-SCOP, 38000 Grenoble, France

August 2023

## 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. It is of interest to estimate the long-memory, long-run covariance and general phase parameters of time series generated by this wide class of models although they are not necessarily Gaussian nor stationary. This estimation is thus not directly 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 local 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 neuroscience dataset is presented, 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. Many examples can be cited such as hydrology [Whitcher and Jensen, 2000], finance

---

\*Institute of Engineering Univ. Grenoble Alpes

[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. Some approaches proposed a simple definition, where the covariance matrix is real [Lobato, 1999, Shimotsu, 2007]. However, this simple model is not able to address any multivariate models. For example, in [Lobato, 1997] two models were introduced, FIVARMA and VARFIMA, from this approach. Long-memory models with a complex covariance matrix give a solution to overcome this problem [Kechagias and Pipiras, 2020, Baek et al., 2020]. Following these modelings, the long-memory model studied in this paper admits a complex long-run covariance matrix, where a phase-term is added to the covariance structure.

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

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

where  $\circ$  denotes the Hadamard product, and  $d_m^* = d_m - D_m \in (-0.5, 0.5)$  for all  $m$ . The process  $(1 - \mathbb{L})^{D_m} X_m$  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 parameters  $\mathbf{d}$  throughout the paper.

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

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

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

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

(M-2) There exists  $C_f > 0$  and  $\beta > 0$  such that  $\sup_{0 < \lambda < \pi} \sup_{\ell, m=1, \dots, N} \frac{|f_{\ell, m}^S(\lambda) - 1|}{\lambda^\beta} \leq C_f$ .

In particular, our definition agrees with the one given in [Kechagias and Pipiras \[2014\]](#) if  $D_m = 0$  for all  $m = 1, \dots, p$ . Definition (M-1) includes both stationary and non-stationary time series. It has the advantage of including multivariate fractional Brownian motion [[Coeurjolly et al., 2013](#)].

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

$$\Theta_{\ell, m} = \Omega_{\ell, m} e^{i\phi_{\ell, m}}, \quad \ell, m = 1, \dots, p,$$

with  $\Omega = (\Omega_{\ell, m})_{\ell, m=1, \dots, p}$  a real symmetric non-negative semi-definite matrix and  $\Phi = (\phi_{\ell, m})_{\ell, m=1, \dots, p}$  an anti-symmetric matrix. Let the bar above denote the conjugate operator. The matrix  $\Theta$  satisfies  $\Theta^T = \overline{\Theta}$  since  $\mathbf{f}^T(\cdot) = \overline{\mathbf{f}(\cdot)}$ . We will use  $\|\Omega\|$  to denote the infinity norm, that is,  $\|\Omega\| = \max_{\ell, m=1, \dots, p} \Omega_{\ell, m}$ . 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 \[1999\]](#), [Shimotsu \[2007\]](#) and [Nielsen and Frederiksen \[2011\]](#), 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. Our model includes among other the co-integration case [Robinson, 2008, Nielsen, 2011, Baek et al., 2020]. Observe that the Fourier-based local Whittle procedure proposed in Baek et al. [2020] is very closed to the one developed here.

Introducing wavelets is motivated by their flexibility for real data applications. In particular, it allows to consider non-stationary processes thanks to an implicit differentiation. The introduction of a general phase term challenges the choice of the wavelet filters. Due to condition (M-1), we need to consider complex filters for identifying the imaginary part of  $\Theta$ . Indeed, as illustrated in Gannaz et al. [2017], with real wavelet filters it is not possible to recover both the real and the imaginary part of the matrix  $\Theta$ . Complex wavelet filters, with quasi-analytic properties, are described in Section 2. The main properties of the filters are displayed and an approximation of the covariance of the wavelet coefficients is derived in Section 3. Section 4 recalls the definition of the wavelet local Whittle estimators. We prove 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  $\mathbf{X} = \{\mathbf{X}(t), t \in \mathbb{Z}\}$ .

Let  $(h^{(L)}(\cdot), h^{(H)}(\cdot))$  and  $(g^{(L)}(\cdot), g^{(H)}(\cdot))$  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  $(h^{(L)}(\cdot), h^{(H)}(\cdot))$ . They can be defined through their Fourier transforms as

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

Let us define similarly  $(\varphi_g(\cdot), \psi_g(\cdot))$  the father and the mother wavelets associated with the wavelet

filters  $g^{(L)}(\cdot)$  and  $g^{(H)}(\cdot)$ . Their Fourier transforms are

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

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

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

Wavelet  $\psi(\cdot)$  is said to be analytic if its Fourier transform is only supported on the positive frequency semi-axis. In particular, it is sufficient to show that the pair  $(\psi_g(\cdot), \psi_h(\cdot))$  is a Hilbert pair, that is,  $\widehat{\psi}_g(\lambda) = -i \operatorname{sign}(\lambda) \widehat{\psi}_h(\lambda)$ , for all  $\lambda \in \mathbb{R}$ , where  $\operatorname{sign}(\lambda)$  denotes the sign function taking values  $-1, 0$  and  $1$  for  $\lambda < 0$ ,  $\lambda = 0$  and  $\lambda > 0$ , respectively.

From Paley-Wiener Theorem, analytic filters with finite support do not exist. Selesnick's common factor filters propose compact wavelet filters with a relaxation of the strict analytic condition.

## 2.1 Selesnick's common factor filters

We choose here to consider the quasi-analytic filters introduced by [Thiran \[1971\]](#), [Selesnick \[2002\]](#). The *common-factor* wavelets, defined by [Selesnick \[2002\]](#), have a compact support and are quasi-analytic. They are parameterized by a degree  $L$  quantifying the approximation of 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 of their properties.

Let  $\widehat{d}_L(\lambda)$ ,  $\lambda \in \mathbb{R}$ , be defined by

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

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

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

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

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 Achard et al. [2020].

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

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

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  $\widehat{q}_{L,M}(\cdot)$  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. Since perfect reconstruction is not necessary for deriving estimation procedures, we can assume that  $\widehat{q}_{L,M}(\cdot)$  is a constant equal to 1.

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

The two main characteristics are the compact support and the quasi-analiticity. The compact support property for CFW-C( $M, L$ ) filters is given below.

**Proposition 2.** *Let  $M, L$  be strictly positive integers. The functions  $\varphi(\cdot)$ , and  $\psi(\cdot)$  for CFW-C( $M, L$ ) have supports of respective length  $M + 2L + 1$  and  $M + L + 1/2$ .*

The proof is given in Appendix.

Concerning the functions  $\varphi(\cdot)$ , and  $\psi(\cdot)$  for CFW-PR( $M, L$ ), in practice, they have supports of respective length  $2M + 3L$  and  $2M + 2L + 1/2$ . Yet, there is no theoretical proof that these lengths are indeed achieved for all  $(L, M)$ . See Section 4 of Achard et al. [2020].

Let us now recall the main result concerning the analytic approximation established in Achard et al. [2020].

**Theorem 3** (Achard et al. [2020]). For all  $\lambda \in \mathbb{R}$ , for all  $\widehat{q}_{L,M}(\cdot)$  real polynomial of  $(e^{-i\lambda})$ ,

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

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

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

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

where  $U_L(\cdot)$  is a  $\mathbb{R} \rightarrow [0, 2]$  function satisfying, for all  $\lambda \in \mathbb{R}$ ,

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

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  $\alpha_L(\cdot)$  is well defined on  $\mathbb{R}$ .

Theorem 3 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. However, the higher  $L$ , the larger the wavelets support.

### 3 Moments approximations of the wavelet coefficients

Let  $\{\mathbf{W}_{j,k}, j \geq 0, k \in \mathbb{Z}\}$  denote the wavelet coefficients of the process  $\mathbf{X}$  associated to the wavelet pair  $(\varphi(\cdot), \psi(\cdot))$ . At a given resolution  $j \geq 0$ , 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

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



where  $\widetilde{\mathbf{X}}(t) = \sum_{k \in \mathbb{Z}} \mathbf{X}(k) \varphi(t - k)$ . Given any  $j \geq 0$  and any  $k \in \mathbb{Z}$ ,  $\mathbf{W}_{j,k}$  is a  $p$ -dimensional vector  $\mathbf{W}_{j,k} = \left( W_{j,k}(1) \ W_{j,k}(2) \ \dots \ W_{j,k}(p) \right)^T$  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  $\text{Cov}(\mathbf{W}_{j,k})$ , defined as follows

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

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

In practice, a finite number of observations of the process  $\mathbf{X}$  are available,  $\mathbf{X}(1), \mathbf{X}(2), \dots, \mathbf{X}(N)$ . As the wavelets have a compact support, only a finite number of coefficients are non-zero at each scale  $j$ . More precisely, for every  $j \geq 0$ , let  $n_j$  denote the number of coefficients  $\mathbf{W}_{j,k}$  evaluated from the observations. Note that only the coefficients evaluated without boundary effects are taken into account (see the definition of  $n_j$  in Lemma 4). For all  $k < 0$  or  $k > n_j$ , the coefficients  $\mathbf{W}_{j,k}$  are set to zero. In the following, we will assume that  $M$  is fixed and finite, and that  $L$  may go to infinity. Hence, the length of the wavelets support is equivalent to  $L$  when  $N$  goes to infinity. If additionally  $2^j N^{-1} L \rightarrow 0$ , then  $n_j$  is equivalent to  $2^{-j} N$ . In that case, the behavior of  $n_j$  is similar to the framework of [Moulines et al. \[2008\]](#) and [Achard and Gannaz \[2016\]](#).

**Lemma 4.** *Let  $(\varphi(\cdot), \psi(\cdot))$  be a CFW-C( $M, L$ ) wavelet pair, with  $M, L \geq 1$ . Let  $\{\widehat{\mathbf{W}}_{j,k}, j \geq 0, k \in \mathbb{Z}\}$  denote the wavelet coefficients evaluated from  $\mathbf{X}(1), \mathbf{X}(2), \dots, \mathbf{X}(N)$  by  $\widehat{\mathbf{W}}_{j,k} = \int_{\mathbb{R}} \widehat{\mathbf{X}}(t) \psi_{j,k}(t) dt$ ,  $j \geq 0, k \in \mathbb{Z}$ , where  $\widehat{\mathbf{X}}(t) = \sum_{k=1}^N \mathbf{X}(k) \varphi(t - k)$ . Then, at each scale  $j$ , the number of coefficients  $n_j$  such that  $\widehat{\mathbf{W}}_{j,k} = \mathbf{W}_{j,k}$  is*

$$n_j = \max\{0, \lfloor 2^{-j}(N - 2L - M - 1) - L - M - 1/2 \rfloor\}.$$

*Suppose that  $N^{-1}L \rightarrow 0$  when  $N$  goes to infinity. Then, for all  $j$  such that  $2^j N^{-1}L \rightarrow 0$  when  $N$  goes to infinity,  $n_j 2^j N^{-1} \rightarrow 1$  when  $N$  goes to infinity.*

### 3.1 Motivations

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

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

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  $\Theta_{\ell,m} = \Omega_{\ell,m} e^{i\phi_{\ell,m}}$ ,  $\phi_{\ell,m} = \pi(d_1 - d_2)/2$ . Let  $\mathbf{d}$  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 to infer  $\Theta_{\ell,m}$ .

We simulate  $\mathbf{X}(1), \dots, \mathbf{X}(2^J)$ , with  $J = 12$ . For each scale  $j \geq 0$ , we evaluate the sample wavelet covariances as  $\widehat{\Sigma}(j) = \frac{1}{n_j} \sum_{k=0}^{n_j-1} W_{j,k}(1) \overline{W_{j,k}(2)} - \left(\frac{1}{n_j} \sum_{k=0}^{n_j-1} W_{j,k}(1)\right) \left(\frac{1}{n_j} \sum_{k=0}^{n_j-1} \overline{W_{j,k}(2)}\right)$ , and the wavelet sample correlations as  $\widehat{\Sigma}_{1,2}(j) / \sqrt{\widehat{\Sigma}_{1,1}(j) \widehat{\Sigma}_{2,2}(j)}$ .

Figure 1 shows the behavior of sample wavelet correlations with respect to scale  $j$  over 100 realizations of  $(\mathbf{X}(1), \dots, \mathbf{X}(2^J))$ . First observe that for real wavelets (left column), the wavelet sample covariance  $\widehat{\Sigma}_{1,2}(j)$  tends to 0 when the scale  $j$  increases. This confirms the impossibility to identify  $\Theta_{1,2}$ . In addition, the plots displayed in the middle and right columns illustrate 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  $\Omega_{1,2} e^{i\phi_{1,2}} / \sqrt{\Omega_{1,1} \Omega_{2,2}}$  as the frequency decreases.

### 3.2 Theoretical results

We will now develop the theory of the behavior of  $\text{Cov}(\mathbf{W}_{j,k})$ . 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 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.

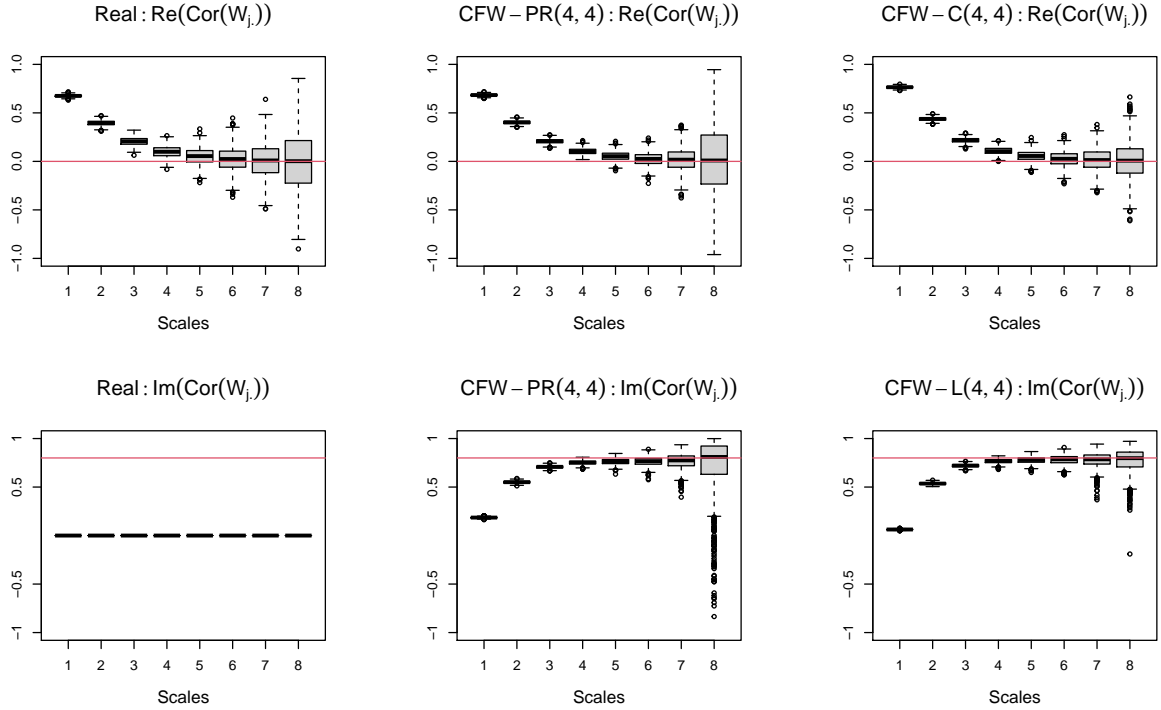


Figure 1: Boxplots of sample correlations between wavelet coefficients at different scales for the bivariate process defined in (8). First row gives the real part of the correlations and second row gives the imaginary part. Each column corresponds to different wavelet filters, from left to right: Daubechies' 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.

$$(C-a) \quad -M + \beta + 1 < 2d_\ell < M \quad \text{for all } \ell = 1, \dots, p, M \geq 2.$$

Parameter  $M$  is the number of vanishing moments and it also corresponds to the regularity of CFW-C( $M,L$ ) filters. Parameters  $(d_\ell)_{\ell=1,\dots,p}$  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 5.** *Let  $\mathbf{X}$  be a  $p$ -multivariate long range dependent process with long memory pa-*

rameters  $d_1, \dots, d_p$  with generalized spectral density  $\mathbf{f}(\cdot)$  satisfying (M-1) with short-range behavior (M-2). Consider  $\{W_{j,k}(\ell), (j, k) \in \mathbb{Z}, \ell = 1, \dots, p\}$  the wavelet coefficients obtained with CFW-C(M,L) filters. Suppose that (C-a) hold. Then we have, for all  $j \geq 0, k \in \mathbb{Z}$ ,

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

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

The proof is given in Appendix.

The result follows from the fact that CFW-C(M,L) 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 5 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

$$\left| 2^{-j(d_\ell+d_m)} \text{Cov}(W_{j,k}(\ell), W_{j,k}(m)) - 2 \Omega_{\ell,m} \cos(\phi_{\ell,m}) \int_0^\infty |\lambda|^{-\delta} |\widehat{\psi}(\lambda)|^2 d\lambda \right| \leq C \|\Omega\| 2^{-j\beta}. \quad (9)$$

It is straightforward to check 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 3, leads to the following result.

**Proposition 6.** *Let  $\mathbf{X}$  be a  $p$ -multivariate long range dependent process with long memory parameters  $d_1, \dots, d_p$  with generalized spectral density  $\mathbf{f}(\cdot)$  satisfying (M-1)–(M-2).*

*Consider  $\{W_{j,k}(\ell), (j, k) \in \mathbb{Z}, \ell = 1, \dots, p\}$  the wavelet coefficients obtained with CFW-C(M,L)*

filters. Suppose that (C-a) hold and that  $L$  goes to infinity, with  $L2^{-2j} \rightarrow 0$  when  $j$  goes to infinity.

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

$$\left| 2^{-j(d_\ell+d_m)} \text{Cov}(W_{j,k}(\ell), W_{j,k}(m)) - 4\Theta_{\ell,m} \int_0^\infty |\lambda|^{-d_\ell-d_m} \left| \widehat{\psi}_h(\lambda) \right|^2 d\lambda \right| \leq C_2 (2^{-j\beta} + L2^{-2j} + L^{-M-1}), \quad (10)$$

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

The proof is given in Appendix.

Convergence (10) can be written as follows: when  $2^{-j\beta} + L2^{-2j} + L^{-1} \rightarrow 0$ , for all  $(\ell, m) \in \{1, \dots, p\}^2$ ,

$$\lim_{j \rightarrow \infty} 2^{-j(d_\ell+d_m)} \text{Cov}(W_{j,k}(\ell), W_{j,k}(m)) = G_{\ell,m}, \quad (11)$$

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

Common-factor wavelets, as stated by Proposition 6, have the ability of recovering simultaneously the magnitude and the phase. Observe that with real wavelets the upper bound in (9) is  $2^{-j\beta}$ , up to a multiplicative constant. With complex wavelets, the rate depends on  $L$ , and the parameter will need to be calibrated accordingly.

The specificity of CFW-C(M,L) filters is that the quality of the analytic approximation is based only on parameter  $L$ , as written in Proposition 6. Nevertheless, if we want to have an approximation with the same quality as that obtained with real wavelets, the choice of  $L$  is more constrained. This trade-off 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.*  $\geq 8$ ) is not manageable. As shown by the simulations in Section 5, however, the results are of good quality even with a smaller value of  $L$ .

### 3.3 Quality of approximation

To empirically assess the accuracy of the approximation, let us compare the empirical covariances of the example of Section 3.1 to the approximation of Proposition 5. Figure 2 displays the sample covariance of the wavelet coefficients, respectively with real Daubechies filters with  $M = 4$ , CFW-PR(4,4) and CFW-C(4,4) filters. As for Figure 1,  $N = 2^{12}$  observations were considered. Observe that the covariance term is complex, and only the magnitude is represented in Figure 2.

Figure 2 shows the difference between our theoretical findings given in Proposition 5 and the simulations for both CFW-PR(M,L) and CFW-C(M,L). To better evaluate the quality of the approximation with CFW-C(M,L) filters, the same figure without the first scale is provided in Figure 3. It shows that indeed the approximation improves when the scale  $j$  increases. Nevertheless, the difference between the results obtained with the simulations at first scales (corresponding to the highest frequencies) and the approximation given in Proposition 5 is higher with CFW-C(4,4) filters in comparison with Daubechies and CFW-PR(4,4) filters. Therefore, the lowest scale used in estimation may be higher with CFW-C(M,L) filters. This choice may reduce the bias but increase the variance.

## 4 Estimation

Let  $j_0$  and  $j_1$ ,  $j_1 \geq j_0 \geq 1$  be respectively the lower and the upper resolution levels used in the estimation procedure. The estimation is based on the vectors of wavelet coefficients  $\{\mathbf{W}_{j,k}, j_0 \leq j \leq j_1, k \in \mathbb{Z}\}$ . The total number of non-zero coefficients used for estimation is then  $n = \sum_{j=j_0}^{j_1} n_j$ . Without restriction of generality, we can assume that  $L = o(N)$ .

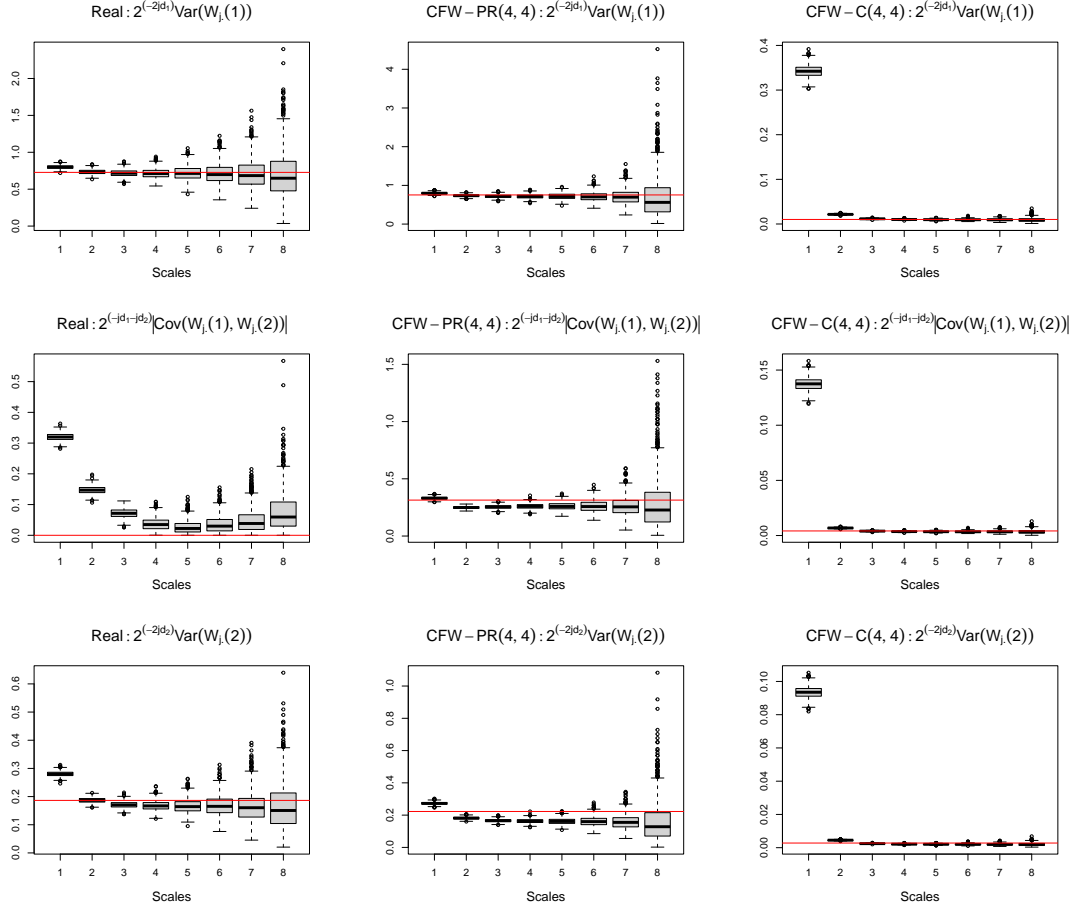


Figure 2: Boxplots of normalized sample covariances between wavelet coefficients at different scales for the bivariate process defined in (8). Let  $\mathbf{V}_j = \begin{pmatrix} 2^{-2j d_1} \text{Var}\{W_{j,k}(1), k \in \mathbb{Z}\} & 2^{-j(d_1+d_2)} \text{Cov}\{(W_{j,k}(1), W_{j,k}(2)), k \in \mathbb{Z}\} \\ 2^{-j(d_1+d_2)} \text{Cov}\{(W_{j,k}(1), W_{j,k}(2)), k \in \mathbb{Z}\} & 2^{-2j d_2} \text{Var}\{W_{j,k}(2), k \in \mathbb{Z}\} \end{pmatrix}$ . The first row gives the sample version of  $V_{j11}$ , the second row gives the sample version of  $|V_{j12}|$  and the third row gives the sample version of  $V_{j22}$ . 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 approximation given by Proposition 5.

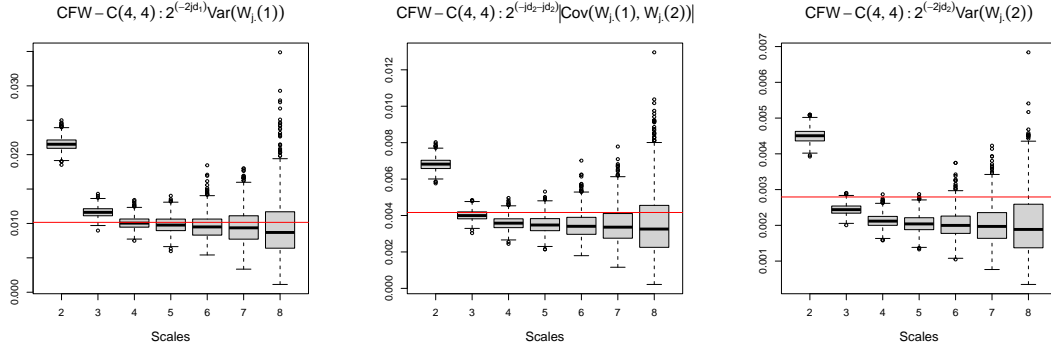


Figure 3: Boxplots of normalized sample covariances between CFW-C(4,4) coefficients at different scales for the bivariate process defined in (8). Plots are the same as the right column of Figure 2 but without the first wavelet scale. From left to right, panels correspond respectively to the variance 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 5.

#### 4.1 Estimation procedure

Based on approximation (11), the objective function  $\mathcal{L}(\cdot)$  is defined by the wavelet Whittle approximation of the negative log-likelihood (see Achard and Gannaz [2016])

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

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  $G_{\ell,m} = \Theta_{\ell,m} K(d_\ell + d_m)$ ,  $1 \leq \ell, m \leq p$ . We can rewrite  $\mathcal{L}(\cdot)$  as

$$\mathcal{L}(\mathbf{G}, \mathbf{d}) = \frac{1}{n} \sum_{j=j_0}^{j_1} \left[ n_j \log \det (\mathbf{\Gamma}_j(\mathbf{d}) \mathbf{G} \mathbf{\Gamma}_j(\mathbf{d})) + \text{trace} \left( (\mathbf{\Gamma}_j(\mathbf{d}) \mathbf{G} \mathbf{\Gamma}_j(\mathbf{d}))^{-1} \mathbf{I}(j) \right) \right], \quad (12)$$

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 a positive definite Hermitian matrix, for all  $j \geq 0$  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  $\text{trace} \left( (\mathbf{\Gamma}_j(\mathbf{d}) \mathbf{G} \mathbf{\Gamma}_j(\mathbf{d}))^{-1} \mathbf{I}(j) \right)$  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 (12) with respect to the matrix  $\mathbf{G}$  yields

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

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 whereas in [Baek et al. \[2020\]](#) the authors consider a general form of phases. In both [Shimotsu \[2007\]](#) and [Baek et al. \[2020\]](#), with a Fourier-based approach, a real matrix  $\mathbf{G}(\mathbf{d})$  and complex valued matrices  $\mathbf{\Gamma}_j(\mathbf{d})$ , including the phases  $(\phi_{\ell,m})_{\ell,m=1,\dots,p}$ , are considered.  $\mathbf{G}(\mathbf{d})$  and  $\mathbf{\Gamma}_j(\mathbf{d})$  are estimated in a second step, together with parameter  $\mathbf{d}$ . They minimize the objective function obtained when replacing  $\mathbf{G}$  by  $\widehat{\mathbf{G}}(\mathbf{d})$  in (12). However, our procedure makes it possible 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

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

Since  $\mathbf{\Gamma}_j(\mathbf{d}) = \text{Diag}(2^{-j\mathbf{d}})$ , 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  $\widehat{\mathbf{d}} = \text{argmin}_{\mathbf{d}} R(\mathbf{d})$ .

In a second step of estimation we define  $\widehat{\mathbf{G}}(\widehat{\mathbf{d}})$ , estimator of  $\mathbf{G}$ . And we recover an estimation of  $\Theta$  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  $\{\mathbf{I}(j), j \geq 0\}$ . Examples of linear processes satisfying this condition can be found in Proposition 4 of [Achard and Gannaz \[2016\]](#) for real wavelets. With complex wavelets, to obtain the convergence results, we need to define the parameter  $L$  depending on  $N$ . We omit the dependence in the notation except when it is necessary. Therefore, the wavelet bases depend on  $N$  as the parameter  $L$  depends on  $N$ . Hence the wavelet scalogram  $\{\mathbf{I}(j), j \geq 0\}$  depends on  $N$ , via the number of observations used in the calculation of the coefficients and via  $L$ .

### Condition (C)

$L$  is a sequence of  $N$ ,  $L = L(N)$ , such that,

$$\text{for all } \ell, m = 1, \dots, p, \quad \sup_N \sup_{j \geq 0} \frac{|\text{Var}(I_{\ell, m}(j))|}{n_j 2^{2j(d_\ell + d_m)}} < \infty.$$

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

**Theorem 7.** *Suppose that (C-a) and assumptions of Proposition 6 hold. Assume that Condition (C) is satisfied. Denote  $j_N = \max\{j, n_j \geq 1\}$ .*

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

*Consider CFW-C( $M, L$ ) filters with  $M \geq 2$  and  $2^{-2j_0}L + N^{-1}2^{j_1}L + \log(N)^3 L^{-M-1} \rightarrow 0$ .*

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

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

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

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

Elements of proof are given in Appendix.

The convergence rate  $\widehat{\mathbf{d}} - \mathbf{d}^0 = O_{\mathbb{P}}(N^{-\beta/(1+2\beta)})$  is optimal in minimax sense [Giraitis et al., 1997].

The condition on the scales used in the estimation is  $\log(N)^2(2^{-j_0\beta} + N^{-1/2}2^{j_0/2}) \rightarrow 0$ . As explained in Achard and Gannaz [2019], this indicates that highest frequencies should be removed. The number of scales to remove depends on the short-range dependence via  $\beta$ . 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-a) and  $2^{-2j_0}L + N^{-1}2^{j_1}L + \log(N)^3 L^{-M-1} \rightarrow 0$ . Condition (C-a) only depends on  $M$ . It is very similar to the one given in Achard and Gannaz [2016] with real filters. It imposes that the number of vanishing moments  $M$  is high enough.

The parameter  $L$  quantifies the quality of the analytic approximation of CFW-C(M,L) filters. The assumption  $2^{j_1}N^{-1}L \rightarrow 0$  results from Lemma 4. This is a technical assumption allowing  $n_j$  to be equivalent to  $2^{-j}N$  as  $N$  goes to infinity. This facilitates the translation of the proofs from real wavelets to complex wavelets. This assumption deals with the highest scale  $j_1$ . It can be formulated on  $j_0$ ,  $N^{-1}2^{j_0}L \rightarrow 0$ , when  $j_1 = j_0 + \Delta$ , with  $\Delta < \infty$ . The condition  $\log(N)^3 L^{-M-1} \rightarrow 0$  guarantees that  $L$  is high enough for the analytic approximation to be satisfactory. Alternatively,  $L$  should not be too high, and condition  $2^{-2j_0}L \rightarrow 0$  ensures that the size of the wavelets support 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 other scales.

*Remark.* Baek et al. [2020] observe that in case of co-integration, the corresponding magnitude  $\Omega_{a,b}$  is equal to 0. Hence, the phase parameter  $\varphi_{a,b}$  is not identifiable. To counter this problem, Baek et al. [2020] propose to use another parametrization, where  $\Theta_{a,b}$  is decomposed into its real and its imaginary parts. Our procedure estimates the complex matrix  $\Theta$ , which is always identifiable, so

this discussion is unnecessary here.

### 4.3 Asymptotic normality

A useful result in estimation is asymptotic normality. For real wavelet-based local Whittle estimation, in a multivariate context, it has been studied by [Gannaz \[2023\]](#). 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  $\sum_{u \in \mathbb{Z}} \max_{a,b=1,\dots,p} |A_{a,b}^{(\mathbf{D})}(u)|^2 < \infty$  and

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

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

The asymptotic normality of the estimator of the long-memory parameters is established by our following theorem. For  $\mathbf{M} \in \mathbb{R}^{p \times p}$ ,  $\text{vec}(\mathbf{M})$  denotes the operation that transforms a matrix  $\mathbf{M} \in \mathbb{R}^{p \times p}$  into a vector of  $\mathbb{R}^{p^2}$ .

**Theorem 8.** *Suppose that conditions of Theorem 7 are satisfied and that assumption (M-3) hold. Let  $j_0 < j_1 \leq j_N$  with  $j_N = \max\{j, n_j \geq 1\}$  such that*

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

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

Consider CFW-C( $M, L$ ) filters with  $M \geq 2$  and

$$N^{-1} 2^{j_1} L + \log(N)^3 N^{1/2} 2^{-j_0/2} (L 2^{-2j_0} + L^{-M-1}) \rightarrow 0.$$

Then,

- $\sqrt{n}(\widehat{\mathbf{d}} - \mathbf{d}^0)$  converges in distribution to a centered Gaussian distribution with a variance  $\mathbf{V}^{(\mathbf{d})}(\Delta)$  defined in Appendix, equation (40).
- $\text{vec}\left(\sqrt{n}\left(\widehat{\mathbf{G}}(\widehat{\mathbf{d}}) - \mathbf{G}^0\right)\right)$  converges in distribution to a centered Gaussian distribution with a variance  $\mathbf{V}^{(\mathbf{G})}(\Delta)$  defined in Appendix, equation (41).

The proof is very similar to the one of Gannaz [2023]. Some points are detailed in Appendix.

The highest scale is  $j_1 = j_0 + \Delta$ . The theorem distinguishes the cases  $\Delta < \infty$  and  $\Delta = \infty$ . Note that when  $\Delta < \infty$ , the condition  $N^{-1/2}2^{j_1/2}L \rightarrow 0$  is equivalent to  $N^{-1/2}2^{j_0/2}L \rightarrow 0$ . Hence, the condition on  $L$  writes as  $\log(N)^3 N^{1/2}2^{-j_0/2}(L2^{-2j_0} + L^{-M-1}) \rightarrow 0$ .

*Remark.* Observe that with the condition on  $j_0$ , the minimax rate is not achieved. But we can take  $2^{j_0} = N^{a_0}$  with  $1/(1+2\beta) < a_0 < 1$ . If  $L$  is defined as  $L = N^{b_0}$ , then it suffices that  $\frac{\beta}{1+2\beta} \cdot \frac{1}{M+1} < b_0 < \frac{1}{2} \min\{1 - a_0, 5a_0 - 1\}$ . For example, we can take  $2^{j_0} = N^{(1+\beta)/(1+2\beta)}$  and  $L = N^{c_0 \cdot \beta/(1+2\beta)}$  with  $\frac{1}{2(M+1)} < c_0 < \frac{1}{2}$ .

As detailed in Gannaz [2023], these results allow to build hypothesis tests on the long-memory parameters and on the long-run covariance.

## 5 Simulation study

In this section, we verify the accuracy of the covariance approximation given in Proposition 6 and the consistency of the parameters estimates provided in Proposition 7 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-C(4,4) filters.

We compare the quality of estimation of parameters  $\mathbf{d}$  to the one given by real wavelets, namely Daubechies' wavelets with 4 vanishing moments. A comparison with a Fourier-based local Whittle procedure is also provided, when the simulated processes are stationary.

Observe that comparisons of Fourier-based and wavelet-based estimation procedures in univariate settings have been done previously in Nielsen and Frederiksen [2005] and Faÿ et al. [2008]. In multivariate settings, comparisons have been done in Achard and Gannaz [2019] with a parametric phase. With a general phase, Baek et al. [2020] proposed Monte Carlo simulations using Fourier-based approach. The specificity here is to extend simulations of the former to non-stationary processes.

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  $\mathbb{E}[\xi(t)\xi(t)^T \mid \mathcal{F}_{t-1}] = \mathbf{\Sigma}$  with  $\mathbf{\Sigma}$  positive definite, where  $\mathcal{F}_{t-1}$  is the  $\sigma$ -field generated by  $\{\xi(s), s < t\}$ . The spectral density of  $\xi$ , denoted  $\mathbf{f}_\xi(\cdot)$ , satisfies  $\mathbf{f}_\xi(\lambda) = \mathbf{\Sigma}$ , for all  $\lambda \in \mathbb{R}$ .

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  $\mathbf{X} = \{\mathbf{X}(t), t \in \mathbb{Z}\}$  by

$$\mathbf{A}(\mathbb{L}) \text{Diag} \left( (1 - \mathbb{L})^{\mathbf{d}} \right) \mathbf{X}(t) = \mathbf{B}(\mathbb{L})\xi(t), t \in \mathbb{Z}. \quad (13)$$

The spectral density of  $\mathbf{X}$  satisfies

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

In particular

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

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

This corresponds to Model A of Lobato [1997]. Note that this model satisfies the definition of LRD processes of Kechagias and Pipiras [2014]. In Kechagias and Pipiras [2014], the process  $\xi$  is not necessarily a white-noise process, and verifies  $\mathbf{f}_\xi(\lambda) \sim_{\lambda \rightarrow 0^+} \boldsymbol{\Sigma}$ , which is indeed true for a white-noise process.

Following (13), we have simulated  $\mathbf{X}(1), \dots, \mathbf{X}(N)$  in (13) with  $N = 2^{12}$ , null  $\mathbf{A}_k$  and  $\mathbf{B}_k$  for  $k \geq 0$ . 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  $\boldsymbol{\Sigma}$  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 6 is valid, especially for high scales (lowest frequencies), even if it has not been established theoretically for such filters. The figure shows that the approximation of Proposition 6 is slightly more accurate for the real part of wavelet correlations than for the imaginary part.

The 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 to choose 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

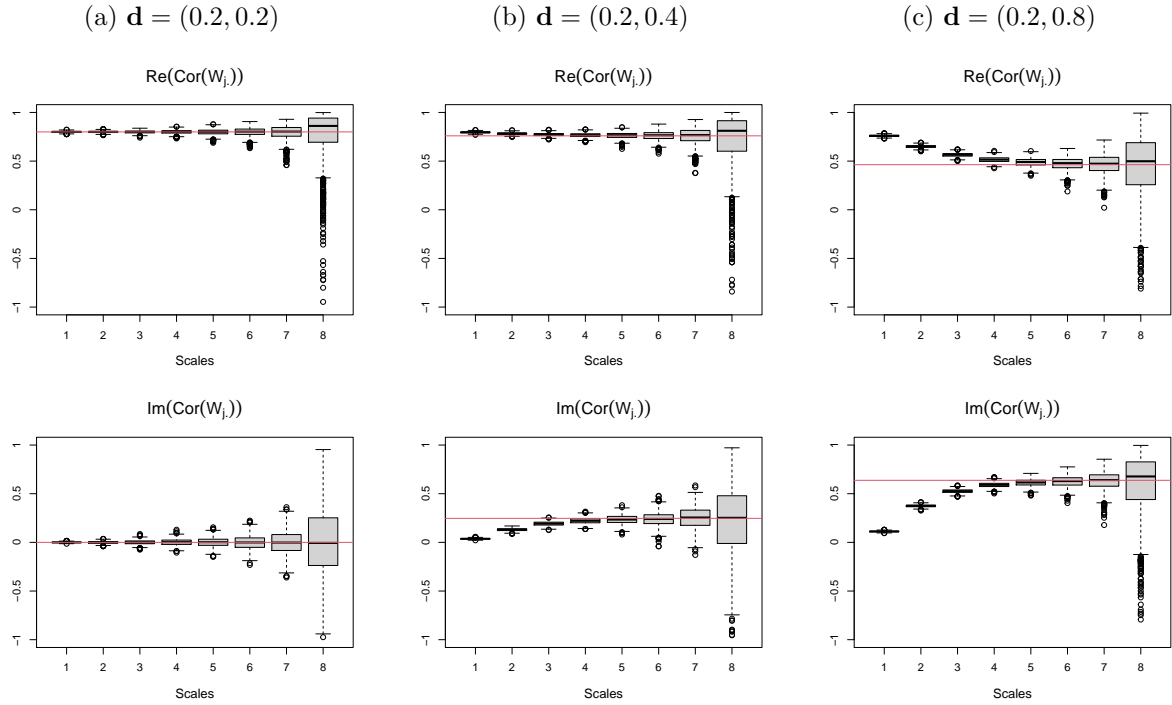


Figure 4: Boxplots of correlation between CFW-PR(4,4) coefficients at different scales for ARFIMA processes. First row gives the real part of the correlations and second row gives the imaginary part. Each column corresponds to a given value of parameter  $\mathbf{d}$ . Horizontal red lines correspond to the approximation given by Proposition 6,  $\rho \cos(\phi)r_K$  for the real part and  $\rho \sin(\phi)r_K$  for the imaginary part, with  $r_K = K(d_1 + d_2)/\sqrt{K(2d_1)K(2d_2)}$ .

scales. Thus, for both of them, we consider  $j_0 = 1$  when the two components of the time series are stationary, and  $j_0 = 2$  when a component is not stationary. These choices are motivated by previous studies [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.

When  $\mathbf{d} = (0.2, 0.2)$  or  $\mathbf{d} = (0.2, 0.4)$ , the processes are stationary, it is then possible to estimate the parameters of the model using a Fourier-based procedure. Following Baek et al. [2020], we implemented a Fourier-Based local Whittle estimation by keeping our parametrization. In particular, the complex matrix  $\Theta$  is not decomposed in real and imaginary parts or in magnitude and phase in the objective function. Let  $\lambda_j = 2\pi j/N$ ,  $j = 1, \dots, m$ , be the Fourier frequencies used in estima-



tion,  $m \in \mathbb{N}$ . Define  $\mathbf{\Lambda}_j^F(\mathbf{d}) = \text{Diag}\left(\lambda_j^{\mathbf{d}} e^{-i\lambda_j \mathbf{d}/2}\right)$ . The Fourier-based estimators  $(\hat{\mathbf{d}}^{MFW}, \hat{\Theta}^{MFW})$  are minimizers of the objective function

$$\frac{1}{m} \sum_{j=1}^m \left[ \log \det \left( \mathbf{\Lambda}_j^F(\mathbf{d}) \Theta \mathbf{\Lambda}_j^F(\mathbf{d})^* \right) + \mathbf{W}^F(\lambda_j)^* \left( \mathbf{\Lambda}_j^F(\mathbf{d}) \Omega(\mathbf{d}) \mathbf{\Lambda}_j^F(\mathbf{d})^* \right)^{-1} \mathbf{W}^F(\lambda_j) \right].$$

Note that this objective function differs from Baek et al. [2020] since we chose a different parametrization (see Section 4.2). We consider  $m = N^{0.65}$ , as suggested by Lobato [1999], Shimotsu [2007], Nielsen [2011].

Concerning the estimation with CFW-C(M,L), it has been shown in Section 3.3 that the approximation of Proposition 5 is not accurate for the first scales. Therefore, the procedure is executed with  $j_0 = 4$  in each case, to reduce the bias of the long-run covariance estimation. Since fewer scales are available in the procedure, the variance increases and the quality of the estimation is lower than the one based on CFW-PR(M,L). Table 2 illustrates that the quality of the estimation with CFW-C(M,L) filters presents a good accuracy, however lower than that with CFW-PR(M,L) filters.

$\mathbf{d}$	bias	std	RMSE	PR/Real	PR/Fourier
0.2	-0.0066	0.0165	0.0178	1.0375	1.0000
0.2	-0.0069	0.0156	0.0171	1.0228	1.0108
0.2	-0.0080	0.0165	0.0183	1.1537	1.0000
0.4	-0.0138	0.0159	0.0211	1.1888	0.9606
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. PR/Fourier denotes the ratio between the RMSE given by CFW-PR(4,4) filter and the RMSE given by a Fourier-based local Whittle procedure.

Table 3 and Table 4 give the results for the estimation of the covariance structure, respectively for CFW-PR(4,4) filter and for CFW-C(4,4) filter. The quality of the 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.

$\mathbf{d}$	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 in Table 1.

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 is for example of order  $\pi/10$  when estimating the phase of  $3\pi/10$ , corresponding to the case  $\mathbf{d} = (0.2, 0.8)$ . However, our results for the stationary cases are comparable or significantly better than those obtained by the Fourier-based Whittle estimation. Similar qualities were observed 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.

It can be observed that the estimation with CFW-C(4,4) filters has a lower bias and a higher variance than the 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 according to the choice of  $j_0$ . Since the bias for the phase estimation is critical, a 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. Thus, to ensure a good approximation and consequently a small bias 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), even if this increases the variance.

Table 5 and Table 6 display the results 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. The results for the long-run covariance are omitted, for simplicity. Table 5 shows that even if the bias of  $\hat{\mathbf{d}}$

$\mathbf{d}$		bias	std	RMSE	PR/Real	PR/Fourier
( 0.2 , 0.2 )	$\Omega_{1,1}$	0.0087	0.0227	0.0243	0.5089	1.6669
	$\Omega_{1,2}$	0.0076	0.0204	0.0218	0.5509	1.6621
	$\Omega_{2,2}$	0.0092	0.0229	0.0247	0.5184	1.6796
	correlation	4e-04	0.0057	0.0057	0.9702	0.6987
	phase	1e-04	0.0081	0.0081	.	1.0869
( 0.2 , 0.4 )	$\Omega_{1,1}$	0.0082	0.0229	0.0243	0.4919	1.7049
	$\Omega_{1,2}$	0.0028	0.0206	0.0208	0.4883	1.4759
	$\Omega_{2,2}$	0.0146	0.0231	0.0273	0.3142	1.3126
	correlation	-0.0063	0.0058	0.0085	0.6079	0.6950
	phase	0.1879	0.0083	0.1880	.	0.1677
( 0.2 , 0.8 )	$\Omega_{1,1}$	-0.0184	0.0365	0.0409	0.8883	.
	$\Omega_{1,2}$	-0.0860	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  $\Theta$  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  $\mathbf{d} \in \{(0.2, 0.8)\}$ . PR/Fourier denotes the ratio between the RMSE given by CFW-PR(4,4) filter and the RMSE given by a Fourier-based local Whittle procedure.

increases, it has a similar order of magnitude. Interestingly, Table 6 highlights that the quality of the estimations with 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 shown in Table 3. With  $j_0 = 4$ , the RMSE is significantly lower than the one of the Fourier-based local Whittle estimator. This approach, hence, improves significantly the estimation of the phase compared to the other procedures. Note that it is nevertheless sensitive to the choice of  $j_0$ .

## 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].

$\mathbf{d}$		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  $\Theta$  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 in Table 3.

The  $p$ -multivariate fractional Brownian motion  $\mathbf{X} = \{\mathbf{X}(t), t \in \mathbb{R}\}$  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  $\mathbf{d} - 1/2$ .

$\mathbf{d}$	bias	std	RMSE	PR/Real	PR/Fourier	C/PR
0.2	-0.0020	0.0412	0.0412	1.1056	1.0000	1.0068
0.2	-0.0037	0.0419	0.0421	1.1177	1.5829	1.0121
0.2	-0.0029	0.041	0.0411	1.0477	1.0000	0.9815
0.4	-0.0035	0.0418	0.0419	1.0560	1.5435	0.9682
0.2	-0.0054	0.0409	0.0412	0.9472	.	1.0173
0.8	-0.0078	0.0423	0.0430	0.9049	.	0.9855

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. PR/Fourier denotes the ratio between the RMSE given by CFW-PR(4,4) filter and the RMSE given by a Fourier-based local Whittle procedure. C/PR denotes the ratio between the RMSE given by CFW-C(4,4) filter in Table 2 and the RMSE given by CFW-PR(4,4) filter.

$\mathbf{d}$		bias	std	RMSE	PR/Fourier	C/PR
( 0.2 , 0.2 )	correlation	8e-04	0.0166	0.0166	0.9944	1.0477
	phase	0	0.0340	0.0340	1.5912	1.0346
( 0.2 , 0.4 )	correlation	2e-04	0.0164	0.0164	0.9745	1.0362
	phase	0.0241	0.0345	0.0421	0.1320	0.9842
( 0.2 , 0.8 )	correlation	-0.0033	0.0169	0.0172	.	1.0291
	phase	0.0654	0.0341	0.0738	.	0.9995

Table 6: Results for the estimation of matrices  $\Theta$  with CFW-PR(4,4) filters on ARFIMA processes with  $j_0 = 4$ . PR/Fourier denotes the ratio between the RMSE given by CFW-PR(4,4) filter and the RMSE given by a Fourier-based local Whittle procedure. C/PR denotes the ratio between the RMSE given by CFW-C(4,4) filter in Table 4 and the RMSE given by CFW-PR(4,4) filter.

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

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

where  $\text{Cor}(X_1, X_2)$  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  $f_{\ell,m}^{(1,1)}(\cdot)$  denotes the cross-spectral density of  $\{(1 - \mathbb{L})X_\ell(t), (1 - \mathbb{L})X_m(t), t \in \mathbb{Z}\}$ , then

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

with

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

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

Let  $\Theta$  be given by  $\Theta = (\Omega_{\ell,m} e^{i\phi_{\ell,m}})_{\ell,m=1,\dots,p}$ . 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  $\{d_\ell, \sigma_\ell, r_{\ell,m}, \eta_{\ell,m}, \ell, m = 1, \dots, p\}$  is not identifiable. Indeed, for  $0 < a < 1$ ,  $\{d_\ell, \sigma_\ell, r_{\ell,m}, \eta_{\ell,m}, \ell, m = 1, \dots, p\}$  and  $\{d_\ell, \sqrt{a} \sigma_\ell, r_{\ell,m}/a, \eta_{\ell,m}/a, \ell, m = 1, \dots, p\}$  lead to the same expressions of  $f_{\ell,m}^{(1,1)}(\cdot)$ . It thus seems reasonable to parameterize the fractional Brownian motion by  $\{d_\ell, \Theta_{\ell,m}, \ell, m = 1, \dots, p\}$ .

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 the approximation is observed except for the highest frequencies. Identical observations are obtained for CFW-C(4,4) filters (figure not provided).

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, considering  $j_0 = 4$  for both filters, the estimation procedures are equivalent for the two common-factor wavelets. Compared to the real wavelet-based estimation (with  $j_0 = 2$  as suggested by Achard and Gannaz [2019]), the RMSE increases. 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.

Table 9 and Table 10 give the results obtained for the estimation of the covariance structure, that

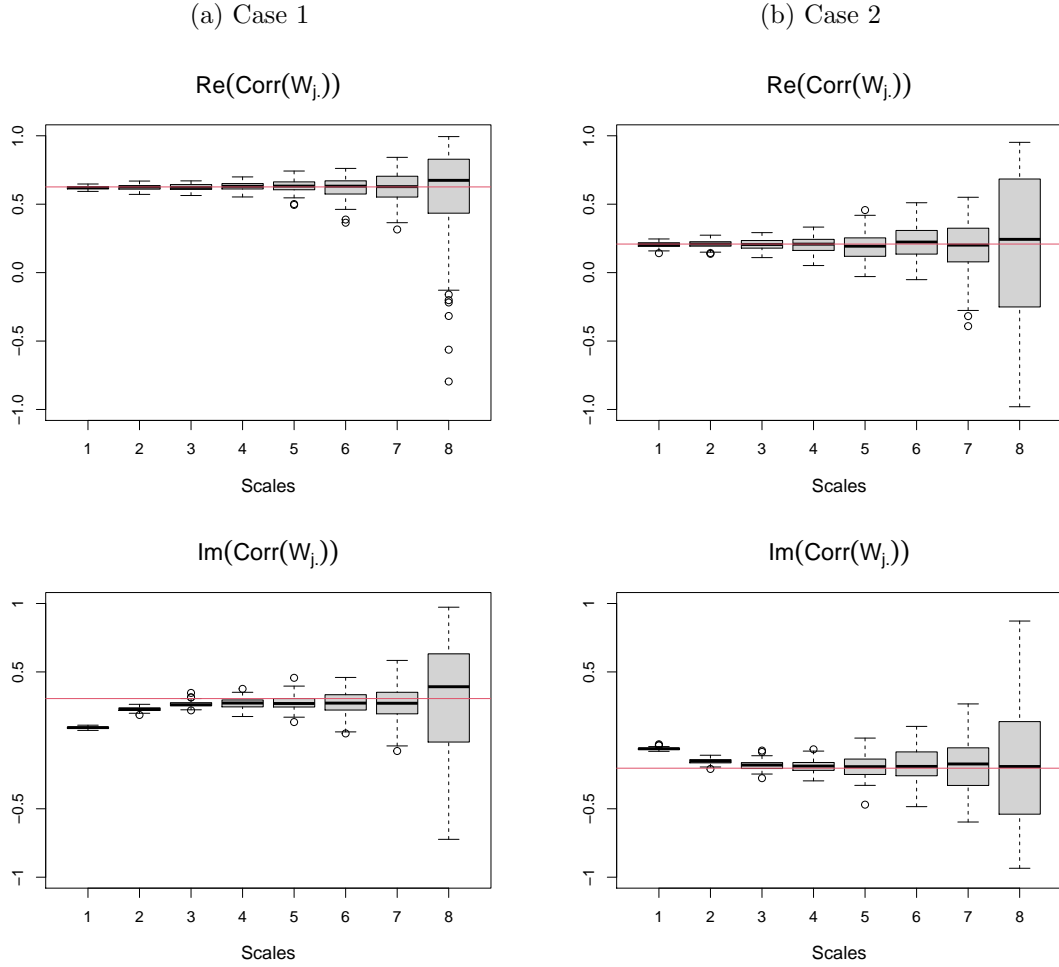


Figure 5: Boxplots of correlation between CFW-PR(4,4) coefficients at different scales for the simulated mFBM in Case 1 (left column-(a)) and in Case 2 (right column-(b)). First row gives the real part of the correlations and second row gives the imaginary part. Horizontal red lines correspond to the approximation given by Proposition 6, that is,  $\rho \cos(\phi)r_K$  for the real part and  $\rho \sin(\phi)r_K$  for the imaginary part, with  $r_K = K(d_1 + d_2)/\sqrt{K(2d_1)K(2d_2)}$ .

is,  $\Omega$ ,  $\rho$  and  $\phi$ . It is not possible to compare our results with alternative non parametric procedures because real wavelet-based procedure estimates the real part of the long-run covariance or of the correlation, and Fourier-based estimations are not valid for non-stationary time series.

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



	$\mathbf{d}$	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.

estimation of  $\rho$  and of  $\phi$ .

		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  $\Theta$  with CFW-PR(4,4) filter on mFBMs. Hyperparameter  $j_0$  satisfies  $j_0 = 4$ .

To conclude, no major difference are observed between CFW-PR and CFW-C filters. As theoretical results are also available for CFW-C filters, it seems preferable to use them in practice.

		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  $\Theta$  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 have applied our framework on fMRI data acquired on rats. We consider functional Magnetic Resonance images (fMRI) of 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 dataset, we know that for dead rats the recordings are just 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 anesthesia. The dataset is 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 dead and live rats (the full 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 scan 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,  $\hat{\mathbf{d}}$ ,
- the magnitude of the correlations,  $\hat{\boldsymbol{\rho}} = \{\hat{\rho}_{\ell,m}, 1 \leq \ell < m \leq p\}$  with  $\hat{\rho}_{\ell,m} = \frac{\hat{\Omega}_{\ell,m}}{\sqrt{\hat{\Omega}_{\ell,\ell}\hat{\Omega}_{m,m}}}$ ,
- the phases,  $\hat{\boldsymbol{\phi}} = \{\hat{\phi}_{\ell,m}, 1 \leq \ell < m \leq p\}$ .

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  $\hat{\mathbf{d}}$ . 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, the densities are not centered around zero and the variance between brain regions is higher than what is observed for dead rats. Long-memories for rats under anesthetic ISO\_W are higher than under other anesthetics.

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

First of all, as expected, the quantities obtained for the dead rats appear significantly different from those of the living rats. For dead rats, the distributions have a small support, that is, only 9 values out of 5100 (0.18%) satisfy  $\hat{\rho} > 0.3$ . Also note that no major differences are observed between the rats. Then, ISO\_W and ETO\_L show quite similar distributions, with possibly high magnitudes. On the other hand, the correlations for the anesthetic MED\_L are weaker. These

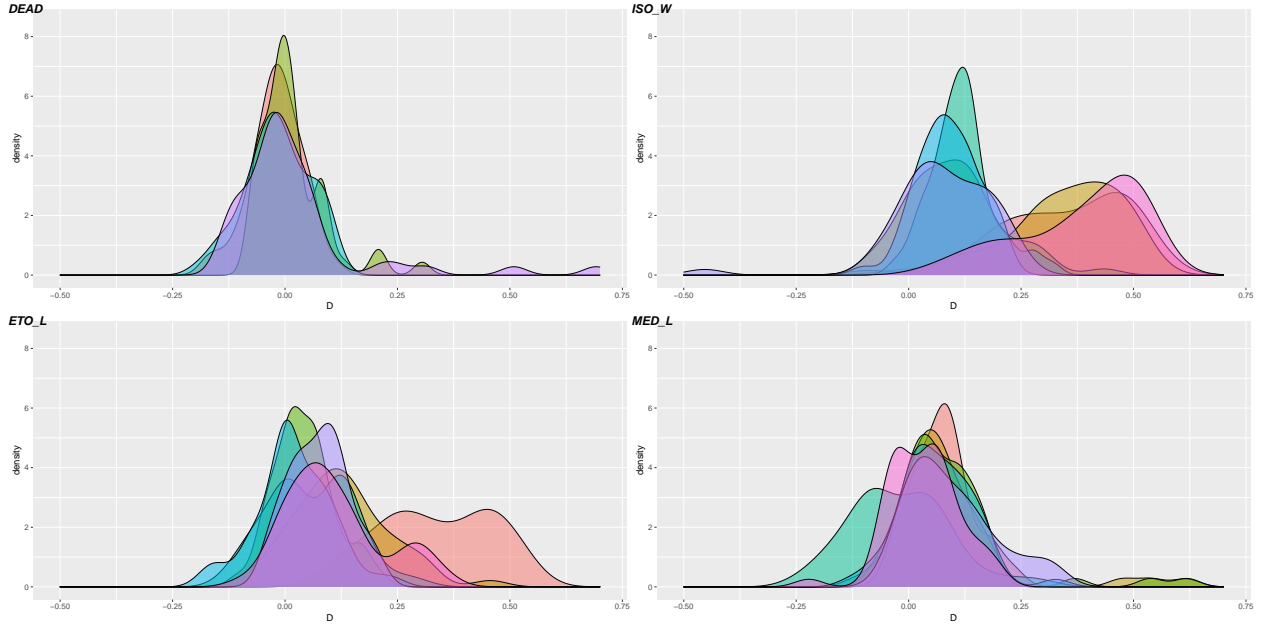


Figure 6: Plot of the empirical distribution of the long memory parameters  $\hat{\mathbf{d}}$  obtained for the 4 groups of rats. Each color corresponds to a rat.

results tend to show that the anesthetic MED\_L is more potent than other anesthetics, resulting in fewer connections between brain regions.

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

For dead rats, we observe mainly uniform distributions. For live rats, Figure 8 shows that the distributions have heavy tails. The tails are heavier for MED\_L anesthetic than for other anesthetics. This can be explained by the fact that the phase is non-informative when the magnitude is close to zero. As indicated previously, this problem of identifiability occurs for example in the case of fractional co-integration. Following Baek et al. [2020] another parametrization could be proposed to overcome it. The parametrization chosen here, nevertheless, seems more appropriate since the magnitude is crucial in this real data application.

To illustrate this fact, Figure 9 shows the distributions of the estimated phases  $\phi$  corresponding

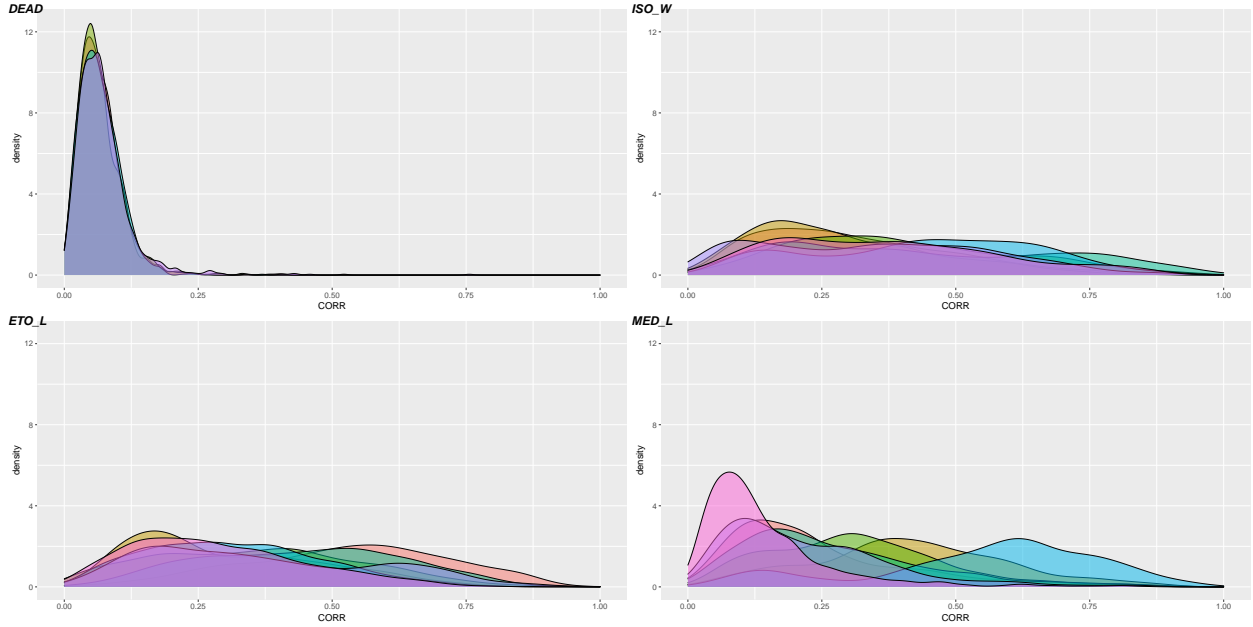


Figure 7: Plot of the empirical distribution of the correlation magnitudes  $\hat{\rho}$  obtained for the 4 groups of rats. Each color corresponds to a rat.

to magnitudes satisfying  $\rho > 0.3$  (this choice is motivated by the observation on the support of dead rats' correlations above). The distributions then have smaller tails. It can be observed that the supports 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 regions, 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 mean of the estimated phase for each detected edge. Figure 10 illustrates the graphs obtained

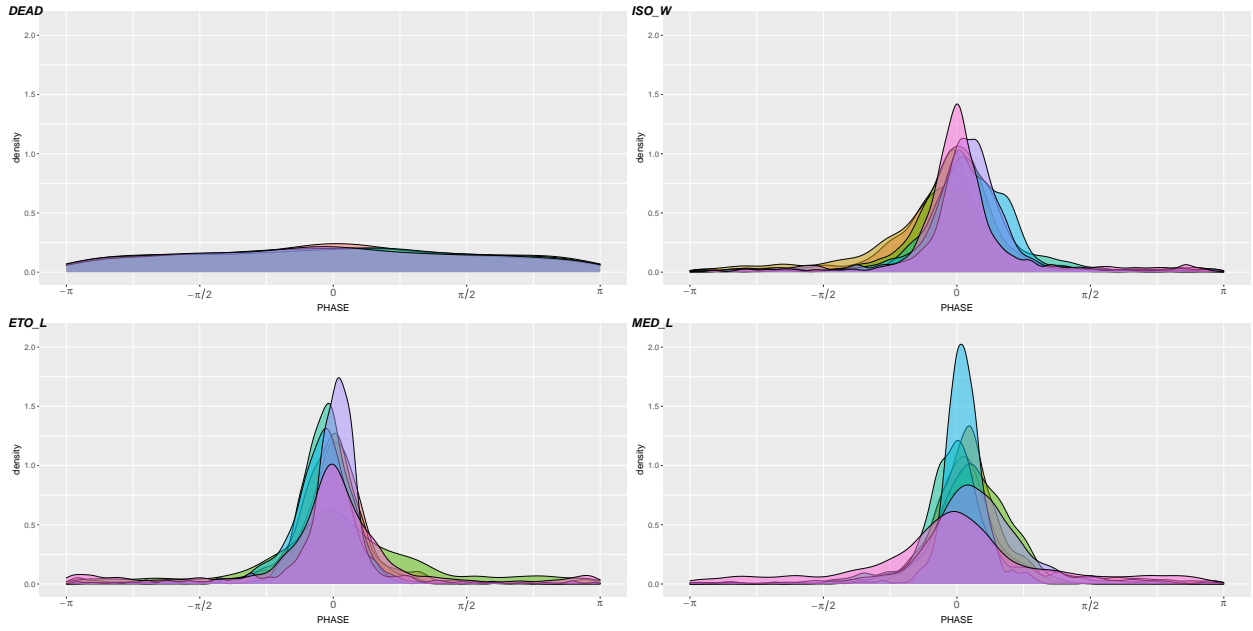


Figure 8: Plot of the empirical distribution of the phases  $\hat{\phi}$  obtained for the 4 groups of rats without thresholding the correlations. Each color corresponds to a rat.

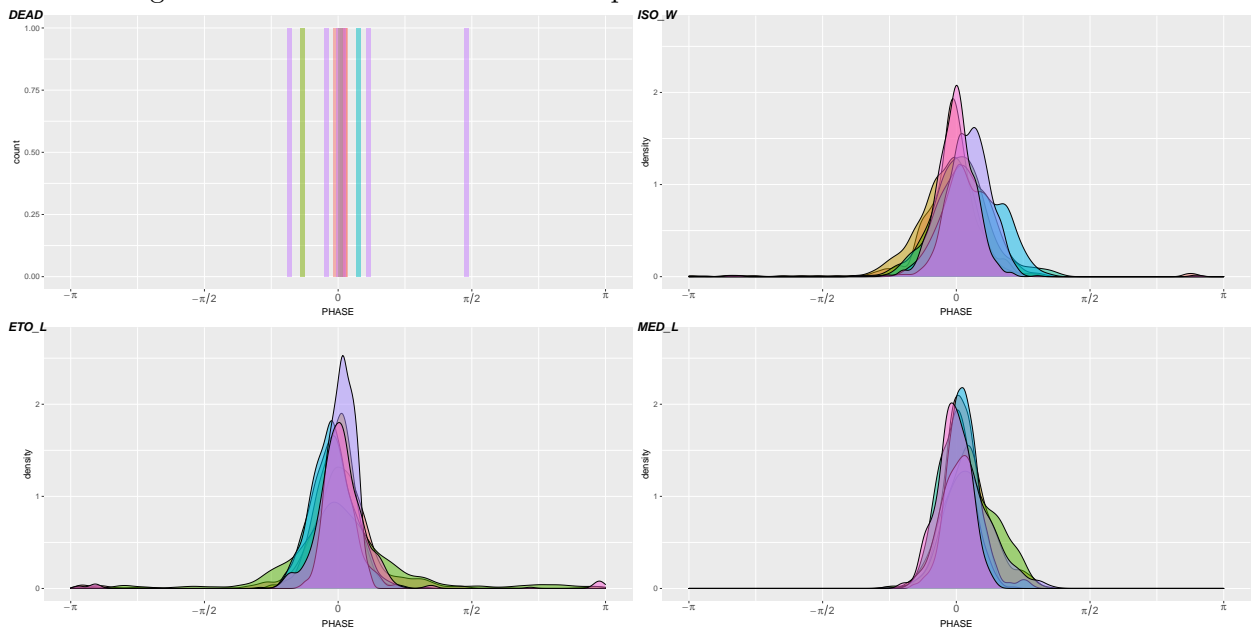


Figure 9: Plot of the empirical distribution of the phases obtained for the 4 groups of rats after first thresholding the correlations. Only phases associated to correlations with a magnitude higher than 0.3 are considered. Each color corresponds to a rat. For dead rats, a bar plot is provided rather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

for the 4 different groups.

We have colored each edge according to the mean 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 modeling used in Achard and Gannaz [2016] with similar data. The more the edges are colored, the more the behavior of the phase differs from the preceding modeling.

The DEAD group has indeed no edges. The MED\_L group has fewer 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 the mean phases are outside the interval  $[-1.1|\phi^*|, 1.1|\phi^*|]$  for ETO\_L and ISO\_W groups, with similar proportions. This observation is interesting because it illustrates that the modeling of these data is complex. The introduction of a general phase enables to take this complexity into account. 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 with the neuronal activity. The observed delay is 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 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 anesthesia.

## 7 Conclusion

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

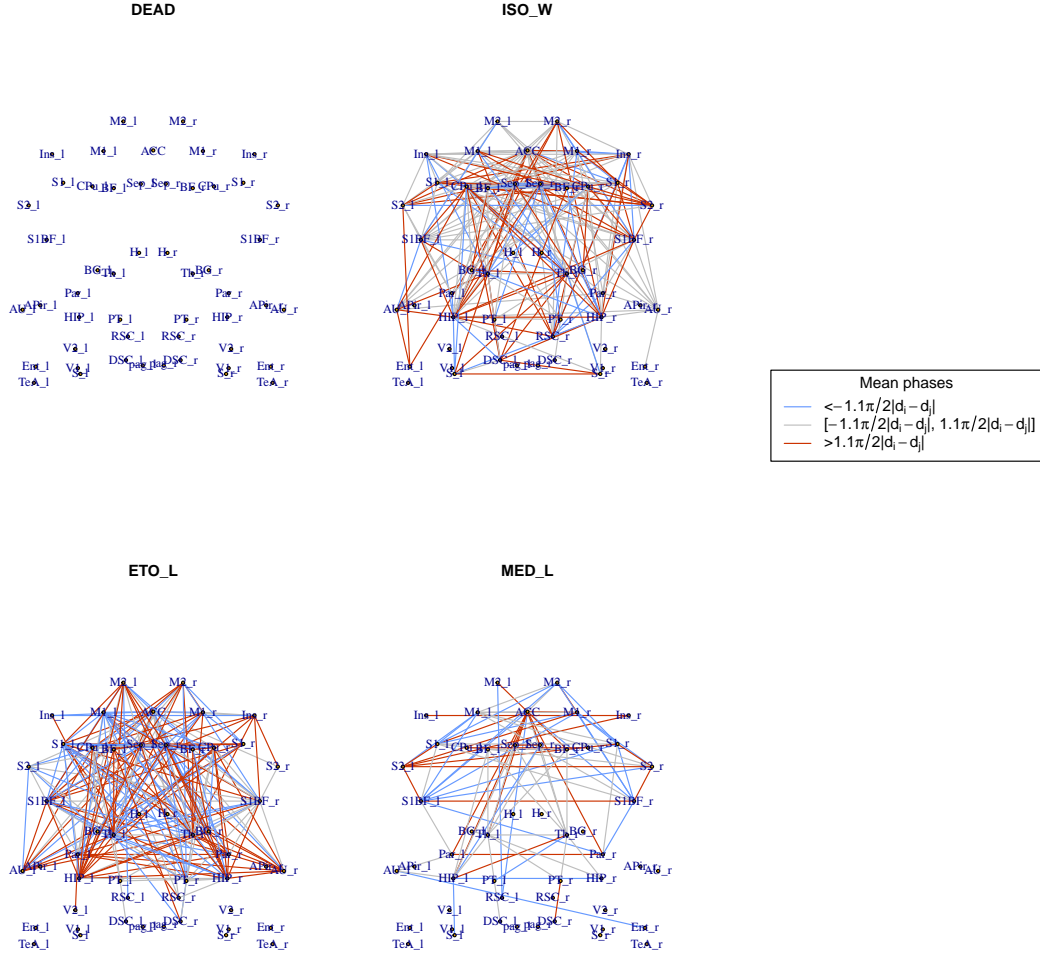


Figure 10: Plot of the average graphs with correlations and phases obtained for 4 groups of rats: DEAD, ISO\_W, ETO\_L and MED\_L. Only edges corresponding to a mean correlation's magnitude higher than 0.3 are displayed. Red edges correspond to positive mean phases higher than  $1.1|\phi^*|$ , blue edges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



phase. A simulation study on linear processes and on multivariate Brownian motions illustrates 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 brain activity more intensively.

## Acknowledgments

The authors are indebted to Marianne Clausel and François Roueff who contributed substantially to this work. We are grateful to the reviewers and to the co-editor who help to significantly improve the paper. We also thank Jean-François Coeurjolly for providing the code for the simulations of multivariate fractional 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  $\widehat{h}^{(L)}(\cdot)$  and the high-pass filter  $\widehat{h}^{(H)}(\cdot)$ :

$$\widehat{h}^{(L)}(\lambda) = \sqrt{2} \left( \frac{1 + e^{-i\lambda}}{2} \right)^M \widehat{q}_{L,M}(\lambda) \widehat{d}_L(\lambda) \quad \text{and} \quad \widehat{h}^{(H)}(\lambda) = \overline{\widehat{h}^{(L)}(\lambda + \pi)} e^{-i\lambda}, \quad (14)$$

for all  $\lambda \in \mathbb{R}$ . All the same, for all  $\lambda \in \mathbb{R}$ ,

$$\widehat{g}^{(L)}(\lambda) = \sqrt{2} \left( \frac{1 + e^{-i\lambda}}{2} \right)^M \widehat{q}_{L,M}(\lambda) \overline{\widehat{d}_L(\lambda)} e^{-i\lambda L} \quad \text{and} \quad \widehat{g}^{(H)}(\lambda) = \overline{\widehat{g}^{(L)}(\lambda + \pi)} e^{-i\lambda}. \quad (15)$$

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

$$\widehat{\varphi}_h(\lambda) = 2^{-1/2} \prod_{j=1}^{\infty} 2^{-1/2} \widehat{h}^{(L)}(2^{-j}\lambda) = 2^{-1/2} \prod_{j=1}^{\infty} \left( \frac{1 + e^{-i2^{-j}\lambda}}{2} \right)^M \widehat{q}_{L,M}(2^{-j}\lambda) \widehat{d}_L(2^{-j}\lambda),$$

and  $\widehat{\varphi}_g(\lambda) = 2^{-1/2} \prod_{j=1}^{\infty} \left( \frac{1 + e^{-i2^{-j}\lambda}}{2} \right)^M \widehat{q}_{L,M}(2^{-j}\lambda) \overline{\widehat{d}_L(2^{-j}\lambda)} e^{-i2^{-j}\lambda L}.$

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

$$\widehat{\varphi}_h(\lambda) = \prod_{j=1}^{\infty} \left( \frac{1 + e^{-i2^{-j}\lambda}}{2} \right)^M \overline{\widehat{d}_L(2^{-j}\lambda)}, \quad (16)$$

$$\widehat{\varphi}_g(\lambda) = \prod_{j=1}^{\infty} \left( \frac{1 + e^{-i2^{-j}\lambda}}{2} \right)^M \overline{\widehat{d}_L(2^{-j}\lambda)} e^{-i2^{-j}\lambda L}. \quad (17)$$

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

$$\begin{aligned} \widehat{\psi}_h(\lambda) &= 2^{-3/2} \widehat{h}^{(H)}(\lambda/2) \prod_{j=2}^{\infty} 2^{-1/2} \widehat{h}^{(L)}(2^{-j}\lambda) \\ &= 2^{-1} \left( \frac{1 - e^{i\lambda/2}}{2} \right)^M \overline{\widehat{q}_{L,M}(\lambda/2 + \pi) \widehat{d}_L(\lambda/2 + \pi)} e^{-i\lambda/2} \prod_{j=2}^{\infty} \left( \frac{1 + e^{-i2^{-j}\lambda}}{2} \right)^M \widehat{q}_{L,M}(2^{-j}\lambda) \widehat{d}_L(2^{-j}\lambda), \end{aligned}$$

and

$$\begin{aligned} \widehat{\psi}_g(\lambda) &= -2^{-1/2} \left( \frac{1 - e^{i\lambda/2}}{2} \right)^M \overline{\widehat{q}_{L,M}(\lambda/2 + \pi) \widehat{d}_L(\lambda/2 + \pi)} e^{+i(\lambda/2 + \pi)L} e^{-i\lambda/2} \\ &\quad \prod_{j=2}^{\infty} \left( \frac{1 + e^{-i2^{-j}\lambda}}{2} \right)^M \widehat{q}_{L,M}(2^{-j}\lambda) \overline{\widehat{d}_L(2^{-j}\lambda)} e^{-i2^{-j}\lambda L}. \end{aligned}$$

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

$$\widehat{\psi}_h(\lambda) = 2^{-1/2} \left( \frac{1 - e^{-i\lambda/2}}{2} \right)^M \overline{\widehat{d}_L(\lambda/2 + \pi)} e^{-i\lambda/2} \prod_{j=2}^{\infty} \left( \frac{1 + e^{-i2^{-j}\lambda}}{2} \right)^M \widehat{d}_L(2^{-j}\lambda)$$

$$\widehat{\psi}_g(\lambda) = -2^{-1/2} \left( \frac{1 - e^{-i\lambda/2}}{2} \right)^M \widehat{d}_L(\lambda/2 + \pi) e^{+i(\lambda/2 + \pi)L} e^{-i\lambda/2} \prod_{j=2}^{\infty} \left( \frac{1 + e^{-i2^{-j}\lambda}}{2} \right)^M \overline{\widehat{d}_L(2^{-j}\lambda)} e^{-i2^{-j}\lambda L}.$$

We shall use the following equality

$$\prod_{\ell=1}^{\infty} \left( \frac{1 + e^{-i2^{-\ell}\lambda}}{2} \right) = \frac{1 - e^{-i\lambda}}{-i\lambda} = e^{-i\lambda/2} \frac{\sin(\lambda/2)}{\lambda/2}.$$

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

$$\widehat{\varphi}_h(\lambda) = 2^{-1/2} \left( \frac{\sin(\lambda/2)}{\lambda/2} \right)^M \prod_{j=1}^{\infty} \overline{\widehat{d}_L(2^{-j}\lambda)}, \quad (18)$$

$$\widehat{\varphi}_g(t) = 2^{-1/2} \left( \frac{\sin(\lambda/2)}{\lambda/2} \right)^M \prod_{j=1}^{\infty} \overline{\widehat{d}_L(2^{-j}\lambda)} e^{-i2^{-j}\lambda L}, \quad (19)$$

and

$$\widehat{\psi}_h(\lambda) = 2^{-1/2} \sin(\lambda/4)^M \left( \frac{\sin(\lambda/4)}{\lambda/4} \right)^M \overline{\widehat{d}_L(\lambda/2 + \pi)} e^{-i\lambda} \prod_{j=2}^{\infty} \widehat{d}_L(2^{-j}\lambda), \quad (20)$$

$$\widehat{\psi}_g(\lambda) = 2^{-1/2} \sin(\lambda/4)^M \left( \frac{\sin(\lambda/4)}{\lambda/4} \right)^M \widehat{d}_L(\lambda/2 + \pi) e^{+i(\lambda + \pi)L} e^{-i\lambda} \prod_{j=2}^{\infty} \overline{\widehat{d}_L(2^{-j}\lambda)} e^{-i2^{-j}\lambda L}. \quad (21)$$

## A.2 A first property

The function  $\widehat{d}_L(\cdot)$  satisfies, for all  $\lambda \in \mathbb{R}$ ,

$$\widehat{d}_L(\lambda) = e^{i\lambda(-L/2+1/4)} [\cos(\lambda/4)^{2L+1} + i(-1)^{L+1} \sin(\lambda/4)^{2L+1}]. \quad (22)$$

We deduce the following lemma.

**Lemma 9.**

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

The proof is straightforward and it is thus omitted.

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

**Lemma 10.**

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

with  $C_\infty = 1$ .

### A.3 Expression of the wavelet filters

Recall that

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

Hence,

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

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

## 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  $M \geq 0$  and  $C_m > 0$  such that for all  $j \geq 0$  and  $\lambda \in \mathbb{R}$ ,

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

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

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

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

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  $k \in \mathbb{Z} \setminus \{0\}$ ,

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

Properties (W1), (W2), (W3) and (W4) correspond respectively to (W1), (W2), (W3) and (W4) of [Moulines et al. \[2007, 2008\]](#), [Achar and Gannaz \[2016\]](#) in the context of real-wavelets. We can establish that they are satisfied by CFW-C(L,M) wavelets. Assumption (W1) is given in [Lemma 4](#). The wavelets supports are finite when  $L$  and  $M$  are finite.

**Proposition 11.** *When  $M \geq 2$ , and  $L \geq 1$ , 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 3](#).

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

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

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

**Proposition 12.** *Suppose (W1), (W2), (W3), and (W4) and Lemma 10 hold. Then, for all  $j \geq 0$ , for all  $\lambda \in \mathbb{R}$ ,*

$$\left| \widehat{\tau}_j(\lambda) - 2^{j/2} \widehat{\varphi}(\lambda) \overline{\widehat{\psi}(2^j \lambda)} \right| \leq C_\tau 2^{j(1/2-\alpha)} |\lambda|^M, \quad (23)$$

$$\left| \widehat{\varphi}(\lambda) \overline{\widehat{\psi}(2^j \lambda)} \right| \leq C_\infty C_m |2^j \lambda|^M, \quad (24)$$

$$\left| \widehat{\varphi}(\lambda) \overline{\widehat{\psi}(2^j \lambda)} \right| \leq \frac{C_\infty C_s}{(1 + 2^j |\lambda|)^\alpha}, \quad (25)$$

$$\left| 2^{-j/2} \widehat{\tau}_j(2^{-j} \lambda) \right| \leq C_{m\tau} |\lambda|^M, \quad (26)$$

and, for all  $j, j' \geq 0$ , for all  $|2^{-j} \lambda| \leq \pi$ ,

$$\left| 2^{-j/2} \widehat{\tau}_j(2^{-j} \lambda) \right| \leq C_{s\tau} (1 + |\lambda|)^{-\alpha}, \quad (27)$$

$$\left| 2^{-j/2} \widehat{\tau}_j(2^{-j} \lambda) \right| \leq C_{ms\tau} |\lambda|^M (1 + |\lambda|)^{-\alpha-M}, \quad (28)$$

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

$$\sup_{|\lambda| \leq \pi} \left| \left| 2^{-j/2} \widehat{\tau}_j(2^{-j} \lambda) \right|^2 - \left| \widehat{\psi}(\lambda) \right|^2 \right| \leq C_a 2^{-\gamma j} |\lambda|^{2M}, \quad (29)$$

$$\sup_{1 \leq |\lambda| \leq 2^j \pi} \left| \left| 2^{-j/2} \widehat{\tau}_j(2^{-j} \lambda) \right|^2 - \left| \widehat{\psi}(\lambda) \right|^2 \right| \leq C_a 2^{-\gamma j} |\lambda|^2 (1 + |\lambda|)^{-\alpha}, \quad (30)$$

with  $\gamma = 2$  and  $C_a = 2 \cdot 5^{2M} \cdot (M + L + 1)$ .

The proof is given in Section B.3.

## B.1 Proof of Proposition 11

We first establish that CFW-C(M,L) filters satisfy properties (W2)–(W4).

### B.1.1 Property (W2)

Let  $\lambda \in \mathbb{R}$ . Recall that  $\widehat{\psi}_h(\lambda)$  and  $\widehat{\psi}_g(\lambda)$  are given respectively by (20) and (21). Observe that  $|\sin(\lambda/4)|^M \leq 4^{-M} |\lambda|^M$ . Since  $\sup |\widehat{d}_L| = 1$ , assumption (W2) follows with a constant  $C_m = 1$ , for all  $M \geq 1$ .

### B.1.2 Property (W3)

Let  $\lambda \in \mathbb{R}$ . Recall that  $\widehat{\psi}_h(\lambda)$  is given by (20). Since  $\sup_{\lambda \in \mathbb{R}} |\widehat{d}_L(\lambda)| \leq 1$ , we obtain

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

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

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

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

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

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

### B.1.3 Property (W4)

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

- If  $k$  is odd,

Using the expressions of  $\widehat{\varphi}_h(\lambda)$  and  $\widehat{\varphi}_g(\lambda)$  given respectively by (16) and (17),

$$\begin{aligned}\widehat{\varphi}_h(\lambda) &= (\cos(\lambda/4))^M \overline{\widehat{d}_L(\lambda/2)} \prod_{j=2}^{\infty} \left( \frac{1 + e^{-i2^{-j}\lambda}}{2} \right)^M \overline{\widehat{d}_L(2^{-j}\lambda)}, \\ \widehat{\varphi}_g(\lambda) &= (\cos(\lambda/4))^M \overline{\widehat{d}_L(\lambda/2)} e^{-i\lambda L/2} \prod_{j=2}^{\infty} \left( \frac{1 + e^{-i2^{-j}\lambda}}{2} \right)^M \overline{\widehat{d}_L(2^{-j}\lambda)} e^{-i2^{-j}\lambda L}.\end{aligned}$$

Observe that

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

With Lemma 9, we obtain that

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

Hence,

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

- If  $k$  is even,  $k \geq 2$ ,

Let us use the expressions of  $\widehat{\varphi}_h(\lambda)$  and  $\widehat{\varphi}_g(\lambda)$  given respectively by (18) and (19). We have,

$$\left| \frac{\sin(\lambda/2)}{\lambda/2} \right|^M = \left| \frac{\sin(\omega/2)}{\omega/2} \right| \left| \frac{\omega}{\lambda} \right|^M \leq \left| \frac{\omega}{\lambda} \right|^M \leq \left| \frac{\omega}{\pi} \right|^M.$$

Using Lemma 9,

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

Hence,

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

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



## B.2 Proof of Proposition 12

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 (23)

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

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

It follows that

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

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

$$\begin{aligned} \widehat{\tau}_j(\lambda) &= \int_{-\infty}^{\infty} \left( \sum_{k \in \mathbb{Z}} \widehat{\varphi}(\lambda + 2k\pi) e^{it(\lambda + 2k\pi)} \right) 2^{-j/2} \psi(2^{-j}t) dt \\ &= \sum_{k \in \mathbb{Z}} \widehat{\varphi}(\lambda + 2k\pi) \int_{-\infty}^{\infty} 2^{-j/2} \psi(2^{-j}t) e^{it(\lambda + 2k\pi)} dt \\ &= 2^{j/2} \sum_{k \in \mathbb{Z}} \widehat{\varphi}(\lambda + 2k\pi) \overline{\widehat{\psi}(2^j(\lambda + 2k\pi))}. \end{aligned}$$

We obtain

$$\left| \widehat{\tau}_j(\lambda) - 2^{j/2} \widehat{\varphi}(\lambda) \overline{\widehat{\psi}(2^{-j}\lambda)} \right| = 2^{j/2} \left| \sum_{k \in \mathbb{Z}, k \neq 0} \widehat{\varphi}(\lambda + 2k\pi) \overline{\widehat{\psi}(2^j(\lambda + 2k\pi))} \right|. \quad (31)$$

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

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

Using the inequality above and (W4) in (31), we get

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

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

### B.2.2 Proof of inequality (24) and inequality (25)

Using (W2), (W3) and Lemma 10, inequalities (24) and (25) are straightforward.

### B.2.3 Proof of inequality (26) and inequality (27)

With inequalities (23), (24) and (25), we get

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

It follows that

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

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 (28)

With (26), inequality (28) is straightforward with  $C_{ms\tau} = C_{m\tau} (1 + \pi)^{\alpha+M}$ .

### B.3 Proof of Proposition 13

Let  $j \geq 1$  and  $|\lambda| \leq 2^j \pi$ . Inequalities (23) and (24) imply that

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

with  $C_{a1} = C_\tau (C_{m\tau} + C_m C_\infty)$  and  $\gamma_1 = M + \alpha$ . Similarly, using rather inequalities (23), (27) and (25),

$$\left| \left| 2^{-j/2} \widehat{\tau}(2^{-j}\lambda) \right|^2 - \left| \widehat{\varphi}(2^{-j}\lambda) \widehat{\psi}(\lambda) \right|^2 \right| \leq C_{b1} 2^{-\gamma_1 j} |\lambda|^M (1 + |\lambda|)^{-\alpha},$$

with  $C_{b1} = C_\tau (C_{s\tau} + C_s C_\infty)$ .

Next,

$$\left| \left| \widehat{\varphi}(2^{-j}\lambda) \widehat{\psi}(\lambda) \right|^2 - \left| \widehat{\psi}(\lambda) \right|^2 \right| \leq \left| \widehat{\psi}(\lambda) \right|^2 \left| |\widehat{\varphi}(2^{-j}\lambda)|^2 - 1 \right|$$

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

**Lemma 14.** *There exists a constant  $C_Z = (M + L + 1)$  such that for all  $j \in \mathbb{N}$ , for all  $|2^{-j}\lambda| < \pi$ ,*

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

*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

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

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  $x \in \mathbb{R} \setminus \{0\}$ ,  $|\sin(x)/x| \leq 1$  and  $|\widehat{d}_L(x)| \leq 1$ .

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

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

We get

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

Hence,

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

with  $C_Z = (M + L + 1)/2$ . □

We deduce that

$$\left| \left| \widehat{\varphi}(2^{-j}\lambda) \overline{\widehat{\psi}(\lambda)} \right|^2 - \left| \widehat{\psi}(\lambda) \right|^2 \right| \leq C_Z |2^{-j}\lambda|^2 \left| \widehat{\psi}(\lambda) \right|^2.$$

Using respectively (W2) and (W3), we get

$$\begin{aligned} \left| \left| \widehat{\varphi}(2^{-j}\lambda)\overline{\widehat{\psi}(\lambda)} \right|^2 - \left| \widehat{\psi}(\lambda) \right|^2 \right| &\leq C_{a2} 2^{-\gamma_2 j} |\lambda|^{2M}, \text{ when } |\lambda| \leq \pi, \\ \left| \left| \widehat{\varphi}(2^{-j}\lambda)\overline{\widehat{\psi}(\lambda)} \right|^2 - \left| \widehat{\psi}(\lambda) \right|^2 \right| &\leq C_{b2} 2^{-\gamma_2 j} |\lambda|^2 (1 + |\lambda|)^{-2\alpha}, \text{ when } 1 \leq |\lambda| \leq 2^j \pi. \end{aligned}$$

with  $C_{a2} = C_m^2 C_Z \pi^2$ ,  $C_{b2} = C_s^2 C_Z$  and  $\gamma_2 = 2$ .

We obtain inequality (29) and (30) with  $C_a = \max\{C_{a1}, C_{a2}, C_{b1}, C_{b2}\}$ ,  $\gamma = \min\{\gamma_1, \gamma_2\}$ . Observe that for all  $M \geq 2$   $\gamma_1 > 2$ , and, hence,  $\gamma = 2$  and we can take  $C_a = 2.5^{2M} (M + L + 1)$ .

## C Proof of Proposition 2 and of Lemma 4

Proposition 7.2 of Mallat [1999] states that  $\varphi_g(\cdot)$  and  $\varphi_h(\cdot)$  have the same support of the conjugate mirror filters associated to  $h^{(L)}(\cdot)$  and  $g^{(L)}(\cdot)$ . When  $\widehat{q}_{L,M}(\cdot) = 1$ , (14) and (15) yield

$$\widehat{h}^{(L)}(\lambda) = 2^{-M+1/2} \left( \sum_{k=0}^M (e^{-i\lambda})^k \right) \widehat{d}_L(\lambda) \quad \text{and} \quad \widehat{g}^{(L)}(\lambda) = 2^{-M+1/2} \left( \sum_{k=0}^M (e^{-i\lambda})^k \right) e^{i\lambda} \overline{\widehat{d}_L(\lambda)},$$

for all  $\lambda \in \mathbb{R}$ . Moreover (22) writes as

$$\widehat{d}_L(\lambda) = 1 + \sum_{\ell=1}^L d(\ell) (e^{-i\lambda})^\ell, \quad \text{with } d(\ell) = (-1)^\ell \binom{L}{\ell} \prod_{k=0}^{\ell-1} \frac{1/2 - L + k}{3/2 + k}, \quad \ell = 1, \dots, L$$

(see *e.g.* Section 2.2 of Achard et al. [2020]). We deduce that  $\widehat{h}^{(L)}(\cdot)$  and  $\widehat{g}^{(L)}(\cdot)$  are polynomials of  $e^{i\lambda}$  with coefficients varying on. They are, hence, associated with conjugate mirror filters defined on  $\{0, \dots, M + L\}$  and on  $\{-1 - L, \dots, M - 1\}$ . Hence, the support of  $\varphi_g(\cdot)$  and  $\varphi_h(\cdot)$  are respectively  $[0, M + L]$  and on  $[-1 - L, \dots, M - 1]$ . Consequently, the support of  $\varphi(\cdot)$  is  $[-1 - L, \dots, M + L]$ , and has length  $M + 2L + 1$ .

Using again Proposition 7.2 of Mallat [1999], we deduce that the supports of  $\psi_g(\cdot)$  and  $\psi_h(\cdot)$  are respectively  $[-(M + L - 1)/2, (M + L + 1)/2]$  and  $[-(M + L)/2, (M + L)/2]$ , and thus that the support of  $\psi(\cdot)$  has length  $M + L + 1/2$ . Proposition 2 follows.

*Remark.* For CFW-PR(M,L) and CFW-C(M,L) filters, the presence of the filter  $\widehat{q}_{L,M}(\cdot)$  in  $\widehat{h}^{(L)}(\cdot)$  and  $\widehat{g}^{(L)}(\cdot)$  changes the supports. No theoretical statement provides the degree of  $\widehat{q}_{L,M}(\cdot)$ , but in practice,  $\widehat{q}_{L,M}(\cdot)$  is a polynomial of  $e^{-i\lambda}$  of degree  $M + L - 1$ . Then, the resulting supports of  $\varphi_h(\cdot)$  and  $\varphi_g(\cdot)$  are  $[0, 2M + 2L - 1]$  and  $[-1 - L, 2M + L - 1]$ . Similarly, we deduce that the supports of  $\psi_h(\cdot)$  and  $\psi_g(\cdot)$  are  $[-M - L, M + L]$  and  $[-M - L + 1/2, M + L + 1/2]$ .

Consider now the wavelet coefficients  $\{\widehat{\mathbf{W}}_{j,k}, j \geq 0, k \in \mathbb{Z}\}$ , as defined in Lemma 4. Denote  $[t_\varphi, T_\varphi]$  the support of  $\varphi(\cdot)$ , and suppose  $T_\varphi - t_\varphi \geq 1$ . The functions  $\widehat{\mathbf{X}}(t) = \sum_{k=1}^N \mathbf{X}(k)\varphi(t-k)$  and  $\widetilde{\mathbf{X}}(t) = \sum_{k \in \mathbb{Z}} \mathbf{X}(k)\varphi(t-k)$  coincide for  $t \in [T_\varphi, N + t_\varphi + 1]$ . Recall that  $n_j$  is the number of coefficients at a scale  $j \geq 0$  such that  $\int_{\mathbb{R}} \widehat{\mathbf{X}}(t)\psi_{j,k}(t)dt = \int_{\mathbb{R}} \widetilde{\mathbf{X}}(t)\psi_{j,k}(t)dt$ . Easy calculation then yields  $n_j = \max\{0, \lfloor 2^{-j}(N - L_\phi + 1) - L_\psi + 1 \rfloor\}$ , where  $L_\phi, L_\psi$  are the respective length of the supports of functions  $\phi(\cdot)$  and  $\psi(\cdot)$ . As a consequence,  $n_j = \max\{0, \lfloor 2^{-j}(N - 2L - M - 1) - L - M - 1/2 \rfloor\}$ , for all  $M \geq 1, L \geq 0$ .

Now, suppose that  $N^{-1}L \rightarrow 0$  and that  $2^j \leq NL^{-1}$ . Then, for all  $j$  such that  $2^j N^{-1}L \rightarrow 0$  when  $N$  goes to infinity,  $n_j 2^j N^{-1} \rightarrow 1$  when  $N$  goes to infinity. This concludes Lemma 4.

## D 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 5 which are stated below. To better highlight 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(M,L) filters, we have  $\alpha = M$  by Proposition 11. Hence, we formulate here the assumption on the parameters on both  $\alpha$  and  $M$ ,

$$(C-b) \quad -\alpha/2 + \beta/2 + 1/2 < d_\ell < M/2 \quad \text{for all } \ell = 1, \dots, p, M \geq 2 \text{ and } 0 < \beta < 2.$$

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

Proposition 5 follows from the following proposition.

**Proposition 15.** *Let  $\mathbf{X}$  be a  $p$ -multivariate long range dependent process with long memory parameters  $d_1, \dots, d_p$  with generalized spectral density  $\mathbf{f}(\cdot)$  satisfying (M-1) with short-range behavior*

(M-2). Consider  $\{W_{j,k}(\ell), \ell = 1, \dots, p, j \geq 0, k \in \mathbb{Z}\}$  the wavelet coefficients of  $\mathbf{X}$  obtained with CFW-C( $M, L$ ) filters,  $M, L \geq 2$ . Then we have, for all  $j \geq 0, k \in \mathbb{Z}$ ,

$$\begin{aligned} \left| 2^{-j(d_\ell + d_m)} \text{Cov}(W_{j,k}(\ell), W_{j,k}(m)) - \Omega_{\ell,m} \int_{-\infty}^{\infty} |\lambda|^{-d_\ell - d_m} e^{\text{sign}(\lambda) \phi_{\ell,m}} \left| \widehat{\psi}(\lambda) \right|^2 d\lambda \right| \\ \leq C'_1 \max\{2^{-j\beta}, L 2^{-2j}\}. \end{aligned}$$

where  $C'_1$  is a constant only depending on  $M$  and  $C_f, \beta, \|\Omega\|, \{d_\ell, \ell = 1, \dots, p\}$ .

## D.1 Proof of Proposition 15

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

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

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

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

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

We introduce

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

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  $\mathbf{f}^S(\cdot)$  at the origin, that is, (M-2), together with inequality (26).
- (b)  $2^{-j(d_\ell+d_m-\beta)} \left| \mathbf{B}_j^{(+)} - \mathbf{I}_j^{(+)\text{inf}} \right|$  is bounded using the convergence of the filter  $\tau_j$  to  $\widehat{\psi}(\cdot)$ , through Proposition 13. We shall need inequality (26) to control the integral around zero and inequality (27) to control the upper part.
- (c)  $2^{-j(d_\ell+d_m-\beta)} \left| \mathbf{I}_j^{(+)\text{sup}} \right|$  is bounded using the regularity of  $\widehat{\psi}(\cdot)$ , 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

$$\mathbf{I}^{(-)} = \int_{-\infty}^0 \mathbf{\Gamma}_j(\mathbf{d})^{-1} \mathbf{\Lambda}(\lambda) \mathbf{\Theta} \mathbf{\Lambda}(\lambda) \mathbf{\Gamma}_j(\mathbf{d})^{-1} \left| \widehat{\psi}(\lambda) \right|^2 d\lambda.$$

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

### D.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,

$$\begin{aligned} \left| \mathbf{A}_j^{(+)} - \mathbf{B}_j^{(+)} \right| &\leq \int_0^\pi |\mathbf{f}(\lambda) - \mathbf{\Lambda}(\lambda) \mathbf{\Theta} \mathbf{\Lambda}(\lambda)| |\widehat{\tau}_j(\lambda)|^2 d\lambda \\ &\leq \int_0^\pi |\mathbf{\Lambda}(\lambda) \mathbf{\Theta} \mathbf{\Lambda}(\lambda)| \circ |\mathbf{f}^S(\lambda) - 1| |\widehat{\tau}_j(\lambda)|^2 d\lambda. \end{aligned}$$

Property (M-2) gives

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

With a change of variable,

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



We split the integral in two parts. First, with (26),

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

As the parameters satisfy (C-b), 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 (27),

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

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

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

This step uses the convergence of the filter  $\widehat{\tau}_j$  to  $\widehat{\psi}(\cdot)$ , through Proposition 13. First,

$$\Gamma_j(\mathbf{d}) 2^{j\beta} \left| \mathbf{B}_j^{(+)} - \mathbf{I}_j^{(+)\text{inf}} \right| \Gamma_j(\mathbf{d}) \leq 2^{j\beta} \int_0^{2^{j\pi}} \mathbf{\Lambda}(\lambda) \mathbf{\Theta} \mathbf{\Lambda}(\lambda) \left| \left| 2^{-j/2} \widehat{\tau}_j(2^{-j}\lambda) \right|^2 - \left| \widehat{\psi}(\lambda) \right|^2 \right| d\lambda.$$

Using (30), for  $1 \leq |\lambda| \leq 2^j\pi$ ,  $\left| \left| 2^{-j/2} \widehat{\tau}_j(2^{-j}\lambda) \right|^2 - \left| \widehat{\psi}(\lambda) \right|^2 \right| \leq C_a 2^{-\gamma j} |\lambda|^{2-\alpha}$ . Thus,

$$\begin{aligned} & \left( \int_1^{2^{j\pi}} \mathbf{\Lambda}(\lambda) \mathbf{\Theta} \mathbf{\Lambda}(\lambda) 2^{j\beta} \left| \left| 2^{-j/2} \widehat{\tau}_j(2^{-j}\lambda) \right|^2 - \left| \widehat{\psi}(\lambda) \right|^2 \right| d\lambda \right)_{(\ell, m)} \\ & \leq \|\mathbf{\Omega}\| C_a 2^{j(\beta-\gamma)} \int_1^{2^{j\pi}} |\lambda|^{-d_\ell - d_m + 2 - \alpha} d\lambda. \end{aligned}$$

Depending on  $-d_\ell - d_m + 2 - \alpha + 1$  being negative, equal to zero or positive, the integral on the right-hand side is bounded up to a constant by 1, by  $j$  or by  $2^{j(-d_\ell - d_m + 2 - \alpha + 1)}$ . In the two first cases the right hand side goes to zero when  $j$  goes to infinity since  $\gamma > \beta$ . In the last case, using (C-b), we obtain the bound  $2^{j(\beta - d_\ell - d_m - \alpha + 1)}$  which goes to zero when  $j$  goes to infinity.

When  $M \geq 2$ ,  $\gamma = 2$ ,  $-d_\ell - d_m + 2 - 2\alpha + 1 < 0$  and  $C_a = 2(M + L + 1)$ . With a fixed  $M$ , this term is, hence, bounded up to a constant by  $L 2^{j(\beta-2)}$ .

It remains to consider the integral on  $(0, 1)$ . Property (29) states that

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

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

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

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

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

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

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

## D.2 Proof of Proposition 6

Recall that

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

Theorem 3 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

$$U_L(\lambda) \leq 2\sqrt{2} \left(\log_2\left(\frac{\max(4\pi, |\lambda|)}{2\pi}\right) + 2\right) \left(1 - \frac{\delta(\lambda, 4\pi\mathbb{Z})}{\max(4\pi, |\lambda|)}\right)^{2L+1}. \quad (32)$$

We deduce from Theorem 3 the following results, which gives inequalities in a form that can be more useful in future developments.

**Corollary 16.** *For all  $\widehat{q}_{L,M}(\cdot)$  real polynomial of  $(e^{-i\lambda})$ , for all  $|\lambda| \leq 2\pi$ ,*

$$\left|\left|\widehat{\psi}(\lambda)\right|^2 - 4\mathbb{1}_{\mathbb{R}_+}(\lambda) \left|\widehat{\psi}_h(\lambda)\right|^2\right| \leq 12\sqrt{2} \left(1 - \frac{|\lambda|}{2\pi}\right)^{2L+1} \left|\widehat{\psi}_h(\lambda)\right|^2.$$

For  $2\pi \leq |\lambda| \leq 4\pi$ ,

$$\left|\left|\widehat{\psi}(\lambda)\right|^2 - 4\mathbb{1}_{\mathbb{R}_+}(\lambda) \left|\widehat{\psi}_h(\lambda)\right|^2\right| \leq 12\sqrt{2} \left(\frac{|\lambda|}{4\pi}\right)^{2L+1} \left|\widehat{\psi}_h(\lambda)\right|^2.$$

The proof is straightforward and it thus omitted.

Let us introduce also

$$\begin{aligned}\mathbf{I}_j^{\text{inf}} &= \int_{-\pi 2^j}^{\pi 2^j} \Gamma_j(\mathbf{d})^{-1} \mathbf{\Lambda}(\lambda) \mathbf{\Theta} \mathbf{\Lambda}(\lambda) \Gamma_j(\mathbf{d})^{-1} \left| \widehat{\psi}(\lambda) \right|^2 d\lambda, \\ \mathbf{I}_j^{\text{inf,analytic}} &= \int_0^{\pi 2^j} \Gamma_j(\mathbf{d})^{-1} \mathbf{\Lambda}(\lambda) \mathbf{\Theta} \mathbf{\Lambda}(\lambda) \Gamma_j(\mathbf{d})^{-1} 4 \left| \widehat{\psi}_h(\lambda) \right|^2 d\lambda, \\ \mathbf{I}_j^{\text{sup,analytic}} &= \int_{\pi 2^j}^{\infty} \Gamma_j(\mathbf{d})^{-1} \mathbf{\Lambda}(\lambda) \mathbf{\Theta} \mathbf{\Lambda}(\lambda) \Gamma_j(\mathbf{d})^{-1} 4 \left| \widehat{\psi}_h(\lambda) \right|^2 d\lambda.\end{aligned}$$

Following the proof of Proposition 15, the steps of the proof are the following:

1.  $2^{-j(d_\ell+d_m-\beta)} \left| \text{Cov}(\mathbf{W}_{j,k}) - \mathbf{I}_j^{\text{inf}} \right|$  is bounded, up to a constant, by  $L2^{-2j}$ . This result was already obtained in the proof of Proposition 15.
2.  $2^{-j(d_\ell+d_m-\beta)} \left| \mathbf{I}_j^{\text{inf}} - \mathbf{I}_j^{\text{inf,analytic}} \right|$  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 Corollary 16.
3.  $2^{-j(d_\ell+d_m-\beta)} \left| \mathbf{I}_j^{\text{sup,analytic}} \right|$  is bounded. This result is straightforward with step 1.(c) in the proof of Proposition 15, since  $|\widehat{\psi}_h(\lambda)| \leq |\widehat{\psi}(\lambda)|$ .

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

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

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

To this aim, we will use (32) (and Corollary 16) and the inequality

$$|\widehat{\psi}_h(\lambda)| \leq 5^M |\sin(\lambda/4)|^M (1 + |\lambda|)^{-M}, \quad (33)$$

for all  $\lambda \in \mathbb{R} \setminus \{0\}$  (see Section B.1.2).

We decompose the integral on the sub-intervals  $(-2\pi, 2\pi)$ ,  $(-4\pi, -2\pi)$ ,  $(2\pi, 4\pi)$ ,  $(-2^j\pi, -4\pi)$ ,  $(-2^j\pi, -4\pi)$  and  $(4\pi, 2^j\pi)$ .

**On  $(-2\pi, 2\pi)$ .**

With assumption (W2), Corollary 16 leads to

$$\begin{aligned}
& \int_{-2\pi}^{2\pi} |\lambda|^{-d_\ell - d_m} \left| \left| \widehat{\psi}(\lambda) \right|^2 - 4 \mathbb{1}_{\mathbb{R}_+}(\lambda) \left| \widehat{\psi}_h(\lambda) \right|^2 \right| d\lambda \\
& \leq 24\sqrt{2} \int_0^{2\pi} \left( 1 - \frac{|\lambda|}{4\pi} \right)^{2L+1} |\lambda|^{-d_\ell - d_m + 2M} d\lambda, \\
& \leq 24\sqrt{2} (4\pi)^{-d_\ell - d_m + 2M+1} B(2M - d_\ell - d_m + 1, 2L + 2),
\end{aligned}$$

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

$$36\sqrt{2} \Gamma(2M - d_\ell - d_m + 1) (4\pi)^{-d_\ell - d_m + 2M+1} (2L + 2)^{-(2M - d_\ell - d_m + 1)}.$$

This bound is negligible with respect to  $L^{-M-1}$ , since  $d_\ell + d_m < M$  under (C-b).

**On  $(-4\pi, -2\pi)$  and  $(2\pi, 4\pi)$ .**

Observe that (33) yields  $|\widehat{\psi}_h(\lambda)| \leq 5^M (4\pi - \lambda)^M$ . With Corollary 16, we get

$$\begin{aligned}
& \int_{2\pi}^{4\pi} |\lambda|^{-d_\ell - d_m} \left| \left| \widehat{\psi}(\lambda) \right|^2 - 4 \mathbb{1}_{\mathbb{R}_+}(\lambda) \left| \widehat{\psi}_h(\lambda) \right|^2 \right| d\lambda \\
& \leq 12\sqrt{2} 5^{2M} (4\pi) \int_{2\pi}^{4\pi} \left( 1 - \frac{\lambda}{4\pi} \right)^{2L+1} (4\pi - \lambda)^{2M} d\lambda, \\
& \leq 12\sqrt{2} 5^{2M} (4\pi)^{2M+2} B(2M + 1, 2L - 1).
\end{aligned}$$

Using Stirling's approximation, for fixed  $M$  and  $L$  going to infinity, the right-hand side is equivalent, up to a multiplicative constant to  $L^{-2M-1}$ . It is therefore lower than  $L^{-M-1}$  for sufficiently large  $L$ .

A similar result is obtained on  $(-4\pi, -2\pi)$ .

On  $(-2^j\pi, -4\pi)$  and  $(4\pi, 2^j\pi)$ .

Let us first consider the integral on an interval  $(4k\pi, 4k\pi + 2\pi)$ , with  $k \in \mathbb{N}$ ,  $k \geq 2$ . Inequality (33) implies that  $|\widehat{\psi}_h(\lambda)| \leq 5^M \left| \frac{4k\pi - \lambda}{\lambda} \right|^{M/2} |\lambda|^{-M/2}$ . With (32), we get

$$\begin{aligned} & \int_{4k\pi}^{4k\pi+2\pi} |\lambda|^{-d_\ell-d_m} \left| \left| \widehat{\psi}(\lambda) \right|^2 - 4 \mathbb{1}_{\mathbb{R}_+}(\lambda) \left| \widehat{\psi}_h(\lambda) \right|^2 \right| d\lambda \\ & \leq 12\sqrt{2} 5^{2M} \log_2(k) \int_{4k\pi}^{4k\pi+2\pi} |\lambda|^{-d_\ell-d_m} \left( \frac{4k\pi}{\lambda} \right)^{2L+1} \left( 1 - \frac{4k\pi}{\lambda} \right)^M |\lambda|^{-M} d\lambda. \end{aligned}$$

Since  $-M - d_\ell - d_m < 0$ ,  $|\lambda|^{-M-d_\ell-d_m} < (4k\pi)^{-M-d_\ell-d_m}$  for all  $\lambda > 4k\pi$ . With the change of variable  $\mu = \frac{4k\pi}{\lambda}$ , we obtain the upper bound

$$\begin{aligned} & 12\sqrt{2} 5^{2M} \log_2(k) (4k\pi)^{-M-d_\ell-d_m+1} \int_0^1 (\mu)^{2L-1} (1-\mu)^M d\mu, \\ & \leq 12\sqrt{2} 5^{2M} (4k\pi)^{-M-d_\ell-d_m+1} \log_2(k) B(M+1, 2L), \end{aligned} \quad (34)$$

with  $B(.,.)$  the Beta function.

Second, we consider the integral on an interval  $(4k\pi - 2\pi, 4k\pi)$ , with  $k \in \mathbb{N}$ ,  $k \geq 2$ . Inequality (33) implies that  $|\widehat{\psi}_h(\lambda)| \leq 5^M \left| \frac{4k\pi - \lambda}{\lambda} \right|^{M/2} |\lambda|^{-M/2}$ . With (32), we get

$$\begin{aligned} & \int_{4k\pi-2\pi}^{4k\pi} |\lambda|^{-d_\ell-d_m} \left| \left| \widehat{\psi}(\lambda) \right|^2 - 4 \mathbb{1}_{\mathbb{R}_+}(\lambda) \left| \widehat{\psi}_h(\lambda) \right|^2 \right| d\lambda \\ & \leq 12\sqrt{2} 5^{2M} \log_2(k) \int_{4k\pi-2\pi}^{4k\pi} |\lambda|^{-d_\ell-d_m} \left( 1 - \frac{4k\pi - \lambda}{\lambda} \right)^{2L+1} \left( \frac{4k\pi - \lambda}{\lambda} \right)^M |\lambda|^{-M} d\lambda \\ & \leq 12\sqrt{2} 5^{2M} \log_2(k) (4k\pi - 2\pi)^{-M-d_\ell-d_m} (4k\pi) \int_0^1 (1-\mu)^{2L-1} (\mu)^M d\mu, \end{aligned}$$

where we have done the change of variable  $\mu = \frac{4k\pi}{\lambda} - 1$ . We obtain the bound

$$12\sqrt{2} 5^{2M} (4(k-1)\pi)^{-M-d_\ell-d_m+1} \log_2(k) B(M+1, 2L), \quad (35)$$

with  $B(.,.)$  the Beta function.

Next,

$$\begin{aligned}
& \int_{4\pi}^{2^j\pi} |\lambda|^{-d_\ell-d_m} \left| \left| \widehat{\psi}(\lambda) \right|^2 - 4 \mathbb{1}_{\mathbb{R}_+}(\lambda) \left| \widehat{\psi}_h(\lambda) \right|^2 \right| d\lambda \\
& \leq \sum_{k=2}^{2^{j-2}} \int_{4k\pi-2\pi}^{4k\pi} |\lambda|^{-d_\ell-d_m} \left| \left| \widehat{\psi}(\lambda) \right|^2 - 4 \mathbb{1}_{\mathbb{R}_+}(\lambda) \left| \widehat{\psi}_h(\lambda) \right|^2 \right| d\lambda \\
& \quad + \sum_{k=1}^{2^{j-2}} \int_{4k\pi}^{4k\pi+2\pi} |\lambda|^{-d_\ell-d_m} \left| \left| \widehat{\psi}(\lambda) \right|^2 - 4 \mathbb{1}_{\mathbb{R}_+}(\lambda) \left| \widehat{\psi}_h(\lambda) \right|^2 \right| d\lambda. \quad (36)
\end{aligned}$$

The two terms on the right hand side are bounded by (34) and (35). Since  $M \geq 2$  and  $d_\ell + d_m > -1$ ,  $-M - d_\ell - d_m < -1$ . Hence,  $\sum_{k=1}^{2^j} \log_2(k) k^{-M-d_\ell-d_m}$  can be bounded by a constant, not depending on  $j$ . We deduce that, for fixed  $M$ , up to a constant, the right-hand side of (36) is bounded by  $B(2M-2, 2L)$ . Stirling's approximation states that  $B(M+1, 2L)$  is equivalent to  $L^{-M-1}$ , for fixed  $M$  and  $L$  going to infinity.

The bound on  $(-2^j\pi, 4\pi)$  is similar.

## E Asymptotic behavior of the estimators

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

First, recall that, for all  $j \geq 0$ ,  $n_j$  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  $2^{-j_0}N \rightarrow \infty$ , the sequence  $n_j$  is equivalent to  $2^{-j}N$  when  $j$  goes to infinity. These assumptions are made in both Theorem 7 and Theorem 8. Hence,  $n_j$  behaves similarly to in Achard and Gannaz [2016] and in Gannaz [2023].

## E.1 Proof of Theorem 7

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

$$S_{\ell,m}^{(1)}(\mu) = \sum_{j=j_0}^{j_1} n_j \mu_j \left[ \frac{\text{Cov}(W_{j,k}(\ell), W_{j,k}(m))}{2^{j(d_\ell^0 + d_m^0)}} - G_{\ell,m}^0 \right],$$

defined on page 499 of Achard and Gannaz [2016], is modified. Consequently, Proposition 8 and Proposition 9 of Achard and Gannaz [2016] do not hold anymore. They are replaced respectively by Proposition 17 and Proposition 18 below.

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

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

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

$$\begin{aligned} S_{\ell,m}^{(0)}(\mu) &= \sum_{j=j_0}^{j_1} \mu_j \frac{1}{2^{j(d_\ell^0 + d_m^0)}} \sum_k (W_{j,k}(\ell)W_{j,k}(m) - \text{Cov}(W_{j,k}(\ell), W_{j,k}(m))), \\ S_{\ell,m}^{(1)}(\mu) &= \sum_{j=j_0}^{j_1} n_j \mu_j \left[ \frac{\text{Cov}(W_{j,k}(\ell), W_{j,k}(m))}{2^{j(d_\ell^0 + d_m^0)}} - G_{\ell,m}^0 \right]. \end{aligned}$$

**Proposition 17.** *Assume that the sequences  $\mu$  belong to the set  $\{\{\mu_j\}_{j \geq 0}, |\mu_j| \leq \frac{1}{n_j}\}$ . Suppose that (C-a) holds. Under condition (C),  $\sup_{\{\mu, |\mu_j| \leq \frac{1}{n_j}\}} S_{\ell,m}(\mu)$  is uniformly bounded by  $2^{-j_0\beta} + L2^{-2j_0} + j_1 L^{-M-1} + N^{-1/2} 2^{j_1/2}$  up to a multiplicative constant, that is,*

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



*Proof.* From Proposition 6, there exists  $C > 0$  such that

$$|S_{\ell,m}^{(1)}(\mu)| \leq C \sum_{j=j_0}^{j_1} (2^{-\beta j} + L2^{-j} + L^{-M-1}) n_j |\mu_j|. \quad (37)$$

Under the assumption  $|\mu_j| \leq \frac{1}{n_j}$ , we have the inequality  $|S_{\ell,m}^{(1)}(\mu)| \leq C \sum_{j=j_0}^{j_1} (2^{-\beta j} + L2^{-j} + L^{-M-1})$ . The right-hand bound is equivalent to  $2^{-j_0\beta} + L2^{-2j_0} + j_1 L^{-M-1}$  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].  $\square$

**Proposition 18.** *Let  $0 < j_0 \leq j_1 \leq j_N$ . Suppose that (C-a) holds. Assume that the sequences  $\mu$  belong to the set*

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

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

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

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

In particular, for any  $0 \leq \gamma < 1$ , under the assumptions  $2^{-j_0\beta} + N^{-1/2}2^{j_0/2} \rightarrow 0$ , and  $L2^{-2j_0} + \log(N)L^{-M-1} \rightarrow 0$ , 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 (37) that,

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

The sums on the right-hand side of the inequality tend 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].  $\square$

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].

## E.2 Proof of Theorem 8

### Expressions of the asymptotic variances.

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

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

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  $\{\mathbf{W}(j+u, 2^u k + \tau), \tau = 0, \dots, 2^u - 1, k \in \mathbb{Z}\}$ ,

$$\begin{aligned} D_{u, \tau}(\lambda; \delta) &= \sum_{t \in \mathbb{Z}} |\lambda + 2t\pi|^{-\delta} \overline{\widehat{\psi}(\lambda + 2t\pi)} 2^{u/2} \widehat{\psi}(2^u(\lambda + 2t\pi)) e^{-i2^u \tau(\lambda + 2t\pi)}, \\ \tilde{D}_{u, \infty}(\lambda; \delta) &= \sum_{\tau=0}^{2^u-1} D_{u, \tau}(\lambda; \delta). \end{aligned}$$

We introduce

$$\begin{aligned} \mathcal{I}_\Delta^d(\delta_1, \delta_2) &= \frac{2}{\kappa_\Delta} \tilde{I}_0(\delta_1, \delta_2) \\ &+ \frac{2}{\kappa_\Delta^2} \sum_{u=1}^{\Delta} (2^{u\delta_1} + 2^{u\delta_2}) 2^{-u} \frac{2 - 2^{-\Delta+u}}{2 - 2^{-\Delta}} ((u + \eta_{\Delta-u} - \eta_\Delta)(\eta_{\Delta-u} - \eta_\Delta) + \kappa_{\Delta-u}) \tilde{I}_u(\delta_1, \delta_2) \end{aligned}$$

if  $\Delta < \infty$ ,

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

Define also

$$\mathbf{G} \cdot \mathcal{I}^d \cdot \mathbf{G}(\Delta) = \text{Diag}(\text{vec}(\mathbf{G}^0)) (\mathcal{I}_\Delta^d(d_a^0 + d_b^0, d_{a'}^0 + d_{b'}^0)_{(a,b),(a',b') \in \{1, \dots, p^2\}}) \text{Diag}(\text{vec}(\mathbf{G}^0)). \quad (38)$$

Additionally, let us denote

$$\begin{aligned} \mathcal{I}_\Delta^G(\delta_1, \delta_2) &= \tilde{I}_0(\delta_1, \delta_2) + \sum_{u=1}^{\Delta} (2^{u\delta_1} + 2^{u\delta_2}) 2^{-u} \frac{2 - 2^{-\Delta+u}}{2 - 2^{-\Delta}} \tilde{I}_u(\delta_1, \delta_2) && \text{if } \Delta < \infty, \\ \mathcal{I}_\infty^G(\delta_1, \delta_2) &= \tilde{I}_0(\delta_1, \delta_2) + \sum_{u=1}^{\infty} (2^{u\delta_1} + 2^{u\delta_2}) 2^{-u} \tilde{I}_u(\delta_1, \delta_2) && \text{if } \Delta = \infty. \end{aligned}$$

Let us also define

$$\mathbf{G} \cdot \mathcal{I}^G \cdot \mathbf{G}(\Delta) = \text{Diag}(\text{vec}(\mathbf{G}^0)) (\mathcal{I}_\Delta^G(d_a^0 + d_b^0, d_{a'}^0 + d_{b'}^0)_{(a,b),(a',b') \in \{1, \dots, p^2\}}) \text{Diag}(\text{vec}(\mathbf{G}^0)). \quad (39)$$

Let us reformulate Theorem 8 with the exact expression of the asymptotic variance.

**Theorem 19.** *Suppose that conditions of Theorem 7 are satisfied and that assumption (M-3) hold.*

*Let  $j_0 < j_1 \leq j_N$  with  $j_N = \max\{j, n_j \geq 1\}$  such that*

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

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

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

Then,

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

$$\mathbf{V}^{(\mathbf{d})}(\Delta) = \frac{1}{2 \log(2)^2} (\mathbf{G}^{0-1} \circ \mathbf{G}^0 + \mathbf{I}_p)^{-1} \Upsilon(\Delta) (\mathbf{G}^{0-1} \circ \mathbf{G}^0 + \mathbf{I}_p)^{-1}, \quad (40)$$

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

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

where quantities  $\mathbf{G} \cdot \mathcal{I}^{\mathbf{d}} \cdot \mathbf{G}(\Delta)$  are defined by (38).

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

$$V_{(a,b),(a',b')}^{(\mathbf{G})}(\Delta) = \mathbf{G} \cdot \mathcal{I}^{\mathbf{G}} \cdot \mathbf{G}_{(a,a'),(b,b')}(\Delta) + \mathbf{G} \cdot \mathcal{I}^{\mathbf{G}} \cdot \mathbf{G}_{(a,b'),(a',b)}(\Delta) \quad (41)$$

where quantities  $\mathbf{G} \cdot \mathcal{I}^{\mathbf{G}} \cdot \mathbf{G}(\Delta)$  are defined by (39).

### Proof.

The properties of wavelet filters in the proofs of Gannaz [2023] are used through Proposition 31 of Gannaz [2023]. The inequalities (I1) and (I2) in Gannaz [2023] correspond respectively to (28) and (23) of Proposition 12. Inequality (I3) in Gannaz [2023] follows with the proof of Proposition 31 of Gannaz [2023]. The constants in (28) and (23) do not depend of  $L$ , which allows to use these inequalities as in the proofs of Gannaz [2023].

The other property of the wavelets used in the proofs of Gannaz [2023] is the convergence of  $|\widehat{\varphi}(2^{-j}\lambda)|$  to 1 when  $j$  goes to infinity (page 29 of Gannaz [2023]). When  $L2^{-2j}$  goes to zero, Lemma 14 yields

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

which is the desired result.

Finally, the approximation of the sample wavelet covariance is changed. It is sufficient, to use

results of Proposition 6 (instead of Proposition 1 of Gannaz [2023]), to check that, for all  $(\ell, m) \in \{1, \dots, p\}^2$ ,

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

Hence, based on Proposition 6 and Lemma 4, when  $N^{-1}2^{j_1}L \rightarrow 0$ , it is sufficient to have

$$\lim_{j \rightarrow \infty} N^{1/2}2^{-j/2}(L2^{-2j} + L^{-M-1}) = 0,$$

since when  $j$  goes to infinity,  $j_0 \leq j \leq j_1$ ,  $n_j$  is equivalent to  $N2^{-j}$ .

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. 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. 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.
- G. Faÿ, E. Moulines, F. Roueff, and M. S. Taquq. Estimators of long-memory: Fourier versus Wavelets. *The Journal of Econometrics*, 2008. To appear.
- I. Gannaz. Asymptotic normality of wavelet covariances and multivariate wavelet Whittle estimators. *Stochastic Processes and their Applications*, 155:485–534, 2023.
- 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.
- L. Giraitis, P. M. Robinson, and A. Samarov. Rate optimal semiparametric estimation of the memory parameter of the Gaussian time series with long range dependence. *J. Time Ser. Anal.*, 18:49–61, 1997.
- A. Hjørungnes 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.
- S. Kechagias and V. Pipiras. Modeling bivariate long-range dependence with general phase. *Journal of Time Series Analysis*, 41(2):268–292, 2020.
- 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.
- F. S. Nielsen. Local Whittle estimation of multi-variate fractionally integrated processes. *Journal of Time Series Analysis*, 32(3):317–335, 2011.
- M. Ø. Nielsen and P. Frederiksen. Fully modified narrow-band least squares estimation of weak fractional cointegration. *The Econometrics Journal*, 14(1):77–120, 2011.
- M. Ø. Nielsen and P. H. Frederiksen. Finite sample comparison of parametric, semiparametric, and wavelet estimators of fractional integration. *Econometric Reviews*, 24(4):405–443, 2005.
- 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.