Whittle estimation with (quasi-)analytic wavelets

Sophie Achard, Irène Gannaz

Whittle estimation with (quasi-)analytic wavelets

Sophie Achard

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

Irène Gannaz

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

June 2021

Abstract

The notion of long-memory is considered in the case of multivariate time series, not necessarily Gaussian nor stationary. The long-memory characteristics are defined by the long-memory parameters describing the autocorrelation structure of each process and the long-run covariance measuring the coupling between time series. A phase term is present in the model to widen the classes of models. We introduce a representation of the time series by quasi-analytic wavelets for inference in this setting. We first show that the covariance of the wavelet coefficients provides an adequate estimator of the covariance structure of the processes, including the phase term. Consistent estimators are then proposed which is based on a Whittle approximation. Simulations highlight a satisfactory behavior of the estimation on finite samples on some linear time series and on multivariate fractional Brownian motions. An application on a real dataset in neuroscience is displayed, where long-memory and brain connectivity are inferred.

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

1 Introduction

Multivariate processes are often observed nowadays thanks to the recordings of multiple sensors simultaneously. Numerous examples can be cited such as hydrology [Whitcher and Jensen, 2000], finance [Gençay et al., 2001] or neuroscience [Achard and Gannaz, 2016]. When in addition the time series have also the property of long-range dependence, the definition of the model gets complicated

and several definitions can be proposed. A simple model is to fix the long memory parameter for each univariate process with a covariance structure added between the processes [Robinson, 1995a]. However phenomena such as co-integration cannot be modeled using this simple definition. An alternative was proposed by Robinson [2008], Kechagias and Pipiras [2014], where a phase-term is added to the covariance structure.

Following Kechagias and Pipiras [2014], consider a multivariate long-range dependence process $\mathbf{X}(t) = \begin{bmatrix} X_1(t) & \dots & X_p(t) \end{bmatrix}^T$, $t \in \mathbb{Z}$ with long memory parameters $\mathbf{d} = (d_1, d_2, \dots, d_p)$. Let $\mathbf{\Delta}$ denote the difference operator, $(\mathbf{\Delta X})_t = \mathbf{X}_{t+1} - \mathbf{X}_t$. The k-th difference operator, $\mathbf{\Delta}^k$, $k \in \mathbb{N}$, is defined by k recursive applications of $\mathbf{\Delta}$. For any $\mathbf{D} > \mathbf{d} - 1/2$, we suppose that the multivariate process $\mathrm{Diag}\left(\mathbf{\Delta}^{D_\ell}, \ell = 1, \dots, p\right)\mathbf{X}$ is covariance stationary with a spectral density matrix given by

$$\text{(M-1)} \qquad \boldsymbol{f}^{\Delta}(\lambda) = (\mathbf{D}(\lambda)\,\mathbf{\Omega}\,\mathbf{D}(\lambda)) \circ \boldsymbol{f}^{S}(\lambda), \quad \text{for all } \lambda > 0,$$

where $\circ$ denotes the Hadamard product,

$$\mathbf{D}(\lambda) = \mathrm{Diag}\left(|\lambda|^{-d_1^S}, \dots, |\lambda|^{-d_p^S}\right).$$

and $d_m^S = d_m - D_m$ for all m. When the multivariate time series is already second order stationary, $D_m = 0$ for all m, and our definition is equivalent to the one of Kechagias and Pipiras [2014]. Writing the model as (M-1) enables us to consider non-stationary processes. This is particularly adequate for handling multivariate fractional Brownian motion [Coeurjolly et al., 2013].

The function $\boldsymbol{f}^S(\cdot)$ deals with short-range memory. We assume that $\boldsymbol{f}^S(0) = \mathbf{1}$ to ensure identifiability. We also need an assumption on its regularity:

(M-2) There exists $C_f > 0$ and $\beta > 0$ such that

$$\sup_{0<\lambda<\pi} \sup_{\ell,m=1,\dots,N} \frac{\left|f_{\ell,m}^S(\lambda) - 1\right|}{\lambda^\beta} \leqslant C_f \,.$$

The major interest of this model is the introduction of the matrix $\mathbf{\Omega}$. Let the bar above denote the conjugate operator. The matrix $\mathbf{\Omega}$ satisfies $\mathbf{\Omega}^T = \overline{\mathbf{\Omega}}$ since $\boldsymbol{f}^T(\cdot) = \overline{\boldsymbol{f}(\cdot)}$. More generally, let

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

with $(\omega_{\ell,m})_{\ell,m=1,\dots,p}$ real symmetric non-negative semi-definite matrix and $\mathbf{\Phi} = (\phi_{\ell,m})_{\ell,m=1,\dots,p}$ anti-symmetric. By symmetry of $\boldsymbol{f}(\cdot)$, $\boldsymbol{f}(-\lambda) = \boldsymbol{f}(\lambda)^T$,

$$\boldsymbol{f}(\lambda) = (\mathbf{D}(\lambda)\,\mathbf{\Omega}^T\,\mathbf{D}(\lambda)) \circ \boldsymbol{f}^S(\lambda) \,, \quad \text{for all } \lambda < 0,$$

which is not equal to $(\mathbf{D}(\lambda)\,\mathbf{\Omega}\,\mathbf{D}(\lambda)) \circ \boldsymbol{f}^S(\lambda)$ if $\mathbf{\Phi} \neq 0$. Let $\|\mathbf{\Omega}\|$ denote the infinity norm, that is, $\|\mathbf{\Omega}\| = \max_{\ell,m=1,\dots,p} |\Omega_{\ell,m}|$. This is a generalisation of multivariate long range dependence models used in Lobato [1997], Shimotsu [2007], Achard and Gannaz [2016], where the phase term was taken depending on the difference of long memory parameter, that is, $\phi_{\ell,m} = \pi(d_\ell - d_m)/2$.

In the classical case of 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 can be encountered for univariate estimation: 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 Whittle estimation.

In a multivariate setting, estimation procedures are also following these different classes depending on the choice of the model. 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 Whittle estimation. For a fixed phase term, $\phi_{\ell,m} = \pi(d_\ell - d_m)/2$, estimation of both the covariance structure and the long-memory was proposed by Lobato [1999] and Shimotsu [2007], with a Fourier-based Whittle estimation, and by Achard and Gannaz [2016] with a similar procedure based on a 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. Introducing wavelets is motivated by their flexibility for real data applications. It enables in particular to consider non stationary processes thanks to an implicit differentiation. The introduction of a general phase term challenges the choice of the wavelet filters. The particular definition of the generalized spectral density satisfying (M-1) imposes the use of analytic filters. Quasi-analytic wavelet filters are described in Section 2. Main properties of the filters are displayed and an approximation of the covariance of wavelet coefficients is derived. Section 3 recalls the wavelet Whittle estimators and gives their consistency and their convergence rate. Section 4 reports some simulation results, on ARFIMA linear models and on multivariate fractional Brownian motions. Section 5 provides an empirical application on neuroscience data. Proofs are collected in Appendix.

2 Transform of the multivariate process

We first define the filters used to transform the multivariate time series $\{\mathbf{X}(t),\, t \in \mathbb{Z}\}$ for estimation. Throughout the paper, we adopt the convention that large values of the scale index j correspond to coarse scales (low frequencies).

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

$$\widehat{\phi}_h(\omega) = \prod_{j=1}^{\infty} \left[2^{-1/2} \widehat{h}^{(L)}(\mathrm{e}^{\mathrm{i}2^{-j}\omega}) \right],$$
$$\widehat{\psi}_h(\omega) = 2^{-1/2} \widehat{h}^{(H)}(\mathrm{e}^{\mathrm{i}\,\omega/2})\, \widehat{\phi}_h(\omega/2).$$

We define similarly (ϕ_g, ψ_g) the father and the mother wavelets associated with the wavelet filters $g^{(L)}$ and $g^{(H)}$. The complex father and mother wavelets $(\phi(\cdot), \psi(\cdot))$ are then defined by

$$\widehat{\phi}(\omega) = \widehat{\phi}_h + \mathrm{i}\, \widehat{\phi}_g,$$
$$\widehat{\psi}(\omega) = \widehat{\psi}_h + \mathrm{i}\, \widehat{\psi}_g.$$

2.1 Selesnick's common factor filters

We cannot obtain analytic FIR filters. We choose here to consider quasi-analytic filters introduced by Thiran [1971], Selesnick [2002]. The *common-factor* wavelets of Selesnick [2002] are parametrized by a degree L on the analytic property of the derived complex wavelet. We refer to Selesnick [2001, 2002], Achard et al. [2020] for a detailed description of the construction of the wavelets and their properties.

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

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

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

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

and

$$\widehat{h}^{(H)}(\lambda) = \overline{\widehat{h}^{(L)}(\lambda+\pi)} e^{-i\lambda},$$

with $\widehat{q}_{L,M}(\lambda)$ a real polynomial of $(e^{-i\lambda})$ such that $\widehat{q}_{L,M}(0) = 1$. All the same, we define

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

and

$$\widehat{g}^{(H)}(\lambda) = \overline{\widehat{g}^{(L)}(\lambda+\pi)} e^{-i\lambda}.$$

We finally introduce

$$\widehat{\tau}_j(\lambda) = \widehat{h}_j(\lambda) + \mathrm{i}\,\widehat{g}_j(\lambda), \tag{2}$$

where

$$\widehat{h}_j(\lambda) = \widehat{h}^{(H)}(2^{j-1}\lambda)\prod_{\ell=0}^{j-2}\widehat{h}^{(L)}(2^{\ell}\lambda),$$
$$\widehat{g}_j(\lambda) = \widehat{g}^{(H)}(2^{j-1}\lambda)\prod_{\ell=0}^{j-2}\widehat{g}^{(L)}(2^{\ell}\lambda).$$

Quantities $\{\tau_j(s),\, s\in\mathbb{Z}\}$ are defined by $\widehat{\tau}_j(\lambda) = \sum_{s\in\mathbb{Z}}\tau_j(s)\mathrm{e}^{-\mathrm{i}\,s\,\lambda}$.

Common factor wavelets of Selesnick [2001] are introduced with $\widehat{q}_{L,M}$ such that the filters $\widehat{\tau}_j$ satisfy the perfect reconstruction condition. In that case, $\widehat{q}_{L,M}$ is defined as a solution of

$$|\widehat{q}_{L,M}(\lambda)|^2\, s(\lambda) + |\widehat{q}_{L,M}(\lambda+\pi)|^2\, s(\lambda+\pi) = 1\ , \tag{3}$$

where we have set $s(\lambda) = 2^{-M}(1+\cos(\lambda))^M\left|\widehat{d}_L(\lambda)\right|^2$. Achard et al. [2020] proved that the existence of $\widehat{q}_{L,M}$ is acquired.

Another possibility is to consider that $\widehat{q}_{L,M}$ is a constant equal to 1. Perfect reconstruction is not ensured but it is not necessary for estimation procedures. The properties of filters are then easier to establish, since $\widehat{q}_{L,M}$ does not have an explicit expression with perfect reconstruction.

Definition 1 (Common Factor Wavelets (CFW))**.** *Let M, L be strictly positive integers. Let $(\tau_j)_{j\in\mathbb{N}}$ be a family of Common Factor filters defined by equation (2). If filter $\widehat{q_{L,M}}$ satisfies perfect reconstruction condition (3), $(\tau_j)_{j\in\mathbb{N}}$ will be denoted as CFW-PR(M,L) filters. If $\widehat{q}_{L,M}$ is a constant polynomial equal to 1, $(\tau_j)_{j\in\mathbb{N}}$ will be denoted as CFW-C(M,L) filters.*

Recall the main results established in Achard et al. [2020].

For all $\lambda\in\mathbb{R}$, for all $\widehat{q}_{L,M}$ real polynomial of $(e^{-i\lambda})$,

$$\widehat{\tau}_\infty(\lambda) = \widehat{h}_\infty(\lambda) + \mathrm{i}\,\widehat{g}_\infty(\lambda) = \left(1-\mathrm{e}^{\mathrm{i}\eta_L(\lambda)}\right)\widehat{h}_\infty(\lambda)\ , \tag{4}$$

with

$$\alpha_L(\lambda) = 2(-1)^L \operatorname{atan}\left(\tan^{2L+1}(\lambda/4)\right)\ , \tag{5}$$
$$\eta_L(\lambda) = -\alpha_L(\lambda/2+\pi) + \sum_{j=1}^{\infty}\alpha_L(2^{-j-1}\lambda)\ .$$

In equation (5), we use the convention $\operatorname{atan}(\pm\infty) = \pm\pi/2$ so that α_L is well defined on $\mathbb{R}$. We deduce the following result.

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

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

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

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

and, for all $\lambda \in \mathbb{R}$ and $A \subset \mathbb{R}$, $\delta(\lambda, A)$ denotes the distance of λ to A defined by

$$\delta(\lambda, A) = \inf\left\{|\lambda - x|\,,\ x \in A\right\} \ .$$

2.2 Properties of the filters

We introduce the following properties:

(A-a) *Finite support.* For each j, $\{\tau_j(s)\}_{s\in\mathbb{Z}}$ has finite support.

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

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

with C_m possibly depending of L and M.

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

$$\left|2^{-j/2}\widehat{\tau}_j(2^{-j}\lambda)\right| \leqslant \frac{C_s}{(1+|\lambda|)^\alpha} \quad \text{and} \quad |\widehat{\tau}_\infty(\lambda)| \leqslant \frac{C_s}{(1+|\lambda|)^\alpha} \ .$$

with C_s possibly depending of L and M.

(A-d) *Asymptotic behavior.* There exist $\gamma > 0$ and C_a depending of L and M such that for all $j \geqslant 1$,

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

(A-e) *Quasi-analyticity.* There exist $\gamma > 0$ and C_a depending of L and M such that for all $j \geqslant 1$,

$$\sup_{|\lambda|\leqslant\pi}\left|\left|2^{-j/2}\widehat{\tau}_j(2^{-j}\lambda)\right|^2 - 2\mathbb{1}_{\mathbb{R}^+}(\lambda)\left|\widehat{h}_\infty(\lambda)\right|^2\right| \leqslant C_a\,2^{-\gamma\,j}\,|\lambda|^{2M} \ .$$

Notice that if $(\widehat{\tau}_j)_{j\in\mathbb{Z}}$ denotes the real Daubechies' wavelet filters with M vanishing moments, then Assumptions (A-a) to (A-d) are satisfied, with $\gamma = 2$ in (A-d). We refer to Proposition 3 of Moulines et al. [2007].

We shall now discuss about the properties for CFW-PR(M,L) and CFW-C(M,L) filters. First, by construction, the filters $h^{(L)}$ and $h^{(H)}$, $g^{(L)}$ and $g^{(H)}$ have length $M + L + 1$. Then τ_j has finite length and (A-a) holds.

Exact analyticity corresponds to filters $\widehat{\tau}_\infty$ satisfying

$$\sup_{\lambda\in\mathbb{R}} \left|\widehat{\tau}_\infty(\lambda) - (1 + \mathrm{sign}(\lambda))\,\widehat{h}_\infty(\lambda)\right| = 0.$$

CFW-PR(M,L) and CFW-C(M,L) filters are not exactly analytic but only approximately analytic. Indeed, as stated by Selesnick [2001], assumption (A-a) and exact analyticity cannot hold simultaneously. Thus we made a weaker assumption for analyticity, namely assumption (A-e).

Assumptions (A-b) to (A-d) are acquired for CFW-C(M,L) as stated in the following proposition.

Proposition 3. *Let $L \geqslant 1$ and $M \geqslant 1$. Then Assumptions (A-a) to (A-d) are satisfied by CFW-C(M,L) filters with $\alpha = M + 2L + 1$ and $\gamma = 1$.*

Proofs are given in Section 2.2. Notice that the fact that CFW-C(M,L) filters do not enable a perfect reconstruction is not an obstacle for estimation.

Next for CFW-PR(M,L) filters, verifying assumptions is much more challenging. Indeed, no explicit expression of filter $\widehat{q}(\cdot)$ satifying the Bezout equation (3) can be given. For instance, (A-c) cannot be proved for CFW-PR(M,L) filters. Yet we have an estimation of the Sobolev exponent which may justify the assumption. That is, we cannot prove it but it is satisfied numerically, see Section 5.3 of Achard et al. [2020]. The study of assumptions (A-b) to (A-d) is not detailed in the manuscript for CFW-PR(M,L) filters.

The regularity parameter α may depend of L and M, as in Proposition 3. It is reasonable to suppose that for CFW-PR(M,L) filters it is also an increasing function of M and L. Remark *e.g.* that for Daubechies' wavelet, which corresponds to CFW-PR(M,0) filters, α is increasing with M. An exact expression of α for CFW-PR(M,L) filters is utopian when perfect reconstruction conditions holds. It yet can be checked on simulations that indeed the Sobolev exponent increases with respect to L and with respect to M (see Table 1 of Achard et al. [2020]).

2.3 Moments approximations of the wavelet coefficients

Denote $\{\mathbf{W}_{j,k},\, j \geqslant 0, k \in \mathbb{Z}\}$ the wavelet coefficients of the process $\mathbf{X}$ associated to filters $\{\tau_j(\cdot),\, j \in \mathbb{Z}\}$. For given $j \geqslant 0$ and $k \in \mathbb{Z}$, $\mathbf{W}_{j,k}$ is a p-dimensional vector $\mathbf{W}_{jk} = \big(W_{j,k}(1) \quad W_{j,k}(2) \quad \dots \quad W_{j,k}(p)\big)$, where

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

We are interested in the behavior of $\mathrm{Cov}(\mathbf{W}_{jk})$, which can be expressed as

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

We suppose that the filters are regular enough with respect to the spectral density, that is, we introduce the assumption

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

Recall that M and α measure the regularity of the filters and are defined respectively in (A-b), (A-d), (A-c). Parameters $(d_\ell)_{\ell=1,\dots p}$ and β characterize the dependence in the spectral domain (M-1)-(M-2).

The regularity of the filters first allows to obtain the following approximation.

Proposition 4. *Let $\mathbf{X}$ be a p-multivariate long range dependent process with long memory parameters $d_1, \dots, d_p$ with normalized spectral density $f(\cdot)$ satisfying (M-1) with short-range behavior (M-2). Suppose that (A-a), (A-b), (A-c), (A-d) and (C-a) hold. Then we have, for all $j \geqslant 0$, $k \in \mathbb{Z}$,*

$$\left|\mathrm{Cov}(W_{jk}(\ell), W_{jk}(m)) - 2^{j(d_\ell+d_m)}\omega_{\ell,m}\int_{-\infty}^{\infty} |\lambda|^{-d_\ell-d_m}\, \mathrm{e}^{\mathrm{sign}(\lambda)\,\phi_{\ell,m}}\, |\widehat{\tau}_\infty(\lambda)|^2 \, \mathrm{d}\lambda\right| \leqslant C_1 \, \|\mathbf{\Omega}\| \, 2^{j\,(d_\ell+d_m-\beta)}\,,$$

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

This result corresponds to Proposition 2 of Achard and Gannaz [2016]. It holds for real wavelets. For example, Daubechies' wavelets basis satisfy assumptions (A-a), (A-b), (A-c) and (A-d).

Next, the quasi-analytic property of the filter τ_∞ given by (A-e) implies the following proposition.

Proposition 5. *Let* $\mathbf{X}$ *be a p-multivariate long range dependent process with long memory parameters* $d_1, \ldots, d_p$ *with normalized spectral density satisfying (M-1)–(M-2). Suppose that (A-a), (A-b), (A-c), (A-d), (A-e) and (C-a) hold. Then we have, for all* $j \geqslant 0$, $k \in \mathbb{Z}$,

$$\left| \mathrm{Cov}(W_{jk}(\ell), W_{jk}(m)) - 4\, 2^{j(d_\ell + d_m)} \Omega_{\ell,m} \int_0^\infty |\lambda|^{-d_\ell - d_m} \left| \widehat{h}_\infty(\lambda) \right|^2 \mathrm{d}\lambda \right| \leqslant C_2 \, \|\mathbf{\Omega}\| \, 2^{j\,(d_\ell + d_m - \beta)} \,, \quad (6)$$

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

Inequality (6) can be written as follows: for all $\ell, m = 1, \ldots, p$,

$$\left| \mathrm{Cov}(W_{jk}(\ell), W_{jk}(m)) - 2^{j(d_\ell + d_m)} \, \Omega_{\ell,m} \, K(d_\ell + d_m) \right| \leqslant C_2 \, \|\mathbf{\Omega}\| \, 2^{j\,(d_\ell + d_m - \beta)} \,, \quad (7)$$

with

$$K(\delta) = 4 \int_0^\infty |\lambda|^{-\delta} \left| \widehat{h}_\infty(\lambda) \right|^2 \mathrm{d}\lambda \,.$$

With CFW-PR(L,M) filters, (A-e) may not hold and the control of quasi-analyticity is given by Theorem 2. But it leads to a similar result:

Proposition 6. *Let* $\mathbf{X}$ *be a p-multivariate long range dependent process with long memory parameters* $d_1, \ldots, d_p$ *with normalized spectral density* $f^S(\cdot)$ *satisfying (M-1)–(M-2). Suppose that we consider CFW-C(M,L) filters and that*

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

Then for all $j \geqslant 0$ *there exists* $L_j \geqslant 1$ *such that for all* $L \geqslant L_j$, *for all* $k \in \mathbb{Z}$, *inequality (7) holds, with* C_2 *constant only depending on* $M, C_f, \beta, \|\mathbf{\Omega}\|, \{d_\ell, \ell = 1, \ldots, p\}$.

Assumption (C-b) corresponds to Assumption (C-a) with $\alpha = M/2 + 3/2$ and $\gamma = 1$. It is stronger than the assumption on parameters needed in Proposition 5 where we could consider $\alpha = M + 2L + 1$. This is due to technical conditions in the proof, and also to the fact that we choose not to write assumptions with a dependence on the scales j. Yet, notice that (C-b) does not seem restrictive.

Result of Proposition 6 justifies the use of (quasi-)analytic wavelets. It should be compared to Proposition 3 of Achard and Gannaz [2016]. With real wavelets, the phase term appears in the approximation of the covariance as a cosine term. That is, we would obtain in this framework an approximation of the form

$$\left| \mathrm{Cov}(W_{jk}(\ell), W_{jk}(m)) - 2^{j(d_\ell + d_m)} \, \omega_{\ell,m} \cos(\phi_{\ell,m}) \, K(d_\ell + d_m) \right| \leqslant C \, \|\mathbf{\Omega}\| \, 2^{j\,(d_\ell + d_m - \beta)}.$$

Consequently, parameters $\{\omega_{\ell,m}, \phi_{\ell,m}\}$ are not identifiable. Estimation is derived in Achard and Gannaz [2016] in the case of a parametric phase, $\phi_{\ell,m} = \frac{\pi}{2}(d_\ell - d_m)$. Common factor wavelets, as stated by Proposition 6, have the ability of recovering simultaneously the magnitude and the phase.

3 Estimation

Let $(\phi(\cdot), \psi(\cdot))$ be respectively a father and a mother wavelets defined by respectively by filters $h^{(L)} + \mathrm{i}\, g^{(L)}$ and $h^{(H)} + \mathrm{i}\, g^{(H)}$. At a given resolution $j \geqslant 0$, for $k \in \mathbb{Z}$, we define the dilated and translated functions $\phi_{j,k}(\cdot) = 2^{-j/2}\phi(2^{-j}\cdot - k)$ and $\psi_{j,k}(\cdot) = 2^{-j/2}\psi(2^{-j}\cdot - k)$. The Fourier transform is $\widehat{\tau}_j(\lambda) = \int_{-\infty}^{\infty} \psi(t) e^{-i\lambda t} dt$ with filter $\widehat{\tau}_j(\cdot)$ defined in Section 2. CFW-PR(M,L) and CFW-C(M,L) filters will be considered.

In practice, the process $\mathbf{X}$, is observed on discrete time points $\mathbf{X}(1), \ldots \mathbf{X}(N)$. Let $\widetilde{\mathbf{X}}(t) = \sum_{s=1}^{N} \mathbf{X}(s)\phi(t-s)$. The empirical wavelet coefficients of the process $\mathbf{X}$ are defined by

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

Equivalently,

$$\mathbf{W}_{j,k} = \sum_{s=1}^{N} \mathbf{X}(s)\tau_j(2^j k - s) \quad j \geqslant 0, k \in \mathbb{Z}.$$

For given $j \geqslant 0$ and $k \in \mathbb{Z}$, $\mathbf{W}_{j,k}$ is a p-dimensional vector $\mathbf{W}_{jk} = \big(W_{j,k}(1) \quad W_{j,k}(2) \quad \ldots \quad W_{j,k}(p)\big)$, where $W_{j,k}(\ell) = \int_{\mathbb{R}} \widetilde{X_\ell}(t)\psi_{j,k}(t)dt$.

Since the wavelets have a compact support only a finite number n_j of coefficients are non null at each scale j. Suppose without loss of generality that the support of $\psi(\cdot)$ is included in $[0, T_\psi]$ with $T_\psi \geqslant 1$. For every $j \geqslant 0$, define

$$n_j := \max\left(0, 2^{-j}(N - T_\psi + 1)\right).$$

Then for every $k < 0$ and $k > n_j$, the coefficients $\mathbf{W}_{j,k}$ are set to zero because all the observations are not available. In the following, $n = \sum_{j=j_0}^{j_1} n_j$ denotes the total number of non-zero coefficients used for estimation.

Let $j_U \geqslant j_L \geqslant 1$ be respectively the upper and the lower resolution levels that are used in the estimation procedure. The estimation is based on the vectors of wavelets coefficients $\{\mathbf{W}_{j,k},\, j_L \leqslant j \leqslant j_U,\, k \in \mathbb{Z}\}$. Let n_j be the number of non-zero coefficients $\{\mathbf{W}_{j,k},\, k \in \mathbb{Z}\}$ for a given scale $j \geqslant 0$. The n_js are finite since the filter is finite (assumption (A-a)).

3.1 Estimation procedure

The estimation procedure is similar than the one developped in Achard and Gannaz [2016]. The wavelet Whittle approximation of the negative log-likelihood is denoted by $\mathcal{L}(\cdot)$. It is based on

approximation (7). The wavelet Whittle criterion is defined as

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

with $\mathbf{G}$ the matrix with elements $G_{\ell,m} = \Omega_{\ell,m}K(d_\ell + d_m)$, $1 \leqslant, \ell, m \leqslant p$. The exponent H denotes the conjugate transpose operator. We have

$$\mathcal{L}(\mathbf{G},\mathbf{d}) = \frac{1}{n}\sum_{j=j_L}^{j_U}\left[n_j \log\det\left(\mathbf{\Gamma}_j(\mathbf{d})\,\mathbf{G}\,\mathbf{\Gamma}_j(\mathbf{d})\right) + \mathrm{trace}\left(\left(\mathbf{\Gamma}_j(\mathbf{d})\,\mathbf{G}\,\mathbf{\Gamma}_j(\mathbf{d})\right)^{-1}\mathbf{I}(j)\right)\right], \tag{8}$$

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

Remark that when $\mathbf{G}$ is an Hermitian positive definite matrix, for all $j \geqslant 0$, for all $\mathbf{d}$, $\det(\mathbf{\Gamma}_j(\mathbf{d})\,\mathbf{G}\,\mathbf{\Gamma}_j(\mathbf{d}))$ is real and strictly positive and trace $\left((\mathbf{\Gamma}_j(\mathbf{d})\,\mathbf{G}\,\mathbf{\Gamma}_j(\mathbf{d}))^{-1}\mathbf{I}_j\right)$ is real. The criterion $\mathcal{L}(\mathbf{G},\mathbf{d})$ is hence well-defined for $\mathbf{G}$ in the set of Hermitian matrices and for all $\mathbf{d} \in \mathbb{R}^p$, and it takes its values in $\mathbb{R}$.

Differentiating expression (8) with respect to the matrix $\mathbf{G}$ yields to

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

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

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

Remark that Shimotsu [2007] and Baek et al. [2020], in the Fourier based approach, consider a real matrix $\mathbf{G}(\mathbf{d})$ and complex valued matrices $\mathbf{\Gamma}_j(\mathbf{d})$, including the phases $\left(\phi_{\ell,m}\right)_{\ell,m=,,,\ldots,p}$. They are estimated in a second step, together with parameter $\mathbf{d}$, minimizing the criterion obtained when replacing $\mathbf{G}$ by $\widehat{\mathbf{G}}(\mathbf{d})$ in (8). In Shimotsu [2007] the resulting criterion only depends of $\mathbf{d}$ since the phases are parametric while Baek et al. [2020] deal with a general form of phases. A similar approach can be derived in our context. Yet, our procedure enables to estimate the magnitude of the correlation even when the phase if equal to $\pi/2$, leading to imaginary terms in $\mathbf{G}$.

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

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

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

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

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

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

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

3.2 Asymptotic behaviour

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

Condition (C)

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

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

Theorem 7. *Suppose that assumptions of Proposition 5 or of Proposition 6 hold. Assume that Condition (C) is satisfied. If j_0 and j_1 are chosen such that $2^{-j_0\beta} + N^{-1/2}2^{j_0/2} \to 0$ and $j_0 < j_1 \leqslant j_N$ with $j_N = \max\{j, n_j \geqslant 1\}$, then*

$$\widehat{\mathbf{d}} - \mathbf{d}^0 = o_{\mathbb{P}}(1).$$

If j_0 and j_1 are chosen such that $\log(N)^2(2^{-j_0\beta} + N^{-1/2}2^{j_0/2}) \to 0$ and $j_0 < j_1 \leqslant j_N$ then

$$\widehat{\mathbf{d}} - \mathbf{d}^0 = O_{\mathbb{P}}(2^{-j_0\beta} + N^{-1/2}2^{j_0/2}),$$

$$\forall (\ell, m) \in \{1, \dots, p\}^2,\ \widehat{G}_{\ell,m}(\widehat{\mathbf{d}}) - G_{\ell,m}(\mathbf{d}^0) = O_{\mathbb{P}}(\log(N)(2^{-j_0\beta} + N^{-1/2}2^{j_0/2})),$$
$$\widehat{\Omega}_{\ell,m} - \Omega^0_{\ell,m} = O_{\mathbb{P}}(\log(N)(2^{-j_0\beta} + N^{-1/2}2^{j_0/2})).$$

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

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

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

4 Simulation study

In this section, we verify the accuracy of the covariance approximation given in Proposition 6 and the consistency of the parameters estimates provide 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-PR(4,4) filters. We also compare to quality of estimation of parameters $\mathbf{d}$ to the one given by real wavelets, namely Daubechies' wavelets with 4 vanishing moments.

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).

4.1 ARFIMA models

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

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

Define the process $\mathbf{X}$ as

$$\mathbf{A}(\mathbb{L})\,\text{Diag}\left((\mathbb{1}-\mathbb{L})^{\mathbf{d}}\right)\,\mathbf{X}_t = \mathbf{B}(\mathbb{L})\xi_t. \tag{9}$$

The spectral density satisfies

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

In particular

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

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

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

We simulate $\{\mathbf{X}(1), \dots \mathbf{X}(N)\}$ in (9) with $N = 2^{12}$, null $\mathbf{A}_k$ and $\mathbf{B}_k$ for $k \geqslant 0$. That is, there is no short-range terms in the model. We consider three sets of values for $\mathbf{d}$, $\mathbf{d} \in \{(0.2, 0.2),\ (0.2, 0.4),\ (0.2, 0.8)\}$. Matrix $\mathbf{\Sigma}$ was taken 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].

Figure 1 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 accurate, especially for the high scales (lowest frequencies), even if it has not been established theoretically for such filters. A similar figure can be obtained with CFW-C(4,4) filter. It is not displayed here since no difference with the CFW-C(4,4) filter can be observed. Remark that 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.

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. They show that the estimation is satisfactory and that its quality is similar to the Daubechies' real wavelet-based estimation. The quality with complex filters procedure increases with respect to real filter procedure when the difference between the long memory parameters d_1 and d_2 increases. This can be explained by the fact that we take into account the imaginary part of the matrix $\mathbf{G}$ in estimation, which is not the case with real filters. Notice that CFW-PR(4,4) and CFW-C(4,4) filters give the same results.

Table 3 and Table 4 give the results for the covariance structure estimation, respectively for CFW-PR(4,4) filter and for CFW-C(4,4) filter. Very similar results are obtained by CFW-PR(4,4) and CFW-C(4,4) filters. Estimation is satisfactory for the covariance structure $|\mathbf{\Omega}|$ and for the correlation ρ. Results for the phase parameter ϕ are less satisfactory. A bias term, increasing as the phase ϕ increases, can be observed. It has for example an order of $\pi/10$ when estimating the phase of $3\pi/10$ corresponding to the case $\mathbf{d} = (0.2, 0.8)$. Yet as illustrated *e.g.* in the simulation study of Baek et al. [2020], estimating the phase is challenging and our result seems comparable to Baek et al. [2020]'s Fourier-based Whittle estimator.

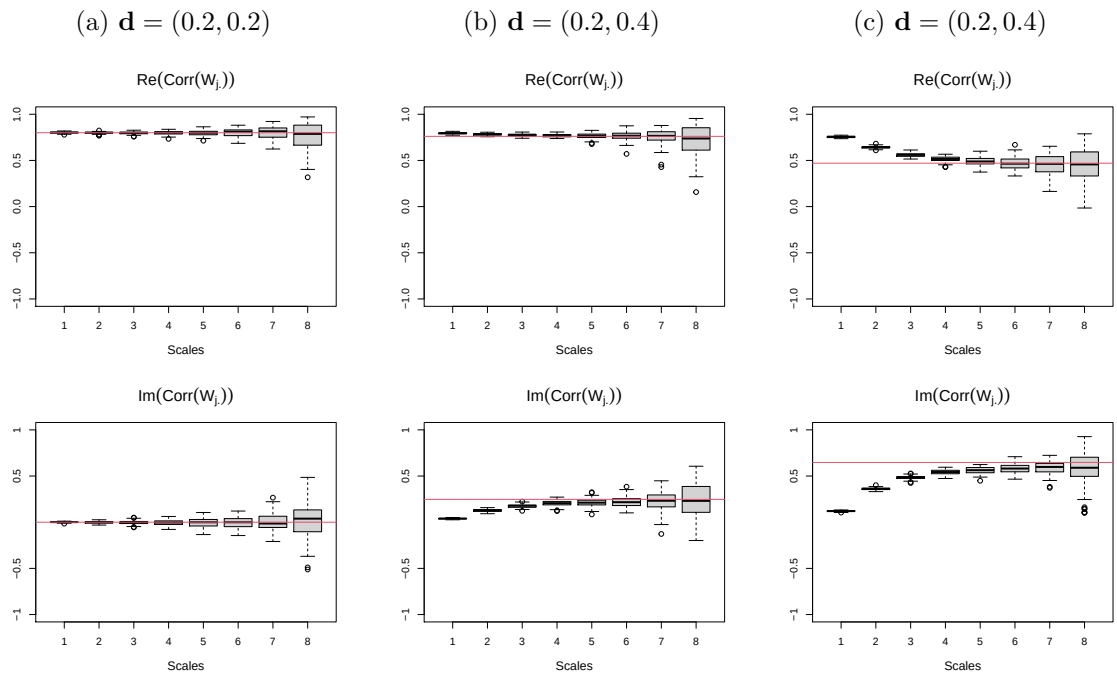

Figure 1: Boxplots of correlation between CFW-PR(4,4) coefficients at different scales for ARFIMA processes. First row gives the real part of the correlations and second row gives the imaginary part. Each column corresponds to a given value of parameter $\mathbf{d}$. Horizontal red lines correspond to the approximation given by Proposition 6, that is, $\rho\cos(\phi)$ for the real part and $\rho\sin(\phi)$ for the imaginary part.

$\mathbf{d}$	bias	std	RMSE	CFW-PR(4,4)/Real
0.2	0.0130	0.0110	0.0170	1.0097
0.2	0.0131	0.0109	0.0170	1.0057
0.2	0.0172	0.0111	0.0205	1.2854
0.4	-0.0041	0.0111	0.0118	0.7127
0.2	0.0203	0.0173	0.0266	0.8769
0.8	4e-04	0.0174	0.0174	0.7937

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

d	bias	std	RMSE	CFW-C(4,4)/CFW-PR(4,4)
0.2	0.013	0.0110	0.017	1
0.2	0.0131	0.0109	0.017	1
0.2	0.0172	0.0111	0.0205	1
0.4	-0.0041	0.0111	0.0118	1
0.2	0.0203	0.0173	0.0266	1
0.8	4e-04	0.0174	0.0174	1

Table 2: Results for the estimation of long-memory parameters $\mathbf{d}$ with CFW-C(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 $\mathbf{d} \in \{(0.2, 0.8)\}$. CFW-C(4,4)/CFW-PR(4,4) denotes the ratio between the RMSE given by CFW-C(4,4) filter and the RMSE given by CFW-PR(4,4) filter.

4.2 Multivariate Brownian motions

We now consider a multivariate fractional Brownian motion (mFBM). 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 Proposition 11 of Amblard and Coeurjolly [2011].

The p-multivariate fractional Brownian motion (mFBM) $(\mathbf{X}(t))_{t\in\mathbb{R}}$ 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 $a > 0$, $(\mathbf{X}_1(at), \ldots, \mathbf{X}_p(at))$ has the same distribution as $(a^{d_1-1/2}\mathbf{X}_1(t), \ldots, a^{d_p-1/2}\mathbf{X}_p(t))$;
- the increments are stationary.

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

d		bias	std	RMSE
(0.2 , 0.2)	$\lvert\Omega_{1,1}\rvert$	0.005	0.0232	0.0238
	$\lvert\Omega_{1,2}\rvert$	0.0044	0.0208	0.0213
	$\lvert\Omega_{2,2}\rvert$	0.0056	0.0233	0.024
	ρ	1e-04	0.0058	0.0058
	ϕ	-8e-04	0.0138	0.0139
(0.2 , 0.4)	$\lvert\Omega_{1,1}\rvert$	0.007	0.0233	0.0243
	$\lvert\Omega_{1,2}\rvert$	0.0246	0.0214	0.0326
	$\lvert\Omega_{2,2}\rvert$	0.074	0.025	0.0781
	ρ	-0.0071	0.0058	0.0092
	ϕ	0.1117	0.0143	0.1126
(0.2 , 0.8)	$\lvert\Omega_{1,1}\rvert$	-0.0072	0.0343	0.0351
	$\lvert\Omega_{1,2}\rvert$	-0.0175	0.0303	0.035
	$\lvert\Omega_{2,2}\rvert$	0.1114	0.0415	0.1189
	ρ	-0.0551	0.0083	0.0557
	ϕ	0.3063	0.0173	0.3068

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

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

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

where $\mathrm{Cor}(X_1, X_2)$ denotes the Pearson correlation between variables X_1 and X_2. The quantities $(\eta_{\ell,m})_{\ell,m=1,\dots,p}$ measure the disymmetry of the process. The 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 it is equivalent for zero-mean multivariate Gaussian stationary processes $\mathbf{X}$ 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 (ℓ, m).

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

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

d		bias	std	RMSE	CFW-C(4,4)/CFW-PR(4,4)
(0.2 , 0.2)	$\lvert\Omega_{1,1}\rvert$	0.005	0.0248	0.0253	1.0641
	$\lvert\Omega_{1,2}\rvert$	0.0044	0.0222	0.0226	1.0646
	$\lvert\Omega_{2,2}\rvert$	0.0056	0.0248	0.0254	1.0618
	ρ	4e-04	0.0058	0.0058	1.002
	ϕ	-8e-04	0.0138	0.0139	1.0000
(0.2 , 0.4)	$\lvert\Omega_{1,1}\rvert$	0.007	0.025	0.026	1.0673
	$\lvert\Omega_{1,2}\rvert$	0.0246	0.0231	0.0337	1.0342
	$\lvert\Omega_{2,2}\rvert$	0.074	0.0271	0.0788	1.0091
	ρ	-0.0065	0.0058	0.0088	0.9519
	ϕ	0.1117	0.0143	0.1126	1.0000
(0.2, 0.8)	$\lvert\Omega_{1,1}\rvert$	0.0368	0.0357	0.0513	0.9979
	$\lvert\Omega_{1,2}\rvert$	0.0172	0.0277	0.0326	0.9063
	$\lvert\Omega_{2,2}\rvert$	0.1606	0.0438	0.1665	1.0007
	ρ	-0.0551	0.0083	0.0557	1.0000
	ϕ	0.3063	0.0173	0.3068	1.0000

Table 4: Results for the estimation of matrices $\mathbf{\Omega}$ with CFW-C(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 $\mathbf{d} \in \{(0.2, 0.8)\}$. CFW-C(4,4)/CFW-PR(4,4) denotes the ratio between the RMSE given by CFW-C(4,4) filter and the RMSE given by CFW-PR(4,4) filter.

with

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

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

Let $\mathbf{G}$ be given by: $\mathbf{G} = (R_{\ell,m} e^{i\phi_{\ell,m}})_{\ell,m=1,\dots,p}$. When λ tends to 0, the spectral density $f_{\ell,m}^{(1,1)}(\lambda)$ is equivalent to $G_{\ell,m}|\lambda|^{-(d_\ell + d_m - 2)}$. 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.

Remark 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 $\{d_\ell, G_{\ell,m}, \ell, m = 1, \dots, p\}$.

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

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

The phase $\phi_{1,2}$ is approximately equal to $\pi/7$ and $R \simeq \begin{pmatrix} 0.318 & 0.223 \\ 0.223 & 0.320 \end{pmatrix}$, 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 $R \simeq \begin{pmatrix} 0.318 & 0.093 \\ 0.093 & 0.319 \end{pmatrix}$, 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.

We first check that the quality of the approximation of the wavelet correlation resulting from Proposition 6 is satisfactory. Figure 2 represents the boxplots of CFW-PR(4,4) wavelet correlations at different scales in Case 1 and in Case 2. In both settings, it can be seen that the approximation holds, except for highest frequencies. As identical observations were done for CFW-C(4,4) filters the figure is not displayed here.

We now consider the Whittle estimation of the parameters. Table 5 and Table 6 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, estimation procedures are equivalent for both common-factor wavelets.

	$\mathbf{d}$	bias	std	RMSE	ratio CFW-PR(4,4)/Real
Case 1	1	-0.0125	0.0182	0.0221	0.7254
	1.2	-8e-04	0.0198	0.0198	0.8333
Case 2	1	-0.0069	0.0198	0.0210	0.6863
	1.2	0.0047	0.0219	0.0224	0.9127

Table 5: Results for the estimation of long-memory parameters $\mathbf{d}$ with CFW-PR(4,4) filter on mFBMs. Hyperparameter j_0 satisfies $j_0 = 2$. CFW-PR(4,4)/Real denotes the ratio between the RMSE given by CFW-C(4,4) filter and the RMSE given by Daubechies' real filter.

Table 7 and Table 8 give the results obtained for the estimation of the covariance structure, that is, $|\mathbf{\Omega}|$, ρ and ϕ. Estimation of $|\mathbf{\Omega}|$ by CFW-PR(4,4) filters is slightly more accurate than estimation

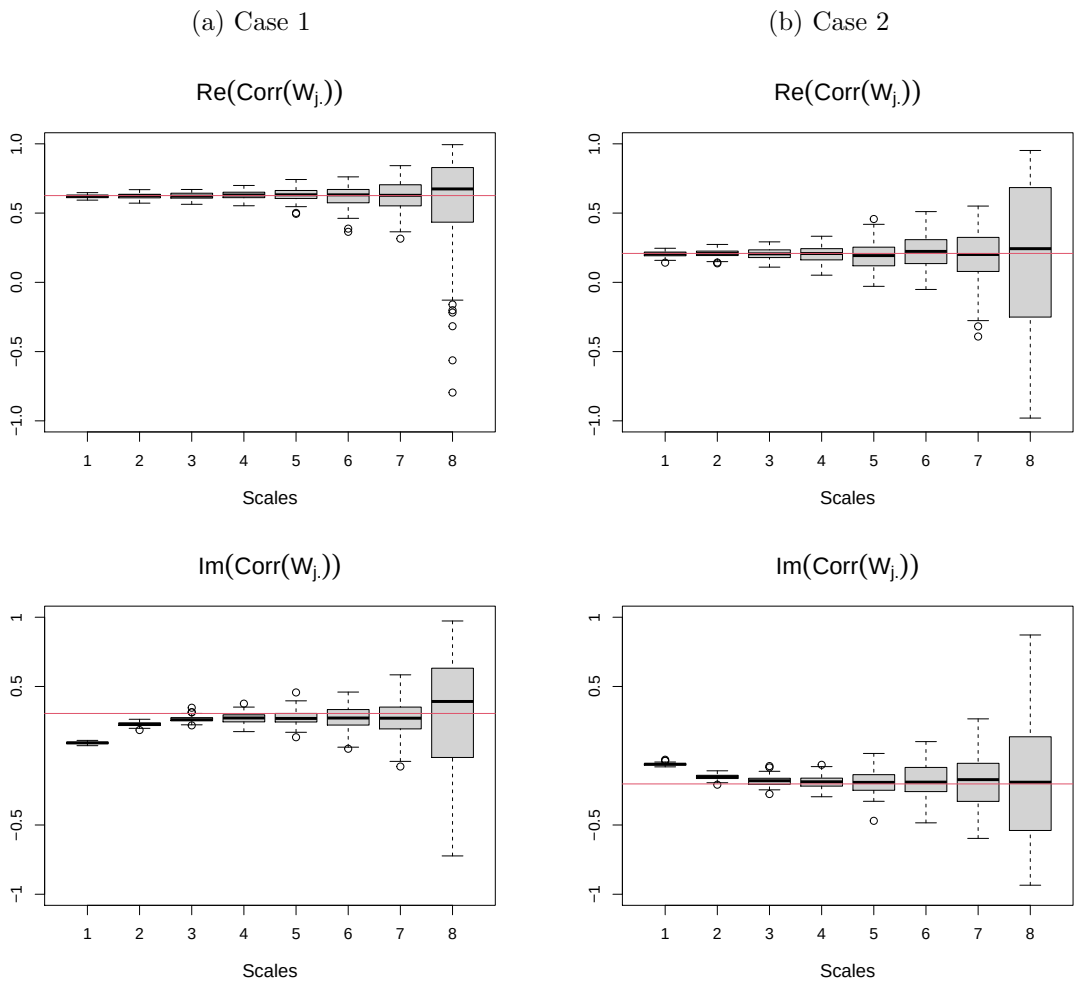

Figure 2: Boxplots of correlation between CFW-PR(4,4) coefficients at different scales for the simulated mFBM in Case 1 (left column–(a)) and in Case 2 (right column–(b)). First row gives the real part of the correlations and second row gives the imaginary part. Horizontal red lines correspond to the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

with CFW-C(4,4) filters, but qualities for the estimation of ρ and of ϕ are equivalent. Bias as standard deviations of the estimations of ρ and ϕ seem reasonable.

	d	bias	std	RMSE	ratio CFW-C(4,4)/CFW-PR(4,4)
Case 1	1	-0.0125	0.0182	0.0221	1
	1.2	-8e-04	0.0198	0.0198	1
Case 2	1	-0.0069	0.0198	0.0210	1
	1.2	0.0047	0.0219	0.0224	1

Table 6: Results for the estimation of long-memory parameters $\mathbf{d}$ with CFW-C(4,4) filter on mFBMs. Hyperparameter j_0 satisfies $j_0 = 2$. CFW-C(4,4)/CFW-PR(4,4) denotes the ratio between the RMSE given by CFW-C(4,4) filter and the RMSE given by CFW-PR(4,4) filter.

		bias	std	RMSE
Case 1	$\lvert\Omega_{1,1}\rvert$	0.1617	0.0453	0.1679
	$\lvert\Omega_{1,2}\rvert$	0.0753	0.0357	0.0833
	$\lvert\Omega_{2,2}\rvert$	0.1370	0.0461	0.1446
	ρ	-0.0252	0.0126	0.0282
	ϕ	0.0743	0.0222	0.0776
Case 2	$\lvert\Omega_{1,1}\rvert$	0.1554	0.0450	0.1618
	$\lvert\Omega_{1,2}\rvert$	0.0134	0.0264	0.0296
	$\lvert\Omega_{2,2}\rvert$	0.1297	0.0469	0.1379
	ρ	-0.0246	0.0201	0.0318
	ϕ	-0.0972	0.0802	0.1260

Table 7: Results for the estimation of matrices $\mathbf{\Omega}$ with CFW-PR(4,4) filter on mFBMs. Hyperparameter j_0 satisfies $j_0 = 2$.

5 Application on a neuroscience dataset

We applied our framework on fMRI data acquired on rats. We use functional Magnetic Resonance images (fMRI) of both dead and alive rats. The aim is to estimate the brain connectivity, that is, the significant correlations between brain regions where fMRI signals are recorded. For this data set, we know that for dead rats we are under the full null hypothesis as no legitimate functional activity should be detected. Thus the estimated graphs should be empty. We also expect non-empty graphs for alive rats under anesthetic, as brain activity keeps on during anesthetic. The dataset are freely available at https://10.5281/zenodo.2452871 [Becq et al., 2020a,b].

		bias	std	RMSE	CFW-C(4,4)/CFW-PR(4,4)
Case 1	$\lvert\Omega_{1,1}\rvert$	0.1617	0.0511	0.1696	1.0097
	$\lvert\Omega_{1,2}\rvert$	0.0753	0.0395	0.0850	1.0206
	$\lvert\Omega_{2,2}\rvert$	0.1370	0.0523	0.1467	1.0145
	ρ	-0.0252	0.0126	0.0282	1.0000
	ϕ	0.0743	0.0222	0.0776	1.0000
Case 2	$\lvert\Omega_{1,1}\rvert$	0.1554	0.0504	0.1633	1.0098
	$\lvert\Omega_{1,2}\rvert$	0.0134	0.0326	0.0352	1.1885
	$\lvert\Omega_{2,2}\rvert$	0.1297	0.0531	0.1401	1.0161
	ρ	-0.0246	0.0201	0.0318	1.0000
	ϕ	-0.0972	0.0802	0.1260	1.0000

Table 8: Results for the estimation of matrices $\mathbf{\Omega}$ with CFW-C(4,4) filter on mFBMs. Hyperparameter j_0 satisfies $j_0 = 2$. CFW-C(4,4)/CFW-PR(4,4) denotes the ratio between the RMSE given by CFW-C(4,4) filter and the RMSE given by CFW-PR(4,4) filter.

5.1 Description of the dataset

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

For each rat, we compute the estimators of

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

Estimation was done with CFW-PR(4,4) filters.

5.2 Results and group comparisons

Figure 3 shows the empirical distribution of the estimated empirical estimators $\widehat{d}$. As expected, the dead rats do not present significant long-memories. The distributions are centered at zero, with a Gaussian-like shape. Rats under anesthetics are not centered at zero and the variance between brain regions are higher than what is observed for dead rats. Long-memories for rats under anesthetic ISO_W are higher than under other anesthetics.

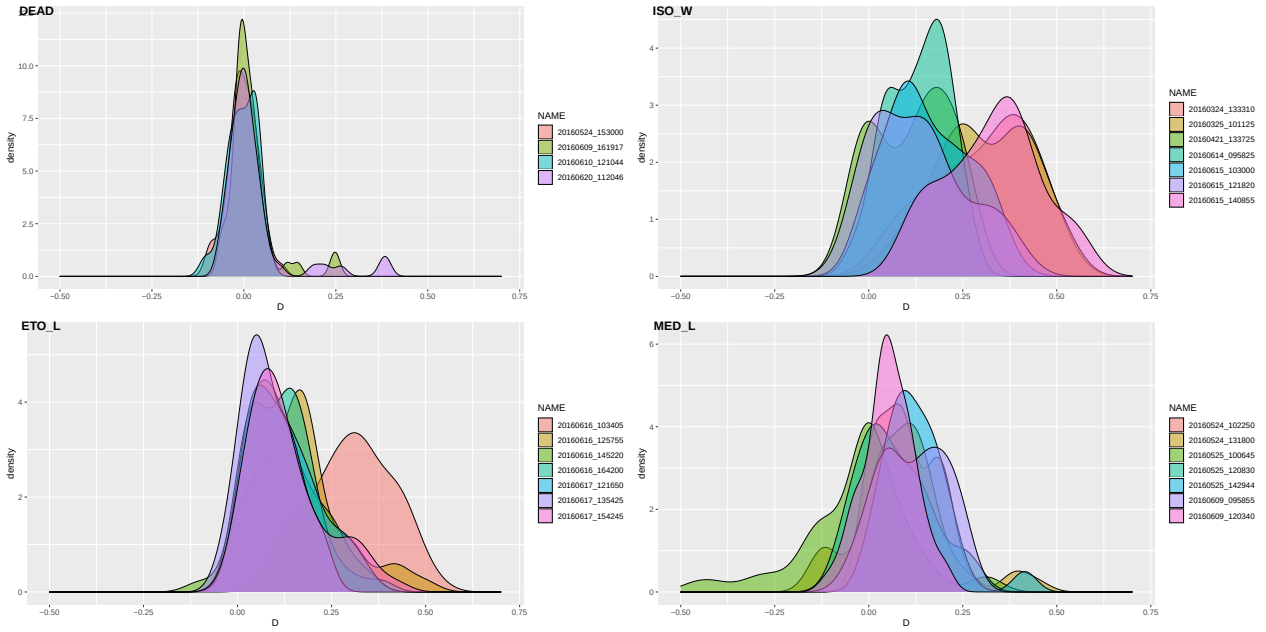

Figure 3: Plot of the empirical distribution of the long memory parameters $\widehat{\mathbf{d}}$ obtained for the 4 groups of rats. Each color corresponds to a rat.

Distributions of magnitude and phases of the estimated correlations, ρ and ϕ, for each rats are displayed respectively in Figure 4 and Figure 5. First, as expected, the magnitudes obtained for the dead rats are significantly different from alive rats. For dead rats, distributions have a small support, that is, only 1.88% of the values satisfy $\widehat{\rho} > 0.3$. Remark also that no significant differences are observed between rats. Next, ISO_W and ETO_L present quite similar distributions, with possibly high magnitudes. By contrast, correlations for MED_L anesthetic are lower. These results tend to show that MED_L anesthetic is stronger than the other anesthetics, leading to less connections between brain regions.

The phase parameter can be interpreted as a shift in the connections between brain regions. Distributions displayed in Figure 5 are the empirical densities of the upper triangular matrices of phases, $\{\phi_{\ell,m}, 1 \leqslant \ell < m \leqslant p\}$. This explains why the distributions are not symmetric.

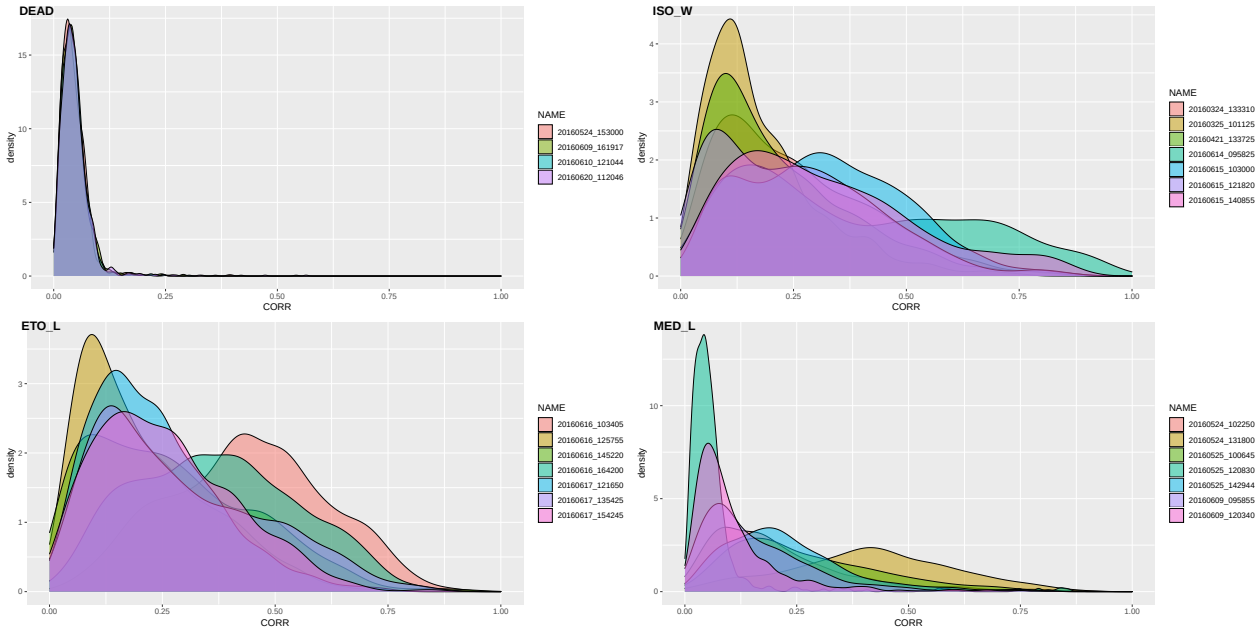

Figure 4: Plot of the empirical distribution of the correlation magnitudes $\widehat{\rho}$ obtained for the 4 groups of rats. Each color corresponds to a rat.

By first approximation, the distributions can be seen as a mixture of a centered Gaussian density and a uniform density on $(-\pi, \pi)$. For dead rats, we observe mainly uniform distributions, while the Gaussian-type distributions dominate for the alive rats. The uniform component is higher 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. To illustrate this fact, Figure 6 shows the distributions of the estimated phases ϕ corresponding to magnitudes satisfying $\rho > 0.3$ (this choice is motivated by the observation on the support of dead rats' correlations above). Only Gaussian-type distributions remain. It can be observed that the support of the phases are larger for alive rats than for dead rats. This shows that the phase brings information on the functional connectivity. Next, the 95%-quantiles of absolute values (*i.e.* q such that 95% of absolute values of phases are lower than q) are respectively 2.9, 2.00, 1.50, 1.53 for Dead rats, ISO_W, ETO_L and MED_L. It seems that ISO_W has a higher support, meaning that shifts appear in the connections between brain region, with respect to other anesthetics. Yet, we have not tested whether the difference is significant.

5.3 Graphs with correlations and phases

Taking into account the rats scanned within each group, we compute a graph to represent the connectivity graph of the group. We first compute the adjacency matrix obtained for each rat

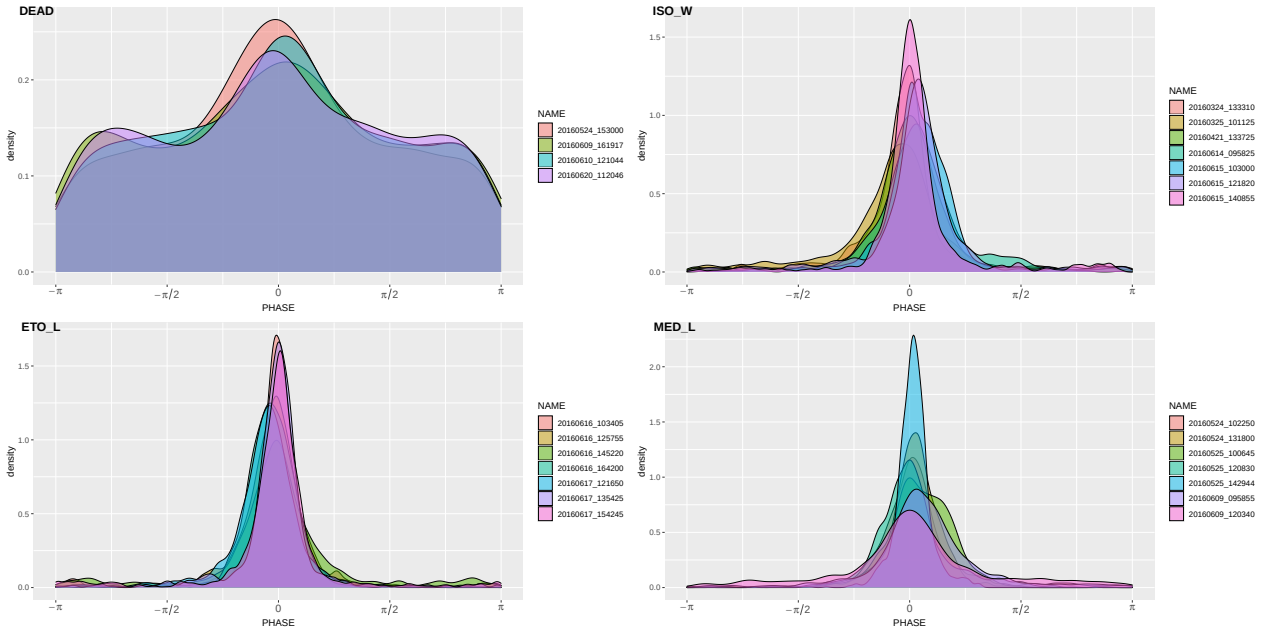

Figure 5: Plot of the empirical distribution of the phases $\widehat{\phi}$ obtained for the 4 groups of rats without thresholding the correlations. Each color corresponds to a rat.

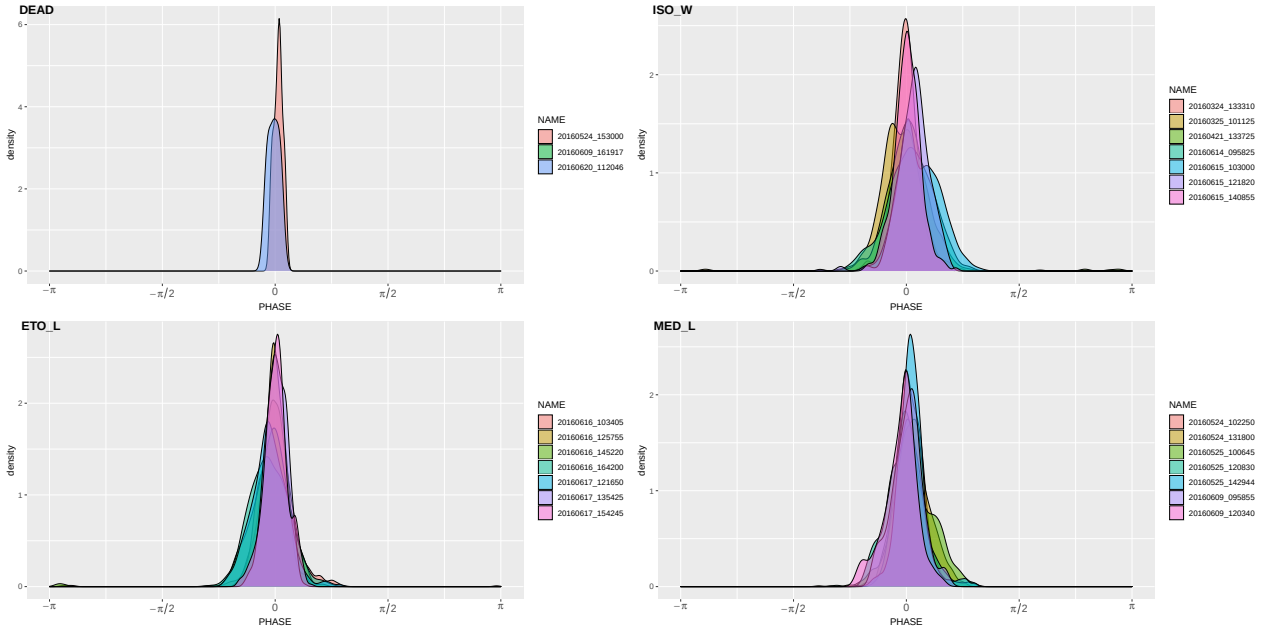

Figure 6: Plot of the empirical distribution of the phases obtained for the 4 groups of rats after first thresholding the correlations. Only phases associated to correlations with a magnitude higher than 0.3 are considered. Each color corresponds to a rat.

within each group. Edges correspond to a magnitude higher than 0.3. The value of the threshold is motivated by the observation of the supports obtained for dead rats. We then select the edges which are present in all the graphs of the rats of the group. One graph is then obtained per group. For each group, we then compute the average of the estimated phase for each detected edge. Figure 7 illustrates the graphs obtained for the 4 different groups.

A color label indicates whether the phases are positive or negative. We colored each edge based on the mean phase when it satisfies $|\phi_{jk}| > 1.1|\phi^*_{jk}|$ where $\phi^*_{jk} = -\frac{\pi}{2}(d_j - d_k)$. The value ϕ^*_{jk} corresponds to the phase of causal linear representations with power-law coefficients [Kechagias and Pipiras, 2014] and to the ARFIMA modelling used in Achard and Gannaz [2016] with similar data. The more colored edges, the more the phase behaviour differs from the previous modelling.

The DEAD group has indeed no edges. The MED_L group has less edges than the two other groups of anesthetic. It hence seems that MED_L anesthetic inhibits more the activity. Next ETO_L group and ISO_W group have a similar number of edges (respectively 133 and 145), but the phases differ. More than half of mean phases are outside the interval $[-1.1|\phi^*|, 1.1|\phi^*|]$ for ETO_L and ISO_W group, with similar proportions. This observation is interesting since it illustrates that the modelling of these data is complex. The introduction of a general phase enables to take into account this complexity. Concerning the physical interpretation, no easy conclusion can be given. In our case with fMRI time series, the presence of latency seems not probable because of the action of hemodynamic function which introduces long delays of around 5 seconds.

6 Conclusion

This work was motivated by a neuroscience application, namely the inference of fractal connectivity from fMRI recordings. We studied the local Whittle estimators for multivariate time series presenting long-range dependence. Our modelling allows a complex covariance structure with phase components which can be interpreted as shifts in the coupling between time series. We introduced quasi-analytic wavelet filters to handle the possible non stationarity of the real data application. The resulting procedures offer a consistent estimation of the main parameters of the model. Indeed, we 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 phase. A simulation study on linear processes and on multivariate Brownian motions illustrate the good performance of the proposed procedure. The real data application highlightes the ability of the procedure to distinguish dead rats from alive rats. It also showed the differences between three anaesthetics and the fact that one of them slows down more intensively the brain activity.

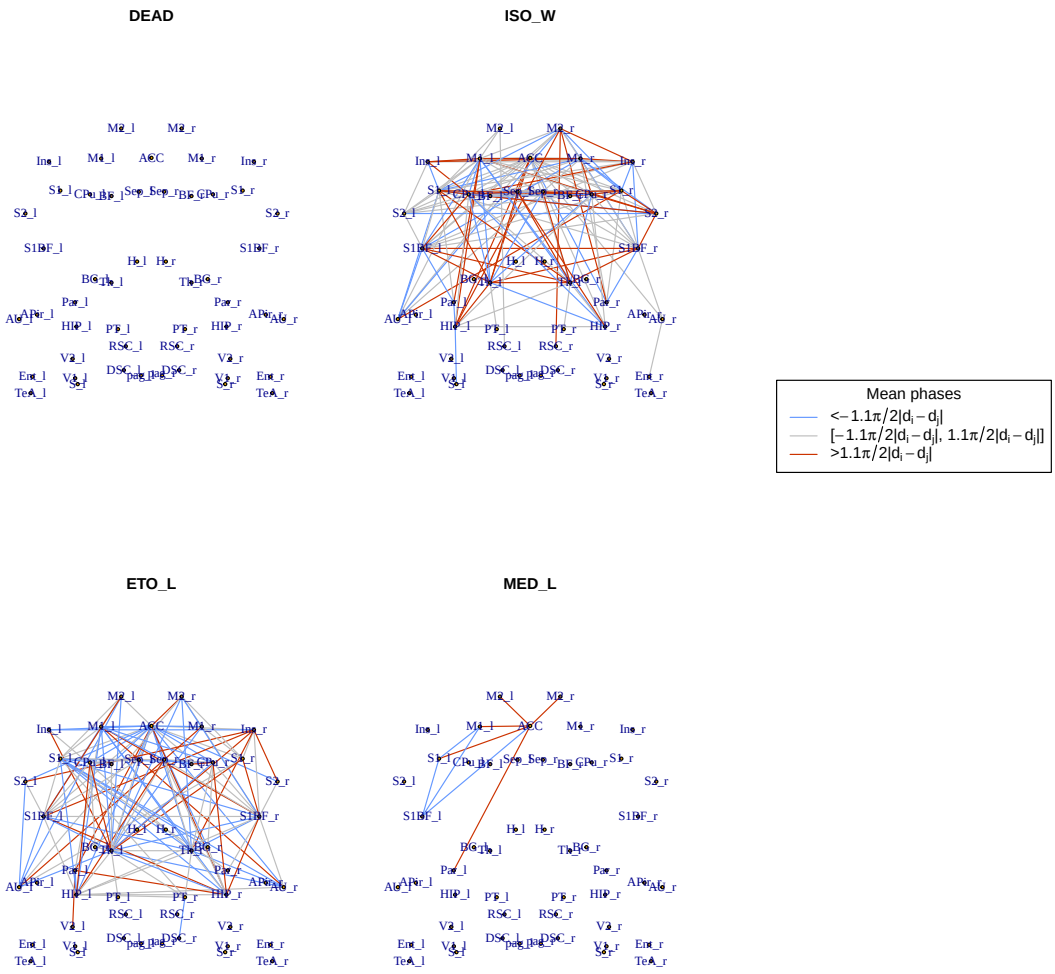

Figure 7: Plot of the mean graphs with correlations and phases obtained for 4 groups of rats: DEAD, ISO_W, ETO_L and MED_L. Only edges corresponding to a mean of correlation's magnitude higher than 0.3 are displayed. Red edges correspond to positive means of phase higher than $1.1|\phi^*|$, blue edges correspond to negative mean of phase lower than -$1.1|\phi^*|$, and grey edges to mean of phases between -$1.1|\phi^*|$ and $1.1|\phi^*|$. The quantities ϕ^* are equal to $\phi^*_{j,k} = -\frac{\pi}{2}(d_j - d_k)$.

Ackowledgments

The authors are grateful to Marianne Clausel and Franҫis Roueff who contributed significantly to this work. We also thank Jean-François Coeurjolly for providing the code for the simulations of

multivariate Brownian motions. We also would like to thank Emmanuel Barbier and Guillaume Becq for providing us the data of the resting state fMRI on the rats.

A Properties of CFW-C(M,L) filters: proof of Proposition 3

The aim of this section is to check whether the assumptions of Section 2.2 are satisfied by CFW-C(M,L).

Let us first give some properties of the filter $\widehat{d}_L$ defined in (1).

Lemma 8.

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

Proof is straightforward and thus omitted.

Next, the modulus of the filter $\widehat{d}_L$ is equal to $\left|\widehat{d}_L(\lambda)\right|^2 = \cos(\lambda/4)^{2(2L+1)} \left(1+\tan(\lambda/4)^{2(2L+1)}\right)$. Recall we defined $\alpha_L(\cdot)$ as $\alpha_L(\lambda) = 2\,(-1)^L \operatorname{atan}\left(\tan(\lambda/4)^{2L+1}\right)$. Then,

$$\widehat{d}_L(\lambda) = \cos(\lambda/4)^{2(2L+1)} \left(1+\tan(\lambda/4)^{2(2L+1)}\right) \mathrm{e}^{\mathrm{i}\,\lambda(-2L+1)/4 - \mathrm{i}\,\alpha_L(\lambda)/2}. \tag{10}$$

A direct consequence is that

$$\prod_{\ell=1}^{\infty} \left|\widehat{d}_L(2^{-\ell}\lambda)\right| \leqslant \left|\frac{\sin(\lambda/4)}{\lambda/4}\right|^{2L+1}. \tag{11}$$

A.1 Assumption (A-a)

By definition of the filters, the filters $h^{(L)}$ and $h^{(H)}$, $g^{(L)}$ and $g^{(H)}$ have length $M+L+1$. Then τ_j has finite length.

A.2 Assumption (A-b)

Recall that

$$2^{-j/2}\widehat{h}_j(\lambda) = \left(\frac{1-\mathrm{e}^{\mathrm{i}2^{j-1}\lambda}}{2}\right)^M \overline{\widehat{d}_L(2^{j-1}\lambda+\pi)}\mathrm{e}^{-\mathrm{i}2^{j-1}\lambda}\prod_{\ell=0}^{j-2}\left(\frac{1+\mathrm{e}^{-\mathrm{i}2^\ell\lambda}}{2}\right)^M \widehat{d}_L(2^\ell\lambda),$$

$$2^{-j/2}\widehat{g}_j(\lambda) = \left(\frac{1-\mathrm{e}^{\mathrm{i}2^{j-1}\lambda}}{2}\right)^M \widehat{d}_L(2^{j-1}\lambda+\pi)\mathrm{e}^{\mathrm{i}2^{j-1}\lambda L}\mathrm{e}^{-\mathrm{i}2^{j-1}\lambda}\prod_{\ell=0}^{j-2}\left(\frac{1+\mathrm{e}^{-\mathrm{i}\,2^\ell\lambda}}{2}\right)^M \overline{\widehat{d}_L(2^\ell\lambda)}\mathrm{e}^{-\mathrm{i}2^\ell\lambda L}.$$

Remark that $\left|\frac{1-\mathrm{e}^{\mathrm{i}2^{j-1}\lambda}}{2}\right|^M = \left|\sin(2^{j-2}\lambda)\right|^M \leqslant 2^{-2M}\left|2^j\lambda\right|^M$. Since $\sup|\widehat{d}_L| = 1$, assumption (A-b) follows with a constant $C_m = 1$.

A.3 Assumption (A-c)

Recall that

$$2^{-j/2}\widehat{h}_j(2^{-j}\lambda) = \left(\frac{1-\mathrm{e}^{\mathrm{i}\lambda/2}}{2}\right)^M \mathrm{e}^{-\mathrm{i}\lambda}\,\overline{\widehat{d}_L(\lambda/2+\pi)}\prod_{\ell=2}^{j}\left(\frac{1+\mathrm{e}^{-\mathrm{i}2^{-\ell}\lambda}}{2}\right)^M \widehat{d}_L(2^{-\ell}\lambda)$$

Since $\sup_{\lambda\in\mathbb{R}}\left|\widehat{d}_L(\lambda)\right| \leqslant 1$,

$$\left|2^{-j/2}\widehat{h}_j(2^{-j}\lambda)\right| \leqslant |\sin(\lambda/4)|^M \prod_{\ell=1}^{\infty}\left|\frac{1+\mathrm{e}^{-\mathrm{i}2^{-\ell}\lambda}}{2}\right|^M \prod_{\ell=1}^{\infty}\left|\widehat{d}_L(2^{-\ell}\lambda\right|.$$

Using inequality (11) and equality

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

we obtain

$$\left|2^{-j/2}\widehat{h}_j(2^{-j}\lambda)\right| \leqslant |\sin(\lambda/4)|^M\left|\frac{\sin(\lambda/2)}{\lambda/2}\right|^M\left|\frac{\sin(\lambda/4)}{\lambda/4}\right|^{2L+1} \leqslant |\sin(\lambda/4)|^M\left|\frac{\sin(\lambda/2)}{\lambda/2}\right|^{M+2L+1},$$

using the fact that $x \mapsto \frac{\sin(x)}{x}$ decreases on $\mathbb{R}^+$. Since $\frac{\sin(x/2)}{|x/2|}(1+|x|) = \left|\frac{\sin(x/2)}{x/2}\right| + 2\,|\sin(x/2)| \leqslant 3$ for any $x \in \mathbb{R}\setminus\{0\}$, it follows that

$$\left|2^{-j/2}\widehat{h}_j(2^{-j}\lambda)\right| \leqslant |\sin(\lambda/4)|^M\left(\frac{3}{1+|\lambda|}\right)^{M+2L+1}. \tag{12}$$

It follows that $\left|2^{-j/2}\widehat{h}_j(\lambda)\right| \leqslant \left(\frac{3}{1+|\lambda|}\right)^{M+2L+1}$. A similar result can be proved for filter $\widehat{g}_j$. By triangular inequality, assumption (A-c) hence holds with $\alpha = M + 2L + 1$ and a constant C_s equal to $2 \cdot 3^{M+2L+1}$ for filter $\widehat{\tau}_j$.

The proof for $\widehat{\tau}_\infty$ is similar and thus omitted.

The constant C_s above may pose a problem, thus we can establish a similar result with a different constant gor large λ. On (π, ∞), we have

$$\left(\frac{\sin(x)}{x}\right)^2 (1 + 2x) \leqslant \frac{1}{x^2} + \frac{2}{x} \leqslant \frac{1}{\pi^2} + \frac{2}{\pi} \leqslant 1.$$

Thus we obtain the following lemma:

Lemma 9. *For all $|\lambda| \geqslant \pi$,*

$$\left|2^{-j/2}\widehat{\tau}_j(2^{-j}\lambda)\right| \leqslant 2\frac{1}{|1+\lambda|^{M/2+L+1/2}}.$$

Inequality (12) also gives a stronger result which will be useful in the proof of assumption (A-d):

Lemma 10. *For all $\lambda \in \mathbb{R}$,*

$$\left|2^{-j/2}\widehat{\tau}_j(2^{-j}\lambda)\right| \leqslant 2 \cdot 3^{M+2L+1}\frac{|\lambda|^M}{|1+\lambda|^{M+2L+1}}.$$

A.4 Assumption (A-d)

We have

$$\widehat{h}_\infty(\lambda) = 2^{-j/2}\widehat{h}_j(2^{-j}\lambda) \cdot \left(\frac{\sin(\lambda/2^{j+1})}{\lambda/2^{j+1}}\right)^M \mathrm{e}^{\mathrm{i}\lambda M/2^{j+1}} \left(\prod_{m=j+1}^{\infty} \widehat{d}_L(2^{-m}\lambda)\right)$$

$$\widehat{g}_\infty(\lambda) = 2^{-j/2}\widehat{g}_j(2^{-j}\lambda) \cdot \left(\frac{\sin(\lambda/2^{j+1})}{\lambda/2^{j+1}}\right)^M \mathrm{e}^{\mathrm{i}\lambda(M+2L)/2^{j+1}} \left(\prod_{m=j+1}^{\infty} \overline{\widehat{d}_L(2^{-m}\lambda)}\right).$$

Thus,

$$\widehat{h}_\infty(\lambda)\, e^{-\mathrm{i}\,\lambda\, 2^{-j-1}(M+L)} = 2^{-j/2}\widehat{h}_j(2^{-j}\lambda) \cdot Z_j(\lambda) \tag{13}$$

$$\widehat{g}_\infty(\lambda)\, e^{-\mathrm{i}\,\lambda\, 2^{-j-1}(M+L)} = 2^{-j/2}\widehat{g}_j(2^{-j}\lambda) \cdot \overline{Z_j(\lambda)}, \tag{14}$$

with

$$Z_j(\lambda) = \left(\frac{\sin(\lambda/2^{j+1})}{\lambda/2^{j+1}}\right)^M e^{-\mathrm{i}\lambda 2^{-j-1}L} \prod_{m=j+1}^{\infty} \widehat{d}_L(2^{-m}\lambda).$$

An inequality on $Z_j(\lambda)$ is given in the following lemma.

Lemma 11. *There exists a constant C_Z depending on L and M such that for all $j \in \mathbb{N}$, for all $\left|2^{-j}\lambda\right| < \pi$,*

$$\left|\mathrm{e}^{\mathrm{i}\,\Phi_j(\lambda)}\, Z_j(\lambda) - 1\right| \leqslant C_Z \left|2^{-j}\lambda\right|^2,$$

with $\Phi_j(\lambda) = \lambda\, 2^{-j-1}(2L - 1/2)$.

Proof. With (10), $Z_j(\lambda)$ can be rewritten as

$$\begin{aligned} Z_j(\lambda) &= \left(\frac{\sin(\lambda/2^{j+1})}{\lambda/2^{j+1}}\right)^M e^{-\mathrm{i}\lambda 2^{-j-1}L} \prod_{m=1}^{\infty} \cos(2^{-m}2^{-j}\lambda/4)^{L+1/2}\, \mathrm{e}^{\mathrm{i}\,2^{-m}2^{-j}\lambda(-2L+1)/4 - \alpha_L(2^{-m}2^{-j}\lambda)/2} \\ &= \left(\frac{\sin(\lambda/2^{j+1})}{\lambda/2^{j+1}}\right)^M \left(\frac{\sin(\lambda/2^{j+2})}{\lambda/2^{j+2}}\right)^{L+1/2} \mathrm{e}^{-\mathrm{i}\,\Phi_j(\lambda) - \mathrm{i}\,\sum_{m=1}^{\infty} \alpha_L(2^{-m}2^{-j}\lambda)/2}. \end{aligned}$$

Last equality comes from the fact that for all $x \in \mathbb{R}$ $\frac{\sin(x)}{x} = \frac{\sin(x/2)}{x/2}\cos(x/2)$ and thus $\frac{\sin(x)}{x} = \prod_{m=1}^{\infty}\cos(2^{-m}x)$.

We now study of the different terms.

- First, Taylor inequalities state that for all $x \in \mathbb{R}\setminus\{0\}$, $|\sin(x)/x| \leqslant 1$ and $|\sin(x) - x| \leqslant |x|^3/6$. Using the equality $(x^K - 1) = (x-1)\sum_{m=0}^{K-1} x^m$ for all $x \in \mathbb{R}$, $K \in \mathbb{N}$,

  $$\left|\left(\frac{\sin(\lambda/2^{j+1})}{\lambda/2^{j+1}}\right)^M - 1\right| \leqslant \left|2^{-j-1}\lambda\right|^2 M/6$$

  and a similar result holds for $\left|\left(\frac{\sin(\lambda/2^{j+1})}{\lambda/2^{j+1}}\right)^{L+1/2} - 1\right|$.

- Next,

  $$\left|\mathrm{e}^{-\mathrm{i}\sum_{m=1}^{\infty}\alpha_L(2^{-m}2^{-j}\lambda)/2} - 1\right| = \left|\sin\left(\sum_{m=1}^{\infty}\alpha_L(2^{-m}2^{-j}\lambda)\right)\right| \leqslant \left|\sum_{m=1}^{\infty}\alpha_L(2^{-m}2^{-j}\lambda)\right|.$$

  By definition, $\alpha_L(\lambda)/2 = (-1)^{L+1}\,\mathrm{atan}\left(\tan(\lambda/4)^{2L+1}\right)$, thus for $|\lambda| < \pi$,

  $$|\alpha_L(\lambda)/2| \leqslant |\tan(\lambda/4)|^{2L+1} \leqslant (\lambda\cdot 4/\pi)^{2L+1}.$$

We obtain, for $\left|2^{-j}\lambda\right| < \pi$,

$$\left|\mathrm{e}^{-\mathrm{i}\sum_{m=1}^{\infty}\alpha_L(2^{-m}2^{-j}\lambda)/2} - 1\right| \leqslant (4/\pi)^{2L+1}\left|2^{-j}\lambda\right|^{2L+1}\sum_{m=1}^{\infty}2^{-m(2L+1)} \leqslant (4/\pi)^{2L+1}\left|2^{-j}\lambda\right|^{2L+1}.$$

We get

$$\left|\mathrm{e}^{\mathrm{i}\,\Phi_j(\lambda)}\,Z_j(\lambda) - 1\right| \leqslant C\,\left|2^{-j-1}\lambda\right|^2,$$

with $C = (4\,M + 2\,L + 1)\,\pi/192 + 4^{2L+1}/\pi$, when $\left|2^{-j}\lambda\right| \leqslant \pi$. □

Going back to (13) and (14),

$$\left|\widehat{h}_\infty(\lambda)\mathrm{e}^{-\mathrm{i}\,\lambda\,2^{-j-1}(M+L)+\mathrm{i}\,\Phi_j(\lambda)} - 2^{-j/2}\widehat{h}_j(2^{-j}\lambda)\right| \leqslant C_Z\left|2^{-j}\lambda\right|^2\left|\widehat{h}_\infty(\lambda)\right|, \tag{15}$$

$$\left|\widehat{g}_\infty(\lambda)\mathrm{e}^{-\mathrm{i}\,\lambda\,2^{-j-1}(M+L)-\mathrm{i}\,\Phi_j(\lambda)} - 2^{-j/2}\widehat{g}_j(2^{-j}\lambda)\right| \leqslant C_Z\left|2^{-j}\lambda\right|^2\left|\widehat{g}_\infty(\lambda)\right|. \tag{16}$$

Let

$$R_j(\lambda) = \widehat{\tau}_\infty(\lambda)\,e^{-\mathrm{i}\,\lambda\,2^{-j-1}(M+L)-\mathrm{i}\,\Phi_j(\lambda)} - 2^{-j/2}\widehat{\tau}_j(2^{-j}\lambda) + \widehat{h}_\infty(\lambda)\left(e^{\mathrm{i}\,\Phi_j(\lambda)} - e^{-\mathrm{i}\,\Phi_j(\lambda)}\right)e^{-\mathrm{i}\,\lambda\,2^{-j-1}(M+L)}.$$

Inequalities (15) and (16) and assumption (A-b) lead to

$$|R_j(\lambda)| \leqslant C_R\,2^{-2j}\,|\lambda|^{M+2},$$

with $C_R = 4\,C_m\,C_Z$.

We have

$$|\widehat{\tau}_\infty(\lambda)| - \left|2^{-j/2}\widehat{\tau}_j(2^{-j}\lambda)\right| \leqslant |R_j(\lambda)| + \left|\widehat{h}_\infty(\lambda)\right|\left|e^{\mathrm{i}\,\Phi_j(\lambda)} - e^{-\mathrm{i}\,\Phi_j(\lambda)}\right|.$$

With assumption (A-b),

$$|\widehat{\tau}_\infty(\lambda)| - \left|2^{-j/2}\widehat{\tau}_j(2^{-j}\lambda)\right| \leqslant \left(C_R\,2^{-j}\,|\lambda|^2 + C_m\,\pi(L - 1/4)\,|\lambda|\right)2^{-j}\,|\lambda|^M.$$

Using again assumption (A-b), Assumption (A-d) is straightforward with $\gamma = 1$.

B Properties of CFW-PR(M,L) filters

The aim of this section is to check whether the assumptions of Section 2.2 are satisfied by CFW-PR(M,L).

Recall that

$$2^{-j/2}\widehat{h}_j(\lambda) = \Big(\frac{1-\mathrm{e}^{\mathrm{i}2^{j-1}\lambda}}{2}\Big)^M \overline{\widehat{d}_L(2^{j-1}\lambda+\pi)}\,\overline{\widehat{q}_L(2^{j-1}\lambda+\pi)}\mathrm{e}^{-\mathrm{i}2^{j-1}\lambda}\prod_{\ell=0}^{j-2}\Big(\frac{1+\mathrm{e}^{-\mathrm{i}2^\ell\lambda}}{2}\Big)^M \widehat{d}_L(2^\ell\lambda)\widehat{q}_L(2^\ell\lambda),$$

$$2^{-j/2}\widehat{g}_j(\lambda) = \Big(\frac{1-\mathrm{e}^{\mathrm{i}2^{j-1}\lambda}}{2}\Big)^M \widehat{d}_L(2^{j-1}\lambda+\pi)\widehat{q}_L(2^{j-1}\lambda+\pi)\mathrm{e}^{\mathrm{i}2^{j-1}\lambda L}\mathrm{e}^{-\mathrm{i}2^{j-1}\lambda}$$
$$\prod_{\ell=0}^{j-2}\Big(\frac{1+\mathrm{e}^{-\mathrm{i}\,2^\ell\lambda}}{2}\Big)^M \overline{\widehat{d}_L(2^\ell\lambda)}\,\overline{\widehat{q}_L(2^\ell\lambda)}\mathrm{e}^{-\mathrm{i}2^\ell\lambda L}.$$

with

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

and $s(\lambda) = 2^{-M}(1+\cos(\lambda))^M \left|\widehat{d}_L(\lambda)\right|^2 = (\cos(\lambda/2))^{2M}\left|\widehat{d}_L(\lambda)\right|^2$.

We have the following lemma:

Lemma 12. *For all $\lambda \in \mathbb{R}$,*

$$|\widehat{q}(\lambda)| \leqslant 2^{M+2L+1}\ .$$

Proof. Consider the functions $s : \lambda \mapsto s(\lambda)$ and $s_\pi : \lambda \mapsto s(\lambda+\pi)$. Both functions are 4π-periodic. Based on their variations on $(-2\pi, 2\pi)$, the minimum between the functions is attained when they are equal, thus for $-3\pi/2$ and $3\pi/2$. Consequently the minimum value is $2^{-M}(\cos(3\pi/8)^{4L+2} + \sin(3\pi/8)^{4L+2})$ which is greater than $2^{-M-2L-1}$. □

It is straightforward that, based on Lemma 12,

$$\left|2^{-j/2}\widehat{h}_j(\lambda)\right| \leqslant \left|\frac{1-\mathrm{e}^{\mathrm{i}2^{j-1}\lambda}}{2}\right|^M \overline{\widehat{q}_L(2^{j-1}\lambda+\pi)} \leqslant 2^{M+2L+1}\left|\frac{1-\mathrm{e}^{\mathrm{i}2^{j-1}\lambda}}{2}\right|^M\ .$$

Assumption (A-b) follows since $\left|\frac{1-\mathrm{e}^{\mathrm{i}2^{j-1}\lambda}}{2}\right|^M \leqslant \left|2^j\lambda\right|^M$, with $C_m = 2^{M+2L+2}$.

C Asymptotic behaviour of the wavelet covariance

This section deals with the proofs of results of Section 2.3.

C.1 Proof of Proposition 4

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

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

We now sum up the main points for the convergence of $\mathrm{Cov}(\mathbf{W}_{jk})$.

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

$$\mathbf{B}_j^{(+)} = \int_0^{\pi\, 2^j} \boldsymbol{\Gamma}_j(\mathbf{d})^{-1}\mathbf{D}(\lambda)\boldsymbol{\Omega}\mathbf{D}(\lambda)\boldsymbol{\Gamma}_j(\mathbf{d})^{-1}\; 2^{-j} \left|\widehat{\tau}_j(2^{-j}\lambda)\right|^2 \,\mathrm{d}\lambda$$
$$\mathbf{I}_j^{(+)\mathrm{inf}} = \int_0^{\pi\, 2^j} \boldsymbol{\Gamma}_j(\mathbf{d})^{-1}\mathbf{D}(\lambda)\boldsymbol{\Omega}\mathbf{D}(\lambda)\boldsymbol{\Gamma}_j(\mathbf{d})^{-1}\; |\widehat{\tau}_\infty(\lambda)|^2 \,\mathrm{d}\lambda$$
$$\mathbf{I}_j^{(+)\mathrm{sup}} = \int_{\pi\, 2^j}^{\infty} \boldsymbol{\Gamma}_j(\mathbf{d})^{-1}\mathbf{D}(\lambda)\boldsymbol{\Omega}\mathbf{D}(\lambda)\boldsymbol{\Gamma}_j(\mathbf{d})^{-1}\; |\widehat{\tau}_\infty(\lambda)|^2 \,\mathrm{d}\lambda$$
$$\mathbf{I}^{(+)} = \mathbf{I}_j^{(+)\mathrm{inf}} + \mathbf{I}_j^{(+)\mathrm{sup}} = \int_0^{\infty} \boldsymbol{\Gamma}_j(\mathbf{d})^{-1}\mathbf{D}(\lambda)\boldsymbol{\Omega}\mathbf{D}(\lambda)\boldsymbol{\Gamma}_j(\mathbf{d})^{-1}\; |\widehat{\tau}_\infty(\lambda)|^2 \,\mathrm{d}\lambda$$

   The steps of the convergence are:

   (a) $\left|\mathbf{A}_j^{(+)} - \mathbf{B}_j^{(+)}\right|$ is bounded using the regularity of the spectral density $\boldsymbol{f}^S(\cdot)$ at the origin, that is, (M-2), together with assumption (A-b).

   (b) $\left|\mathbf{B}_j^{(+)} - \mathbf{I}_j^{(+)\mathrm{inf}}\right|$ is bounded using the convergence of the filter τ_j to τ_∞, through property (A-d). We shall need assumption (A-c).

   (c) $\left|\mathbf{I}_j^{(+)\mathrm{sup}}\right|$ is bounded using the regularity of the filter τ_∞, that is, using (A-c).

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

2. Behaviour 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

$$\mathbf{I}^{(-)} = \int_{-\infty}^{0} \boldsymbol{\Gamma}_j(\mathbf{d})^{-1}\mathbf{D}(\lambda)\boldsymbol{\Omega}\mathbf{D}(\lambda)\boldsymbol{\Gamma}_j(\mathbf{d})^{-1}\; |\widehat{\tau}_\infty(\lambda)|^2 \,\mathrm{d}\lambda.$$

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

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

First notice that $\mathbf{\Gamma}_j(\mathbf{d})^{-1}\mathbf{D}(2^j\lambda) = \mathbf{D}(\lambda)$. Hence,

$$\begin{aligned}\left|\mathbf{A}_j^{(+)} - \mathbf{B}_j^{(+)}\right| &\leqslant \int_0^\pi |\boldsymbol{f}(\lambda) \ - \mathbf{D}(\lambda)\mathbf{\Omega}\mathbf{D}(\lambda)|\, |\widehat{\tau}_j(\lambda)|^2 \ \mathrm{d}\lambda \\ &\leqslant \int_0^\pi |\mathbf{D}(\lambda)\mathbf{\Omega}\mathbf{D}(\lambda)| \circ \left|\boldsymbol{f}^S(\lambda) - 1\right| |\widehat{\tau}_j(\lambda)|^2 \ \mathrm{d}\lambda.\end{aligned}$$

Assumption (M-2) gives:

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

With a change of variable,

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

We split the integral in two parts. First, with assumption (A-b),

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

As the parameters satisfy (C-a), the integral is bounded by a constant depending on $(d_\ell, d_m, \beta, M, L)$.

Next, using the regularity given by assumption (A-c),

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

Assumption (C-a) ensures that the right hand side is bounded by a constant depending of $d_\ell, d_m, \beta, \alpha, M$ and L.

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

This step uses the convergence of the filter τ_j to τ_∞, through property (A-d). First,

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

Using (A-c),

$$\left|\left|2^{-j/2}\widehat{\tau}_j(2^{-j}\lambda)\right|^2 - |\widehat{\tau}_\infty(\lambda)|^2\right| \leqslant \left|\left|2^{-j/2}\widehat{\tau}_j(2^{-j}\lambda)\right|^2 + |\widehat{\tau}_\infty(\lambda)|^2\right| \leqslant 2\, C_s^2 \ |\lambda|^{-2\alpha}$$

Thus,

$$\left(\int_{\pi}^{2^j\pi} \mathbf{D}(\lambda)\mathbf{\Omega}\mathbf{D}(\lambda)\, 2^{j\beta} \left|\left|2^{-j/2}\widehat{\tau}_j(2^{-j}\lambda)\right|^2 - |\widehat{\tau}_\infty(\lambda)|^2\right| \,\mathrm{d}\lambda\right)_{(\ell,m)}$$
$$\leqslant \|\mathbf{\Omega}\|\, 2\, C_s^2\, 2^{j\beta} \int_{\pi}^{2^j\pi} |\lambda|^{-d_\ell-d_m-2\alpha}\, d\lambda$$
$$\leqslant \|\mathbf{\Omega}\|\, 2\, C_s^2\, \pi^{-d_\ell-d_m-2\alpha+1}\, 2^{j(-d_\ell-d_m-2\alpha+\beta+1)}.$$

The right-hand side tends to 0 when j goes to infinity due to (C-a).

It remains to consider the integral on $(0,\pi)$. Assumption (A-d) states that

$$\left(\int_{0}^{\pi} \mathbf{D}(\lambda)\mathbf{\Omega}\mathbf{D}(\lambda)\, 2^{j\beta} \left|\left|2^{-j/2}\widehat{\tau}_j(2^{-j}\lambda)\right|^2 - |\widehat{\tau}_\infty(\lambda)|^2\right| \,\mathrm{d}\lambda\right)_{(\ell,m)}$$
$$\leqslant \|\mathbf{\Omega}\|\, C_a\, 2^{j(\beta-\gamma)} \int_{0}^{\pi} C_a\, |\lambda|^{-d_\ell-d_m+2M}\, d\lambda$$

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

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

This step uses the regularity of the filter τ_∞. Indeed, using (A-c),

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

Assumption (C-a) thus implies that $2^{-j(d_\ell+d_m-\beta)}\left|\mathbf{I}_{j,\ell m}^{(+)\mathrm{sup}}\right|$ is bounded by a constant depending of $(d_\ell, d_m, \beta, M, L)$.

C.2 Proof of Proposition 5

The proof under assumption (A-e) is identical to the proof of Proposition 4 and it is thus omitted.

C.3 Proof of Proposition 6

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

Corollary 13. *For all $\widehat{q}_{L,M}$ real polynomial of $(e^{-i\lambda})$, for all $|\lambda| \leqslant \pi$,*

$$\left| |\widehat{\tau}_\infty(\lambda)|^2 - 4\,\mathbb{1}_{\mathbb{R}_+}(\lambda)\,\left|\widehat{h}_\infty(\lambda)\right|^2 \right| \leqslant 8\,(\log_2(2)+2)\,\left(1-\frac{|\lambda|}{4\pi}\right)^{4L+2} \left|\widehat{h}_\infty(\lambda)\right|^2 .$$

Proof. From (4),

$$\begin{aligned}
\left| |\widehat{\tau}_\infty(\lambda)|^2 - 4\,\mathbb{1}_{\mathbb{R}_+}(\lambda)\,\left|\widehat{h}_\infty(\lambda)\right|^2 \right| &= \left| \left|1-\mathrm{e}^{\mathrm{i}\eta_L(\lambda)}\right|^2 - 4\mathbb{1}_{\mathbb{R}_+}(\lambda) \right| \cdot \left|\widehat{h}_\infty(\lambda)\right|^2 \\
&= 4 \cdot \left|\sin(\eta_L(\lambda)/2)^2 - \mathbb{1}_{\mathbb{R}_+}(\lambda)\right| \cdot \left|\widehat{h}_\infty(\lambda)\right|^2 \\
&= 4\left|\widehat{h}_\infty(\lambda)\right|^2 \cdot \begin{cases} \sin(\eta_L(\lambda)/2)^2 & \lambda < 0 \\ \cos(\eta_L(\lambda)/2)^2 & \lambda > 0 \end{cases}.
\end{aligned}$$

Since $2\,\cos(x/2)^2 = 1+\cos(x)$ and $2\,\sin(x/2)^2 = 1-\cos(x)$ for all $x \in \mathbb{R}$, we have

$$\begin{aligned}
\left| |\widehat{\tau}_\infty(\lambda)|^2 - 4\,\mathbb{1}_{\mathbb{R}_+}(\lambda)\,\left|\widehat{h}_\infty(\lambda)\right|^2 \right| &= \left|\mathrm{e}^{\mathrm{i}\,\eta_L(\lambda)} + \mathrm{sign}(\lambda)\right|^2 \left|\widehat{h}_\infty(\lambda)\right|^2 \\
&= \left|\mathrm{e}^{\mathrm{i}\,\eta_L(\lambda)} - 1 + 2\mathbb{1}_{\mathbb{R}_+}(\lambda)\right|^2 \left|\widehat{h}_\infty(\lambda)\right|^2 .
\end{aligned}$$

The function U_L satisfies, for all $|\lambda| < 2\pi$:

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

which concludes the proof. □

Since Proposition 4 holds, it remains to prove that the following integrals are bounded:

$$\begin{aligned}
&2^{j\beta}\int_{-\pi}^{\pi} |\lambda|^{-d_\ell-d_m} \left| |\widehat{\tau}_\infty(\lambda)|^2 - 4\,\mathbb{1}_{\mathbb{R}_+}(\lambda)\,\left|\widehat{h}_\infty(\lambda)\right|^2 \right| d\lambda, \\
&2^{j\beta}\int_{-\infty}^{-\pi} |\lambda|^{-d_\ell-d_m} \left| |\widehat{\tau}_\infty(\lambda)|^2 - 4\,\mathbb{1}_{\mathbb{R}_+}(\lambda)\,\left|\widehat{h}_\infty(\lambda)\right|^2 \right| d\lambda, \\
&2^{j\beta}\int_{\pi}^{\infty} |\lambda|^{-d_\ell-d_m} \left| |\widehat{\tau}_\infty(\lambda)|^2 - 4\,\mathbb{1}_{\mathbb{R}_+}(\lambda)\,\left|\widehat{h}_\infty(\lambda)\right|^2 \right| d\lambda.
\end{aligned}$$

On $(0,\pi)$.

With assumption (A-b), Corollary 13 leads to

$$2^{j\beta}\int_0^{\pi}|\lambda|^{-d_\ell-d_m}\left||\widehat{\tau}_\infty(\lambda)|^2-4\,\mathbb{1}_{\mathbb{R}_+}(\lambda)\,\left|\widehat{h}_\infty(\lambda)\right|^2\right|d\lambda\leqslant C_m\,2^{j\beta}\int_0^{\pi}\left(1-\frac{|\lambda|}{4\pi}\right)^{2\alpha}|\lambda|^{-d_\ell-d_m+2M}d\lambda$$

with $\alpha=2L+1$ and $C_m=1$.

Using Cauchy-Schwarz inequality,

$$\begin{aligned}&2^{j\beta}\int_0^{\pi}\left(1-\frac{|\lambda|}{4\pi}\right)^{2\alpha}|\lambda|^{-d_\ell-d_m+2M}d\lambda\\&\leqslant 2^{j\beta}\left(\int_0^{\pi}\left(1-\frac{|\lambda|}{4\pi}\right)^{4\alpha}d\lambda\right)^{1/2}\left(\int_0^{\pi}|\lambda|^{-2d_\ell-2d_m+4M}\,\mathrm{d}\lambda\right)^{1/2}\\&\leqslant 2^{j\beta}\frac{1}{2\alpha^{1/2}}\left(\int_0^{\pi}|\lambda|^{-2d_\ell-2d_m+4M}\,\mathrm{d}\lambda\right)^{1/2}.\end{aligned}$$

The integral on the right hand side is bounded with parameters satisfying (C-b). Thus it is sufficient to take

$$L\geqslant 2^{2j\beta}.$$

On $(\pi,2^j\pi)$.

With Lemma 9 and Corollary 13,

$$\left||\widehat{\tau}_\infty(\lambda)|^2-4\,\mathbb{1}_{\mathbb{R}_+}(\lambda)\,\left|\widehat{h}_\infty(\lambda)\right|^2\right|d\lambda\leqslant C_s\,|\lambda|^{-2\alpha}\tag{17}$$

with $C_s=1$ and $\alpha=M/2+L+1/2$. Thus,

$$\begin{aligned}2^{j\beta}\int_{\pi}^{2^j\pi}|\lambda|^{-d_\ell-d_m}\left||\widehat{\tau}_\infty(\lambda)|^2-4\,\mathbb{1}_{\mathbb{R}_+}(\lambda)\,\left|\widehat{h}_\infty(\lambda)\right|^2\right|d\lambda&\\\leqslant C_s\,2^{j(\beta-d_\ell-d_m-M-1-2L)}\pi^{-d_\ell-d_m-M-1-2L}d\lambda\,.&\end{aligned}$$

The integral on the right hand side is bounded under (C-b), for any $L\geqslant 1$.

On $(2^j\pi, \infty)$.

With (17),

$$2^{j\beta}\int_{2^j\pi}^{\infty}|\lambda|^{-d_\ell-d_m}\left||\widehat{\tau}_\infty(\lambda)|^2-4\,\mathbb{1}_{\mathbb{R}_+}(\lambda)\,\left|\widehat{h}_\infty(\lambda)\right|^2\right|d\lambda \leqslant C_s\pi^{-\beta}(2^j\pi)^{-2L}\int_{\pi}^{\infty}|\lambda|^{\beta-d_\ell-d_m+M+1}d\lambda\,.$$

The integral on the right hand side is bounded under (C-b).

The bound of the integral on $(-\infty, -\pi)$ is similar, with a distinction between $(-\infty, -2^j\pi)$ and $(-2^j\pi, \pi)$. 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, 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 and J.-F. Coeurjolly. Identification of the multivariate fractional Brownian motion. *Signal Processing, IEEE Transactions on*, 59(11):5152–5168, 2011.

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

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

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

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.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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