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 multivariate fractional Brownian motions. An application on a real neuroscience dataset is presented, where long-memory and brain connectivity are inferred.

**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 ( 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 (Baek et al., 2020; Kechagias and Pipiras, 2020). Following these modelings, the long-memory model studied in this article 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 $\mathbf{X}(t) = [X_1(t) \ldots X_p(t)]^T$, $t \in \mathbb{Z}$, $p \in \mathbb{N}$, $p \geq 1$, with long memory parameters $\mathbf{d} = (d_1, d_2, \ldots, d_p)$, $\mathbf{d} \in (-0.5, +\infty)^p$. 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 $\mathbf{D} = [\mathbf{d} + 1/2]$, we assume that the multivariate process $\text{Diag}\{(1 - \mathbb{L})^p, \ell = 1, \ldots, p\} \mathbf{X}$ is covariance stationary with a spectral density matrix given by

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

* Correspondence to: Irène Gannaz, Univ. Grenoble Alpes, CNRS, Grenoble INP, G-SCOP, 38000 Grenoble, France. Email: irene.gannaz@grenoble-inp.fr

© 2023 The Authors. *Journal of Time Series Analysis* published by John Wiley & Sons Ltd. This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

**LOCAL WHITTLE ESTIMATION WITH (QUASI-)ANALYTIC WAVELETS**

SOPHIE ACHARD$^*$ AND IRÈNE GANNAZ$^{b*}$

$^*$Univ. Grenoble Alpes, CNRS, Inria, Grenoble INP, LJK, Grenoble, France

$^{b*}$Univ. Grenoble Alpes, CNRS, Grenoble INP, G-SCOP, Grenoble, France

Received 28 June 2021; Accepted 5 August 2023

Keywords: Multivariate processes; long-memory; covariance; phase; wavelets; cerebral connectivity.

MOS subject classification: 60G22; 62M10; 62M15; 62H20; 92C55.
where $\odot$ denotes the Hadamard product, and $d_m^e = d_m - D_m \in (-0.5, 0.5)$ for all $m$. The process $(1 - L)^p X_m$ is said to have long-memory if $d_m^e \in (0, 0.5)$, and to be anti-persistent if $d_m^e \in (-0.5, 0)$ (see for instance Lobato (1999), and Shimotsu (2007)). For simplicity of notation, we will use the term long-memory parameters $d$ throughout the article.

The function $f(\cdot)$ is defined by

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

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

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

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

In particular, our definition agrees with the one given in Kechagias and Pipiras (2015) if $D_m = 0$ for all $m = 1, \ldots, p$. Definition 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), and Achard and Gannaz (2016). Indeed, the matrix $\Theta$ can be written as,

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

with $\Omega = (\Omega_{\ell,m})_{\ell,m=1,...,p}$ a real symmetric non-negative semi-definite matrix and $\Phi = (\phi_{\ell,m})_{\ell,m=1,...,p}$ an anti-symmetric matrix. Let the bar above denote the conjugate operator. The matrix $\Theta$ satisfies $\Theta^T = \Theta$ since $f^S(\cdot) = \overline{f(\cdot)}$. We will use $\|\Omega\|$ to denote the infinity norm, that is, $\|\Omega\| = \max_{\ell,m=1,...,p} \Omega_{\ell,m}$. In Lobato (1997), Shimotsu (2007), and 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, 1995b) and semiparametric estimation based on Whittle approximation (Künsch, 1987; Robinson, 1995a). 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 others 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 close 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 multivariate fractional Brownian motions. Additional simulations, on vector ARFIMA linear models, are provided in Supplementary material. Section 6 presents an empirical application on neuroscience data. The detailed proofs are provided in Supplementary material.

2. TRANSFORM OF THE MULTIVARIATE PROCESS

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

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

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

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

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

$$
\hat{\varphi}(\lambda) = \hat{\varphi}_{h}(\lambda) + i \ \hat{\varphi}_{g}(\lambda) \quad \text{and} \quad \hat{\psi}(\lambda) = \hat{\psi}_{h}(\lambda) + i \ \hat{\psi}_{g}(\lambda).
$$

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, $\hat{\psi}_{g}(\lambda) = -i \ \text{sign}(\lambda) \ \hat{\psi}_{h}(\lambda)$, for all $\lambda \in \mathbb{R}$, where $\text{sign}(\lambda)$ 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) and 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) and Achard et al. (2020) for a fuller description of the construction of the wavelets and of their properties.

Let $\hat{d}_{L}(\lambda), \ \lambda \in \mathbb{R}$, be defined by

$$
\hat{d}_{L}(\lambda) = e^{i \lambda (-L/2+1/4)} \left[ \cos \left( \lambda/4 \right)^{2L+1} + i \ (-1)^{L+1} \sin \left( \lambda/4 \right)^{2L+1} \right].
$$
Next, filters ($\hat{h}^{(L)}$, $\hat{h}^{(H)}$, and ($\hat{g}^{(L)}$, $\hat{g}^{(H)}$) are defined by
\begin{align}
\hat{h}^{(L)}(\lambda) &= 2^{-M+1/2} (1 + e^{-i\lambda})^M \hat{q}_{LM}(\lambda) \hat{d}_L(\lambda) \quad \text{and} \quad \hat{h}^{(H)}(\lambda) = \hat{h}^{(L)}(\lambda + \pi)e^{-i\lambda}, \quad (4) \\
\hat{g}^{(L)}(\lambda) &= 2^{-M+1/2} (1 + e^{-i\lambda})^M \hat{q}_{LM}(\lambda) \hat{d}_L(\lambda)e^{-i\lambda}L \quad \text{and} \quad \hat{g}^{(H)}(\lambda) = \hat{g}^{(L)}(\lambda + \pi)e^{-i\lambda}, \quad (5)
\end{align}
with $\hat{q}_{LM}(\lambda)$ a real polynomial of $(e^{-i\lambda})$ such that $\hat{q}_{LM}(0) = 1$. Observe that the normalization of the filters is different from Achard et al. (2020).

Common-factor wavelets are introduced with $\hat{q}_{LM}(\cdot)$ such that the wavelet decomposition satisfies the perfect reconstruction condition. This condition is classically used for deriving wavelet bases 21 different from Achard et al. (2020). The proof is given in Supplementary material.

Concerning the functions $\hat{q}_{LM}(\cdot)$, the pair $(\phi(\cdot), \psi(\cdot))$ will be denoted by CFW-PR(M,L). If $\hat{q}_{LM}(\cdot)$ is a constant polynomial equal to 1, $(\phi(\cdot), \psi(\cdot))$ will be denoted by CFW-C(M,L) filters.

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

**Proposition 2.** Let $M, L$ be strictly positive integers. The functions $\phi(\cdot)$, 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 Supplementary material.

Concerning the functions $\phi(\cdot)$, 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 $\hat{q}_{LM}(\cdot)$ real polynomial of $(e^{-i\lambda})$,
\begin{align}
\hat{\psi}(\lambda) &= \hat{\psi}_h(\lambda) + i \hat{\psi}_g(\lambda) = (1 - e^{i\eta(\lambda)}) \hat{\psi}_h(\lambda), \\
\text{with} \quad a_L(\lambda) &= 2(-1)^L \tan(2\pi/2L+1(\lambda/4)) \quad \text{and} \quad \eta(\lambda) = -a_L(\lambda/2 + \pi) + \sum_{j=1}^{\infty} a_L(2^{-j-1}\lambda). \quad (7)
\end{align}

Additionally, for all $\lambda \in \mathbb{R}$,
\begin{align}
\left| \hat{\psi}_h(\lambda) + i \hat{\psi}_g(\lambda) - 2\mathbb{I}_{\mathbb{R}^+}(\lambda) \hat{\psi}_h(\lambda) \right| = U_L(\lambda) \left| \hat{\psi}_h(\lambda) \right|.
\end{align}
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 (7), we adopt the convention that $\arctan(\pm \infty) = \pm \pi/2$ so that $\alpha_j(\cdot)$ 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 $\{W_{j,k}, j \geq 0, k \in \mathbb{Z}\}$ denote the wavelet coefficients of the process $X$ associated with 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 $X$ are defined by

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

where $\hat{X}(t) = \sum_{k \in \mathbb{Z}} X(k) \varphi(t - k)$. Given any $j \geq 0$ and any $k \in \mathbb{Z}$, $W_{j,k}$ is a $p$-dimensional vector $W_{j,k} = (W_{j,k}(1), W_{j,k}(2), \ldots, W_{j,k}(p))^T$ where $W_{j,k}(a) = \int_{\mathbb{R}} \hat{X}(t) \psi_{j,k}(t) dt$, $a = 1, \ldots, p$. Throughout the article, we adopt the convention that large values of the scale index $j$ correspond to coarse scales (low frequencies).

We will consider the behavior of $\text{Cov}(W_{j,k})$, defined as follows

$$\text{Cov}(W_{j,k}) = \mathbb{E} \left[ W_{j,k} W_{j,k}^T \right] = \int_{-\infty}^{\infty} f(\lambda) \left| \hat{f}(\lambda) \right|^2 d\lambda,$$

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

In practice, a finite number of observations of the process $X$ are available, $X(1), X(2), \ldots, X(N)$. 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 $W_{j,k}$ 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 $W_{j,k}$ 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 $\{\hat{W}_{j,k}, j \geq 0, k \in \mathbb{Z}\}$ denote the wavelet coefficients evaluated from $X(1), X(2), \ldots, X(N)$ by $\hat{W}_{j,k} = \int_{\mathbb{R}} \hat{X}(t) \psi_{j,k}(t) dt$, $j \geq 0, k \in \mathbb{Z}$, where $\hat{X}(t) = \sum_{k=1}^{N} X(k) \varphi(t - k)$. Then, at each scale $j$, the number of coefficients $n_j$ such that $\hat{W}_{j,k} = W_{j,k}$ 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.
results obtained in Gannaz et al. (2017) are summarized. We begin with a bivariate ARFIMA(0, d, 0) process defined as
\[ X_{\ell}(k) = (1 - L)^{-d_{\ell}} u_{\ell}(k), \quad \ell = 1, 2, \ k \in \mathbb{Z}, \] (8)

where \( L \) is a lag operator and \( \left( u_{\ell}(k) \right) \) i.i.d. with distribution \( \mathcal{N} \left( \left( \begin{array}{c} 0 \\ 0 \end{array} \right), \Omega \right) \), where \( \Omega = \left( \begin{array}{cc} 1 & 0.8 \\ 0.8 & 1 \end{array} \right) \). 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 \( d \) 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 \( X(1), \ldots, X(2^j) \), with \( J = 12 \). For each scale \( j \geq 0 \), we evaluate the sample wavelet covariances as \( \hat{\Sigma}(j) = \frac{1}{n_j} \sum_{k=0}^{n_j-1} W_{jk}(1) \bar{W}_{jk}(2) - \left( \frac{1}{n_j} \sum_{k=0}^{n_j-1} W_{jk}(1) \right) \left( \frac{1}{n_j} \sum_{k=0}^{n_j-1} \bar{W}_{jk}(2) \right) \), and the wavelet sample correlations as \( \hat{\Sigma}_{1,2}(j) / \sqrt{\hat{\Sigma}_{1,1}(j) \hat{\Sigma}_{2,2}(j)} \).

Figure 1 shows the behavior of sample wavelet correlations with respect to scale \( j \) over 100 realizations of \((X(1), \ldots, X(2^j))\). First observe that for real wavelets (left column), the wavelet sample covariance \( \hat{\Sigma}_{1,2}(j) \) 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 Cov(Wj,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 $\hat{q}_{L,M}$ satisfying (6) is available.

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

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

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

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

**Proposition 5.** Let $X$ be a $p$-multivariate long range dependent process with long memory parameters $d_1, \ldots, d_p$ with generalized spectral density $f(\cdot)$ satisfying (M-1) with short-range behavior (M-2). Consider $\{W_{j,k}(\ell), (j,k) \in \mathbb{Z}, \ell = 1, \ldots, p\}$ the wavelet coefficients obtained with CFW-C(M,L) filters. Suppose that (C-a) hold. Then we have, for all $j \geq 0, k \in \mathbb{Z},$

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

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

The proof is given in Supplemental material.

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 Supplemental material). 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 $|\hat{\psi}(\lambda)|^2$. With real wavelets, the approximation is given in Proposition 3 of Achard and Gannaz (2016). Since $|\hat{\psi}(\lambda)|^2$ 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

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

It is straightforward to check that parameters $\{\Omega_{\ell,m}, \phi_{j,m}\}$ are not identifiable. Estimation can be derived in the case of a parametric phase, typically $\phi_{j,m} = \frac{\pi}{2}(d_\ell - d_m)$ (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 $X$ be a $p$-multivariate long range dependent process with long memory parameters $d_1, \ldots, d_p$ with generalized spectral density $f(\cdot)$ satisfying (M-1) and (M-2). Consider $\{W_{j,k}(\ell), (j,k) \in \mathbb{Z}, \ell = 1, \ldots, p\}$ the wavelet coefficients obtained with CFW-C(M,L) filters. Suppose that (C-a) holds and that $L$ goes to infinity, with $L2^{-2j} \to 0$ when $j$ goes to infinity.

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

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

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

---

J. Time Ser. Anal. 45: 421–443 (2024) © 2023 The Authors. Wileyonlinelibrary.com/journal/jtsa

DOI: 10.1111/jtsa.12719 Journal of Time Series Analysis published by John Wiley & Sons Ltd.
The proof is given in Supplementary material.

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

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

with $G_{\ell, m} = \Theta_{\ell, m} K(d_{\ell} + d_m)$ and $K(\delta) = 4 \int_0^{\infty} |\lambda|^{-\delta} |\hat{\varphi}_h(\lambda)|^2 d\lambda.$ (11)

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 $\{W_{j,k}, j_0 \leq j \leq j_1, k \in \mathbb{Z}\}$. 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 $L = o(N)$.

#### 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)

$$\mathcal{L}(G, d) = \frac{1}{n} \sum_{j=j_0}^{j_1} n_j \log \det(G_j(d) G J_j(d)) + \sum_{k=0}^{n_j} W_{j,k}^T (G_j(d) G J_j(d))^{-1} W_{j,k}$$
WHITTLE ESTIMATION WITH (QUASI-)ANALYTIC WAVELETS

429

Figure 2. Boxplots of normalized sample covariances between wavelet coefficients at different scales for the bivariate process defined in (8). Let $V_j = (2^{-j}d_1 \var{W_{j,k}}, k \in \mathbb{Z}) 2^{-jd_1 - jd_2} \text{Cov}((W_{j,k}, W_{k,j}), k \in \mathbb{Z}) 2^{-jd_2} \var{W_{j,k}}$, $k \in \mathbb{Z})$. The first row gives the sample version of $V_{11}$, the second row gives the sample version of $V_{12}$ and the third row gives the sample version of $V_{22}$. 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 approximation given by Proposition 5.

where $\Gamma(d)$ is the diagonal matrix with diagonal entries $2^{-j}d_1$, $..., 2^{-j}d_p$, and $G$ 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

$$\mathcal{L}(G, d) = \frac{1}{n} \sum_{j=0}^{n_j} [\eta_j \log \det(\Gamma_j(d) G \Gamma_j(d)) + \text{trace} \left((\Gamma_j(d) G \Gamma_j(d))^{-1} I(j)\right)],$$

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

Note that when $G$ is a positive definite Hermitian matrix, for all $j \geq 0$ and for all $d \in (-0.5, \infty)^p$, $\det(\Gamma_j(d) G \Gamma_j(d))$ is real and strictly positive and $\text{trace} \left((\Gamma_j(d) G \Gamma_j(d))^{-1} I_j\right)$ is real. The objective function $\mathcal{L}(G, d)$ is hence well-defined for $G$ in the set of Hermitian matrices and for all $d \in \mathbb{R}^p$, and takes its values in $\mathbb{R}$.
Differentiating expression (12) with respect to the matrix $G$ yields

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

Some keys for complex matrix differentiation can be found in Hjorungnes and Gesbert (2007). Hence, the minimum for fixed $d$ is attained at

$$\hat{G}(d) = \frac{1}{n} \sum_{j=j_0}^{j_1} \Gamma_j(d)^{-1} I(j) \Gamma_j(d)^{-1}.$$

In Shimotsu (2007) the resulting objective function only depends of $d$ 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 $G(d)$ and complex valued matrices $\Gamma_j(d)$, including the phases $(\phi_{\ell,m})_{\ell=1,\ldots,p}$, are considered. $G(d)$ and $\Gamma_j(d)$ are estimated in a second step, together with parameter $d$. They minimize the objective function obtained when replacing $G$ by $\hat{G}(d)$ 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 $G$.

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

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

Since $\Gamma_j(d) = \text{Diag}(2^{-jd})$, we obtain

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

The vector of the long-memory parameters $d$ is estimated by $\hat{d} = \arg\min_d R(d)$. 

---

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

$$\hat{\Theta}_{\ell,m} = \frac{\hat{G}_{\ell,m}(\hat{d})}{K(\hat{d}_\ell + \hat{d}_m)}.$$ 

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 $\{I(\ell),\, \ell \geq 0\}$. 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 $\{I(\ell),\, \ell \geq 0\}$ 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,

$$\text{for all } \ell, m = 1, \ldots, p, \sup_{N, j \geq 0} \left| \frac{\text{Var}(I_{\ell,m}(j))}{n_j 2^{2j(\delta_\ell + \delta_m)}} \right| < \infty.$$ 

Let $d^0$, $G^0$ 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 $\log(N)(2^{-j_0 \beta} + N^{-1/2}2^{h_1/2}) \to 0$ and $j_0 < j_1 \leq j_N$.

Consider CFW-C(M,L) filters with $M \geq 2$ and $2^{-2h_0}L + N^{-1/2}L + \log(N)^3 L^{-M-1} \to 0$.

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

$$\hat{d} - d^0 = O_p(L2^{-2h_0} + \log(N)L^{-M-1} + 2^{-j_0 \beta} + N^{-1/2}2^{h_1/2}),$$

$$\hat{G}_{\ell,m}(\hat{d}) - G_{\ell,m}(d^0) = O_p(\log(N)(L2^{-2h_0} + \log(N)L^{-M-1} + 2^{-j_0 \beta} + N^{-1/2}2^{h_1/2})),$$

$$\hat{\Theta}_{\ell,m} - \Theta^0_{\ell,m} = O_p(\log(N)(L2^{-2h_0} + \log(N)L^{-M-1} + 2^{-j_0 \beta} + N^{-1/2}2^{h_1/2})).$$

Taking $2^{h_0} = N^{1/(1+2\beta)}$ and $L = N^{\frac{\beta}{1+2\beta}}$, $\hat{d} - d^0 = O_p(N^{-\beta/(1+2\beta)}).$

Elements of proof are given in Supplementary material.

The convergence rate $\hat{d} - d^0 = O_p(N^{-\beta/(1+2\beta)})$ is optimal in minimax sense (Giraitis et al., 1997).

The condition on the scales used in the estimation is $\log(N)^3(2^{-j_0 \beta} + N^{-1/2}2^{h_1/2}) \to 0$. 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 $2^{-2h_0}L + N^{-1}L + \log(N)^3 L^{-M-1} \to 0$. 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 $2^j N^{-3} L \to 0$ 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 $j_0, N^{-3} 2^j L \to 0$, when $j_1 = j_0 + \Delta$, with $\Delta < \infty$. The condition $\log (N)^3 L^{-M-1} \to 0$ guarantees that $L$ is high enough for the analytic approximation to be satisfactory. Alternatively, $L$ should not be too high, and condition $2^{-2j_0} L \to 0$ 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 $X$.

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

$$ \forall t \in \mathbb{Z}, \ (1 - \mathbb{I}) D X_a(t) = \sum_{u \in \mathbb{Z}} A_{a,b}^{(D)}(t + u) \varepsilon(u) $$

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

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

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

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

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

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

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

Then,

* $\sqrt{n}(\hat{d} - d^0)$ converges in distribution to a centered Gaussian distribution with a variance $V^{(d)}(\Delta)$ defined in Supplementary material, equation (E.4).
• \( \text{vec} \left( \sqrt{n} \left( \hat{G}(d) - G^0 \right) \right) \) converges in distribution to a centered Gaussian distribution with a variance \( V^G(\Delta) \) defined in Supplementary material, equation (E.5).

The proof is very similar to the one of Gannaz (2023). Some points are detailed in Supplementary material.

The highest scale is \( j_1 = j_2 + \Delta \). The theorem distinguishes the cases \( \Delta < \infty \) and \( \Delta = \infty \). Note that when \( \Delta < \infty \), the condition \( N^{-1/2}2^j L \to 0 \) is equivalent to \( N^{-1/2}2^j L \to 0 \). Hence, the condition on \( L \) writes as \( \log(N^3) N^{1/2}2^{-\nu/2} \left(L2^{-2\nu} + L^{-M-1}\right) \to 0 \).

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

We verify the accuracy of the covariance approximation given in Proposition 6 and the consistency of the parameters estimates provided in Theorem 7 on simulated data. We consider 1000 Monte Carlo simulations of bivariate long-memory processes \( X \) observed at \( X(1), \ldots, X(N) \) 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 \( d \) to the one given by real wavelets, namely Daubechies’ wavelets with four vanishing moments.

The estimated parameters are \( d = (d_1, d_2) \), the magnitude of the long-run covariance \( \Omega \), 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, \( \text{RMSE} = \sqrt{\text{bias}^2 + \text{std}^2} \).

We consider multivariate fractional Brownian motions (mFBM). Additional results on vector ARFIMA models, where a linear representation exists, are provided in Supplementary material, in Section F. 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). A comparison with a Fourier-based local Whittle procedure is displayed in Supplementary material, on vector ARFIMA models, when the simulated processes are stationary.

5.1. Multivariate Fractional Brownian Motions

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

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

Another usual parametrization is the one with Hurst parameters, equal to \( d - 1/2 \).

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

\[
\sigma_{\ell m} = \mathbb{E}[X_\ell(1)^2]^{1/2} \\
r_{\ell m} = r_{m,\ell} = \text{Cor}(X_\ell(1), X_m(-1))
\]
\[ \eta_{\ell,m} = -\eta_{m,\ell} = \frac{(\text{Cor}(X_\ell(1), X_m(-1)) - \text{Cor}(X_\ell(-1), X_m(1)))}{c_{\ell,m}} \]

with

\[ c_{\ell,m} = \begin{cases} 
2(1 - 2^{d_\ell + d_m}) & \text{if } d_\ell + d_m \neq 1, \\
2 \log(2) & \text{if } d_\ell + d_m = 1,
\end{cases} \]

where \( \text{Cor}(X_1, X_2) \) denotes the Pearson correlation between variables \( X_1 \) and \( X_2 \). The quantities \( (\eta_{\ell,m})_{\ell,m=1,\ldots,p} \) measure the dissymmetry of the process. A mFBM is time reversible if the distribution of \( X(-t) \) is equal to the distribution of \( X(t) \) for every \( t \). Didier and Pipiras (2011) established that zero-mean multivariate Gaussian stationary processes \( X \) is equivalent to \( \mathbb{E}[X_\ell(t)X_m(s)] = \mathbb{E}[X_\ell(s)X_m(t)] \) for all \( (s, t) \), which corresponds to the definition of time reversibility used in Kechagias and Pipiras (2015). A mFBM is time-reversible if and only if \( \eta_{\ell,m} = 0 \) for all \( (\ell, m) \).

Coeurjolly et al. (2013) characterize the spectral behavior of the increments of a mFBM. If \( f^{(1,1)}_{\ell,m}(\cdot) \) denotes the cross-spectral density of \( \{(1 - L)X_\ell(t), (1 - L)X_m(t), t \in \mathbb{Z} \} \), then

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

with

\[ \Omega_{\ell,m} = \begin{cases} 
\sigma_\ell \sigma_m \Gamma(d_\ell + d_m) \left( r^2_{\ell,m} \cos^2 \left( \frac{\pi}{2} (d_\ell + d_m) \right) + \eta^2_{\ell,m} \sin^2 \left( \frac{\pi}{2} (d_\ell + d_m) \right) \right)^{1/2} & \text{if } d_\ell + d_m \neq 2 \\
\sigma_\ell \sigma_m \Gamma(d_\ell + d_m) \left( r^2_{\ell,m} + \eta^2_{\ell,m} \right)^{1/2} & \text{if } d_\ell + d_m = 2
\end{cases} \]

\[ \phi_{\ell,m} = \begin{cases} 
\text{atan} \left( \frac{\eta_{\ell,m}}{\sqrt{r_{\ell,m}}} \right) & \text{if } d_\ell + d_m \neq 2 \\
\text{atan} \left( \frac{\eta_{\ell,m}}{\eta_{m,\ell}} \right) & \text{if } d_\ell + d_m = 2
\end{cases} \]

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

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

We consider two mFBM, both with parameters \( \sigma_1 = \sigma_2 = 1 \) and \( d = (1, 1, 2) \).

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

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

Case 2. \( \eta_{1,2} = -0.6, r_{1,2} = 0.2 \).

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

Simulations were done using \( \mathbb{R} \) functions provided by J-F Coeurjolly at https://sites.google.com/site/jfc/software.

### 5.2. Simulation Results

Figure 4 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).
(a) Case 1

![Boxplot of Re(Corr(Wj.))](image)

(b) Case 2

![Boxplot of Re(Corr(Wj.))](image)

Figure 4. Boxplots of sample 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 \( r_K = K(d_1 + d_2)/\sqrt{K(2d_1)K(2d_2)} \).

We now consider the local Whittle estimation of the parameters. Based on the discussion of Supplementary material, Section F, and in Figure 4, we fix \( j_0 = 4 \). We refer to Section F of Supplementary material for a discussion on the influence of parameter \( j_0 \).

Tables I and II highlight the good behavior of the estimation of long-memory parameters \( d \), 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 \( j_0 \).

Tables III and IV give the results obtained for the estimation of the covariance structure, that 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 estimation of \( \rho \) and of \( \phi \).
Table I. Results for the estimation of long-memory parameters \( d \) with CFW-PR(4,4) filter on mFBMs

| Case | \( d \) | Bias | Std | RMSE | Ratio PR/real |
|------|------|------|-----|------|---------------|
| 1    | 1    | 0.0065 | 0.0464 | 0.0469 | 1.5577 |
| 1.2  | 0.0059 | 0.0475 | 0.0478 | 2.0886 |
| 1    | 0.0051 | 0.0510 | 0.0513 | 1.6812 |
| 1.2  | 0.0035 | 0.0515 | 0.0516 | 1.9884 |

Note: 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 II. Results for the estimation of long-memory parameters \( d \) with CFW-C(4,4) filter on mFBMs

| Case | \( d \) | Bias | Std | RMSE | Ratio C/PR |
|------|------|------|-----|------|------------|
| 1    | 1    | 0.0155 | 0.0409 | 0.0437 | 0.9323 |
| 1.2  | 0.0133 | 0.0402 | 0.0423 | 0.8849 |
| 1    | 0.0177 | 0.0448 | 0.0482 | 0.9397 |
| 1.2  | 0.0116 | 0.0473 | 0.0487 | 0.9441 |

Note: Hyperparameter \( j_0 \) satisfies \( j_0 = 4 \). C/PR denotes the ratio between the RMSE given by CFW-C(4,4) filter and the RMSE given by CFW-PR(4,4) filter.

Table III. Results for the estimation of matrices \( \Theta \) with CFW-PR(4,4) filter on mFBMs

| Case | \( \Omega_{1,1} \) | Bias | Std | RMSE |
|------|------------------|------|-----|------|
| 1    | \(-0.1999\)      | 0.1544 | 0.2526 |
| 1.2  | \(-0.1592\)      | 0.0925 | 0.1841 |
| 1    | \(-0.2471\)      | 0.1504 | 0.2892 |
| 1.2  | \(-0.02048\)     | 0.1654 | 0.2633 |
| 1    | \(-0.0647\)      | 0.0501 | 0.0818 |
| 1.2  | \(-0.2508\)      | 0.1588 | 0.2969 |
| 1    | \(0.0971\)       | 0.0243 | 0.1001 |
| 1.2  | \(0.0039\)       | 0.0526 | 0.0528 |
| 1    | \(0.0097\)       | 0.0245 | 0.1059 |
| 1.2  | \(0.0043\)       | 0.0493 | 0.0995 |
| 1    | \(0.0043\)       | 0.0493 | 0.0995 |
| 1.2  | \(0.0043\)       | 0.0493 | 0.0995 |
| 1    | \(-0.0967\)      | 0.0434 | 0.1059 |
| 1.2  | \(-0.0087\)      | 0.1549 | 0.1551 |
| 1    | \(0.0039\)       | 0.0526 | 0.0528 |
| 1.2  | \(0.0097\)       | 0.0245 | 0.1059 |
| 1    | \(0.0043\)       | 0.0493 | 0.0995 |
| 1.2  | \(0.0043\)       | 0.0493 | 0.0995 |

Note: Hyperparameter \( j_0 \) satisfies \( j_0 = 4 \).

Table IV. Results for the estimation of matrices \( \Theta \) with CFW-C(4,4) filter on mFBMs

| Case | \( \Omega_{1,1} \) | Bias | Std | RMSE | Ratio C/PR |
|------|------------------|------|-----|------|------------|
| 1    | \(-0.1548\)      | 0.1501 | 0.2156 | 0.8537 |
| 1.2  | \(-0.1305\)      | 0.0865 | 0.1565 | 0.8501 |
| 1    | \(-0.2088\)      | 0.1365 | 0.2495 | 0.8625 |
| 1.2  | \(-0.0965\)      | 0.0245 | 0.0995 | 0.9940 |
| 1    | \(0.0043\)       | 0.0493 | 0.0949 | 0.9376 |
| 1.2  | \(-0.1427\)      | 0.1632 | 0.2168 | 0.8235 |
| 1    | \(-0.052\)       | 0.0527 | 0.074 | 0.9043 |
| 1.2  | \(0.0936\)       | 0.045 | 0.1039 | 0.9805 |
| 1    | \(-0.0061\)      | 0.156 | 0.1561 | 1.0064 |
| 1.2  | \(0.0936\)       | 0.045 | 0.1039 | 0.9805 |

Note: 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.
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.

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://doi.org.10.5281/zenodo.2452871 (Becq et al., 2020a, 2020b).

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 four 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, $\mathbf{d}$,
- the magnitude of the correlations, $\varepsilon = \{\hat{\rho}_{\ell,m}, 1 \leq \ell < m \leq p\}$ with $\hat{\rho}_{\ell,m} = \frac{\hat{\Omega}_{\ell,m}}{\sqrt{\hat{\Lambda}_{\ell,m} \hat{\Lambda}_{m,m}}}$,
- the phases, $\Phi = \{\hat{\phi}_{\ell,m}, 1 \leq \ell < m \leq p\}$.

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 5 shows the empirical distribution of the estimated empirical estimators $\mathbf{d}$. 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 Figures 6 and 7. 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...
Figure 5. Plot of the empirical distribution of the long memory parameters $\hat{d}$ obtained for the four groups of rats. Each color corresponds to a rat.

Figure 6. Plot of the empirical distribution of the correlation magnitudes $\hat{\rho}$ obtained for the four groups of rats. Each color corresponds to a rat.
are weaker. These 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 7 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 7 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 8 shows the distributions of the estimated phases \(\phi\) corresponding to magnitudes satisfying \(\rho > 0.3\) (this choice is motivated by the observation on the support of dead rats’ correlations above). 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 9 illustrates the graphs obtained for the four different groups.
Figure 8. Plot of the empirical distribution of the phases obtained for the four groups of rats after first thresholding the correlations. Only phases associated to correlations with a magnitude higher than 0.3 are considered. Each color corresponds to a rat. For dead rats, a bar plot is provided rather than a density plot due to the low number of values.

Figure 9. Plot of the average graphs with correlations and phases obtained for four 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|φ*|, blue edges correspond to negative mean phases lower than −1.1|φ*|, and gray edges to mean phases between −1.1|φ*| and 1.1|φ*|. The quantities φ* are equal to φ* = −\(\frac{\pi}{2}(d_f - d_m)\).
We have colored each edge according to the mean phase when it satisfies $|\phi_{\ell,m}| > 1.1|\phi^*_{\ell,m}|$ where $\phi^*_{\ell,m} = -\frac{\pi}{2}(d_{\ell} - d_m)$, $(\ell,m) \in \{1, \ldots, p\}^2$. The value $\phi^*_{\ell,m}$ corresponds to the phase of causal linear representations with power-law coefficients (Kechagias and Pipiras, 2015) 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 phase. A simulation study 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.

ACKNOWLEDGEMENTS

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 article. 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.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available in Zenodo at https://doi.org.10.5281/zenodo.2452871, reference number DOI https://doi.org.10.5281/zenodo.2452871, reference number DOI https://doi.org.10.5281/zenodo.2452871.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.
REFERENCES

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

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

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

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

Amblard PO, Coeurjolly JF, Lavancier F, Philippe A. 2013. Basic properties of the multivariate fractional Brownian motion. In Séminaires et congrès, Vol. 28; 65–87.

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

Becq G, Barbier E, Achard S. 2020a. 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 17(4):045012. https://doi.org/10.1088/1741-2552/ab9fec.

Becq G, Habet T, Collomb N, Faucher M, Delon-Martin C, Coizet V, Achard S, Barbier EL. 2020b. Functional connectivity is preserved but reorganized across several anesthetic regimes. NeuroImage 219(116):945. https://doi.org/10.1016/j.neuroimage.2020.116945.

Buxton RB. 2013. The physics of functional magnetic resonance imaging (fMRI). Reports on Progress in Physics 76(9):096601.

Coeurjolly JF, Amblard PO, Achard S. 2013. Wavelet analysis of the multivariate fractional Brownian motion. ESAIM: Probability and Statistics 17:592–604.

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

Gannaz I. 2023. Asymptotic normality of wavelet covariances and multivariate wavelet Whittle estimators. Stochastic Processes and their Applications 155:485–534.

Gannaz I, Achard S, Clausel M, Roueff F. 2017. Analytic wavelets for multivariate time series analysis. In Wavelets and Sparsity XVII, International Society for Optics and Photonics, Vol. 10394, SPIE, San Diego, CA; 103941X.

Genç R, Selduk F, Whitcher BJ. 2001. An Introduction to Wavelets and Other Filtering Methods in Finance and Economics Academic Press.

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

Giraitis L, Robinson PM, Samarov A. 1997. Rate optimal semiparametric estimation of the memory parameter of the Gaussian time series with long range dependence. Journal of Time Series Analysis 18:49–61.

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

Kechagias S, Pipiras V. 2015. Definitions and representations of multivariate long-range dependent time series. Journal of Time Series Analysis 36(1):1–25. https://doi.org/10.1111/jtsa.12086.

Kechagias S, Pipiras V. 2020. Modeling bivariate long-range dependence with general phase. Journal of Time Series Analysis 41(2):268–292.

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

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

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

Moulines E, Roueff F, Taqqu MS. 2008. A wavelet Whittle estimator of the memory parameter of a nonstationary Gaussian time series. The Annals of Statistics 36(4):1925–1956.

Nielsen FS. 2011. Local Whittle estimation of multi-variate fractionally integrated processes. Journal of Time Series Analysis 32(3):317–335.

Nielsen MØ, Frederiksen P. 2011. Fully modified narrow-band least squares estimation of weak fractional cointegration. The Econometrics Journal 14(1):77–120.

Robinson PM. 1994. Semiparametric analysis of long-memory time series. The Annals of Statistics 22(1):515–539.

Robinson PM. 1995a. Gaussian semiparametric estimation of long range dependence. The Annals of Statistics 23(5):1630–1661.

Robinson PM. 1995b. Log-periodogram regression of time series with long range dependence. The Annals of Statistics 23(3):1048–1072.

Robinson PM. 2008. Multiple local Whittle estimation in stationary systems. The Annals of Statistics 36(5):2508–2530.
Sela RJ, Hurvich CM. 2012. The averaged periodogram estimator for a power law in coherency. *Journal of Time Series Analysis* 33(2):340–363.

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

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

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

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

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