Multivariate selfsimilarity: Multiscale eigen-structures for selfsimilarity parameter estimation

Charles-Gérard Lucas[(1)], Gustavo Didier[(2)], Herwig Wendt[(3)], *IEEE Senior*, Patrice Abry[(1)], *IEEE Fellow*,

[(1)] ENSL, CNRS, Laboratoire de physique, F-69342 Lyon, France,
`{patrice.abry,charles.lucas}@ens-lyon.fr`
[(2)] Math. Dept., Tulane University, New Orleans, USA, `gdidier@tulane.edu`,
[(3)] IRIT, Univ. Toulouse, CNRS, Toulouse, France, `herwig.wendt@irit.fr`.

Abstract—Scale-free dynamics, formalized by selfsimilarity, provides a versatile paradigm massively and ubiquitously used to model temporal dynamics in real-world data. However, its practical use has mostly remained univariate so far. By contrast, modern applications often demand multivariate data analysis. Accordingly, models for multivariate selfsimilarity were recently proposed. Nevertheless, they have remained rarely used in practice because of a lack of available reliable estimation procedures for the vector of selfsimilarity parameters. Building upon recent mathematical developments, the present work puts forth an efficient estimation procedure based on the theoretical study of the multiscale eigenstructure of the wavelet spectrum of multivariate selfsimilar processes. The estimation performance is studied theoretically in the asymptotic limits of large scale and sample sizes, and computationally for finite-size samples. As a practical outcome, a fully operational and documented multivariate signal processing estimation toolbox is made freely available and is ready for practical use on real-world data. Its potential benefits are illustrated in epileptic seizure prediction from multi-channel EEG data.

Index Terms—Multivariate selfsimilarity, selfsimilarity parameter vector, wavelet transform, multiscale eigen-structure.

I. INTRODUCTION

Context. The concept of scale-free dynamics results from a change in paradigm: instead of being driven by a few characteristic time scales, thus targeted in estimation, temporal dynamics results from the interactions of a large continuum of time scales, hence all of equal relevance. Scale-free dynamics has been massively used in the past decades to successfully model temporal activities in a large variety of data stemming from several real-world applications very different in nature, ranging from natural phenomena — physics (hydrodynamic turbulence [1], astrophysics [2]), geophysics (rainfalls, [3], seismology [4]), biology (plant growths [5], body rhythms [6], heart rate [7], [8], neurosciences [9], [10], genomics [11]) — to human activity — Internet traffic [12], finance [13], city growth [14], art investigations [15].

However, scale-free dynamics has so far remained restricted to univariate settings only, mostly due to a lack of multivariate models and estimation tools. By contrast, in most applications, one same system is monitored by several sensors, which naturally calls for the joint (or multivariate) analysis of a collection of time series. Multivariate selfsimilarity with the practical and accurate estimation of the vector of selfsimilarity parameters thus constitutes the heart of the present work.

Related work. Fractional Brownian motion (fBm) [16], [17], the only Gaussian, selfsimilar, stationary-increment process, has quasi-exclusively been used in practice as the reference model for scale-free dynamics. The selfsimilarity parameter H is the quantity of interest in applications. Its accurate estimation has thus received considerable attention (cf. [17], [18] for reviews). Notably, multiscale (wavelet) representations have proven to yield accurate and reliable estimation procedures for H [19]: They rely on the crucial facts that the statistics of the wavelet coefficients behave as power laws with respect to scales, and that the exponents of these power laws are controlled by H. To account for the multivariate nature of data in applications, multivariate extensions of fBm were recently proposed based on gathering correlated fBm [20]. Multivariate multiscale-based representations were accordingly devised for the estimation of the resulting vector of selfsimilarity parameters. Nonetheless, these constructs were essentially based on component-wise univariate power law behavior [19], [21]. More recently, a richer model for multivariate selfsimilarity, operator fractional Brownian motion (ofBm), was proposed. Its versatility stems from a major change of perspective: there is no longer a one-to-one association either between each selfsimilarity parameter and each component, or between power laws and scale-free dynamics. Instead, the statistics of each component depend on the entire vector of selfsimilarity parameters [22]–[24]. Very recently, the principles of the estimation for the full vector of selfsimilarity parameters were theoretically discussed, based on the eigenvalues of multivariate wavelet spectra estimated at each scale [25], [26]. This paved the way for preliminary explorations [27] that lead to the core contributions of the present work: The construction of practical statistical signal processing tools for multivariate selfsimilarity analysis.

Goals, contributions and outline. The present work aims to define a reliable and theoretically well-grounded estimation procedure for the vector of selfsimilarity parameters in a multivariate setting and to establish its good performance both in theory and in practice. To that end, Section II recalls state-

of-the-art multivariate selfsimilarity modeling and analysis. As a first contribution, a tractable yet richer model for multivariate selfsimilarity is devised in Section III, and then compared to ofBm; its multivariate wavelet analysis is detailed. As a second and major contribution, an estimation procedure for the vector of selfsimilarity parameters is devised, by exploiting a bias-variance trade-off, and its asymptotic performance is theoretically studied (cf. Section IV). Finite-size estimation performance is studied empirically (from Monte Carlo simulations), together with the derivation of approximations to the covariance structure of the estimates, which is of critical importance for real-world applications (cf. Section V). The estimation performance of the proposed multivariate estimator is compared, in terms of biases, variances, MSE, and covariance structures, to an earlier multivariate estimator defined in [25], [26] and also to the univariate estimator defined in [19]. The potential benefits of the proposed eigen-wavelet estimation procedures are illustrated by means of epileptic seizure prediction from multichannel EEG data. A documented toolbox for both the synthesis of multivariate selfsimilar processes and for the estimation of the vector of selfsimilarity parameters is publicly available at github.com/charlesglucas/ofbm_tools.

II. MULTIVARIATE SELFSIMILARITY: STATE OF THE ART

A. Model: Operator Fractional Brownian Motion (ofBm)

1) Definition: Operator Fractional Brownian Motion, originally proposed in [22]–[24], is defined as a jointly Gaussian M-variate process via the harmonizable representation

$$\mathcal{B}_{\mathbf{H},\mathbf{A}}(t) := \int_{\mathbb{R}} \frac{e^{itf}-1}{if}\left(f_+^{-\mathbf{H}+\frac{1}{2}\mathbb{I}}\mathbf{A} + f_-^{-\mathbf{H}+\frac{1}{2}\mathbb{I}}\overline{\mathbf{A}}\right)\tilde{B}(\mathrm{d}f). \tag{1}$$

In (1), $i^2 = -1$, $\mathbb{I}$ denotes the $M \times M$ identity matrix, $\tilde{B}(\mathrm{d}f)$ a complex-valued Gaussian Hermitian random measure, such that $\mathbb{E}\tilde{B}(\mathrm{d}f)\tilde{B}(\mathrm{d}f)^\star = \mathrm{d}f$ (with $\star$ the Hermitian transposition). Also, $\mathbf{A}$ and $\mathbf{H}$ are two $M \times M$ complex- and real-valued matrices, respectively. Furthermore, hereinafter it is also assumed that [23]:

ASSUMPTION OFBM1: The Hurst matrix has the form

$$\mathbf{H} = \mathbf{W}\mathrm{diag}(H_1,\ldots,H_M)\mathbf{W}^{-1}, \tag{2}$$

with $\mathbf{W}$ a $M \times M$ real-valued invertible matrix, and where

$$0 < H_1 \leq \ldots \leq H_M < 1. \tag{3}$$

ASSUMPTION OFBM2: $\mathbf{A}\mathbf{A}^\star$ is positive definite, i.e., $\det \Re(\mathbf{A}\mathbf{A}^\star) > 0$.
ASSUMPTION OFBM3: $\mathbf{A}\mathbf{A}^\star$ is a real-valued matrix.

OFBM2 ensures that $\mathcal{B}_{\mathbf{H},\mathbf{A}}$ is a well-defined process ; OFBM3 restricts $\mathcal{B}_{\mathbf{H},\mathbf{A}}$ to the special case of *time-reversible* (or covariant under the change $t \to -t$) statistical properties ; $\mathbf{W}$ in OFBM1 is referred to as the mixing matrix and provides an important source of versatility of the model.

From Definition (1), the covariance matrix is obtained as:

$$\mathbb{E}\left[\mathcal{B}_{\mathbf{H},\mathbf{A}}(t)\mathcal{B}_{\mathbf{H},\mathbf{A}}(s)^\star\right] = \int_{\mathbb{R}} \frac{(e^{itf}-1)(e^{-isf}-1)}{|if|^2}$$
$$\left(f_+^{-\left(\mathbf{H}-\frac{1}{2}\mathbb{I}\right)}\mathbf{A}\mathbf{A}^\star f_+^{-\left(\mathbf{H}^\star-\frac{1}{2}\mathbb{I}\right)}\right.$$
$$\left. + f_-^{-\left(\mathbf{H}-\frac{1}{2}\mathbb{I}\right)}\overline{\mathbf{A}\mathbf{A}^\star}f_-^{-\left(\mathbf{H}^\star-\frac{1}{2}\mathbb{I}\right)}\right)\mathrm{d}f, \tag{4}$$

where $\mathbb{E}$ stands for the ensemble average.

2) Multivariate selfsimilarity: Importantly, ofBm satisfies the general multivariate selfsimilarity relation [23]

$$\{\mathcal{B}_{\mathbf{H},\mathbf{A}}(t)\}_{t\in\mathbb{R}} \overset{fdd}{=} \{a^{\mathbf{H}}\mathcal{B}_{\mathbf{H},\mathbf{A}}(t/a)\}_{t\in\mathbb{R}}, \quad \forall a > 0, \tag{5}$$

where $a^{\mathbf{H}} = \sum_{k\geq 0} \log^k a\, \mathbf{H}^k/k!$ and $\overset{fdd}{=}$ denotes equality for all finite-dimensional distributions. In essence, Eq. (5) states that the statistics of $\mathcal{B}_{\mathbf{H},\mathbf{A}}(t)$ are covariant under a change of time scale, $t \to t/a$, up to a change in amplitude $\mathcal{B} \to a^{\mathbf{H}}\mathcal{B}$. Importantly, this holds for any dilation factor $a > 0$. However, the factor controlling the change in amplitude $\to a^{\mathbf{H}}$ involves jointly all entries of the vector of selfsimilarity parameters $\underline{H} = (H_1,\ldots,H_M)$, through the mixing matrix $\mathbf{W}$.

Of particular pedagogical interest is the case when the scaling exponent matrix $\mathbf{H}$ is diagonal, $\mathbf{H} = \mathrm{diag}(H_1,\ldots,H_M)$ (equivalently, $\mathbf{W} = \mathbb{I}$): M-variate selfsimilarity in Eq. (5) then simplifies into a collection of M, component-wise, univariate selfsimilarity relations, each controlled by a single selfsimilarity parameter, i.e., $\forall m = 1,\ldots,M$ and $a > 0$:

$$\{(\mathcal{B}_{\mathbf{H},\mathbf{A}})_m(t)\}_{t\in\mathbb{R}} \overset{fdd}{=} \{a^{H_m}(\mathcal{B}_{\mathbf{H},\mathbf{A}})_m(t/a)\}_{t\in\mathbb{R}}. \tag{6}$$

B. Eigen-wavelet analysis

1) Multivariate wavelet spectrum: Let ψ_0 denote the so-called *mother wavelet* defined as a compact support, unit norm reference pattern $\int_{\mathbb{R}} \psi_0(t)^2\,\mathrm{d}t = 1$, characterized by its number of vanishing moments N_ψ, and satisfying regularity conditions quantified by the decay of its Fourier transform $\hat{\psi}_0$: There is $\alpha > 1$ such that $\sup_{f\in\mathbb{R}} \left|\hat{\psi}_0(f)\right|(1+|f|)^\alpha < +\infty$.

Let Y denote a $M \times N$ multivariate process, with M and N the number of components and the sample size. The multivariate Discrete Wavelet Transform coefficients at scale 2^j consist of a $M \times N$ matrix $D_Y(j,k) = \int_{\mathbb{R}} \psi(j,k)Y(t)\mathrm{d}t$, with $\psi(j,k) = 2^{-j/2}\psi_0(2^{-j}t - k)$ a collection of dilated and translated templates of ψ_0. For further details on wavelet transforms, the reader is referred, e.g., to [28].

The wavelet spectrum is defined as the collection of $M \times M$ scale-dependent random matrices:

$$\mathbf{S}_Y(2^j) = \frac{1}{n_j}\sum_{k=1}^{n_j} D_Y(2^j,k)D_Y(2^j,k)^\top. \tag{7}$$

2) Multivariate wavelet spectrum for ofBm: For $\mathcal{B}_{\mathbf{H},\mathbf{A}}$, it has been shown that wavelet coefficients reproduce selfsimilarity as [25], [26]:

$$\left\{D_{\mathcal{B}_{\mathbf{H},\mathbf{A}}}(2^j,k)\right\}_{k\in\mathbb{Z}} \overset{fdd}{=} \left\{2^{j(\mathbf{H}+\frac{1}{2}\mathbb{I})}D_{\mathcal{B}_{\mathbf{H},\mathbf{A}}}(2^0,k)\right\}_{k\in\mathbb{Z}}. \tag{8}$$

The wavelet spectrum thus becomes:

$$\mathbf{S}_{\mathcal{B}_{\mathbf{H},\mathbf{A}}}(2^j) \stackrel{d}{=} 2^{j(\mathbf{H}+\frac{1}{2}\mathbb{I})}\, S_{\mathcal{B}_{\mathbf{H},\mathbf{A}}}(2^0)\, 2^{j(\mathbf{H}^\top+\frac{1}{2}\mathbb{I})}. \quad (9)$$

When $\mathbf{W} \neq \mathbb{I}$, Eq. (9) shows that (the expectation of) each entry of $\mathbf{S}_{\mathcal{B}_{\mathbf{H},\mathbf{A}}}(2^j)$ generally consists of a mixture of power laws with respect to the scale 2^j. In other words,

$$\mathbb{E}\mathbf{S}_{m,m'}(2^j) = \sum_{k=1}^{M}\sum_{k'=1}^{M} \alpha_{k,k'}^{m,m'} 2^{j(H_k+H_{k'}+1)}, \quad (10)$$

with $\alpha_{k,k'}^{m,m'}$ a priori depending on both $\mathbf{H}$ and $\mathbf{A}$.

3) Estimation of the vector of selfsimilarity parameters: It has been shown in [25], [26] that each eigenvalue $\lambda_m(2^j)$ of $\mathbb{E}\mathbf{S}_{\mathcal{B}_{\mathbf{H},\mathbf{A}}}(2^j)$ behaves asymptotically as a power law with respect to the scale 2^j, where the scaling exponent is controlled by H_m,

$$\lambda_m(2^j) \sim \xi_m(2^0)\,(2^j)^{2H_m+1}, \quad \text{as } j \to +\infty. \quad (11)$$

Eq. (11) naturally suggests that the estimation of the vector of selfsimilarity parameters $\underline{H}$ can be performed by a collection of individual, for each $m = 1, \ldots, M$, linear regressions on the logarithms of the eigenvalues $\hat{\lambda}_m(2^j)$ of $\mathbf{S}_{\mathcal{B}_{\mathbf{H},\mathbf{A}}}(2^j)$ versus the logartihms of the scales 2^j:

$$\hat{H}_m^{\mathrm{M}} = \frac{1}{2}\left(\sum_{j=j_1}^{j_2} w_j \log_2 \hat{\lambda}_m(2^j) - 1 \right), \quad (12)$$

where $j \in (j_1, j_2)$ denote the range of octaves involved and the regression weights satisfy: $\sum_{j=j_1}^{j_2} jw_j = 1$ and $\sum_{j=j_1}^{j_2} w_j = 0$. The asymptotic estimation performance was thoroughly studied theoretically in [25], [26].

III. Multivariate Fractional Brownian Motion

A. Definition

For the specific *nonmixing* case ($\mathbf{W} = \mathbb{I}$), the selfsimilarity relation in Eq. (6) suggests, as a first contribution of this work, the construction of another type of multivariate selfsimilar process. It is hereinafter referred to as multivariate fractional Brownian motion (mfBm) $B_{\mathbf{\Sigma},\mathbf{W},\underline{H}}(t)$ and defined as a mixture of correlated univariate fractional Brownian motions:

$$B_{\mathbf{\Sigma},\mathbf{W},\underline{H}}(t) := \mathbf{W}B_{\mathbf{\Sigma},\underline{H}}(t)^\top, \quad (13)$$

where $\mathbf{W}$ denotes a $M \times M$ real-valued invertible matrix. Also, $B_{\mathbf{\Sigma},\underline{H}}(t) = [B_{\mathbf{\Sigma},H_1}(t), \ldots, B_{\mathbf{\Sigma},H_M}(t)]$ consists of a collection of M fBms, $B_{\mathbf{\Sigma},H_m}$, $m = 1, \ldots, M$, each with a possibly different selfsimilarity parameter H_m ($0 < H_m < 1$) and variance σ_m^2. In addition, these fBms are correlated via an $M \times M$ point matrix $\mathbf{\Sigma} = \mathrm{diag}(\sigma_1^2, \ldots, \sigma_M^2)\boldsymbol{\rho}\,\mathrm{diag}(\sigma_1^2, \ldots, \sigma_M^2)^*$, with $\boldsymbol{\rho}$ the matrix of pairwise correlations.

From the definition of univariate fBm [16], we can show that a mfBm $B_{\mathbf{\Sigma},\mathbf{W},\underline{H}}$ is a zero-mean multivariate Gaussian process with covariance function given by:

$$\mathbb{E}\left[B_{\mathbf{\Sigma},\mathbf{W},\underline{H}}(t)\, B_{\mathbf{\Sigma},\mathbf{W},\underline{H}}(s)^\top \right]$$
$$= \frac{1}{2}\mathbf{W}\left(|t|^{\mathrm{diag}(\underline{H})}\, \mathbf{\Sigma}\, |t|^{\mathrm{diag}(\underline{H})} \right.$$
$$+ |s|^{\mathrm{diag}(\underline{H})}\, \mathbf{\Sigma}\, |s|^{\mathrm{diag}(\underline{H})}$$
$$\left. - |t-s|^{\mathrm{diag}(\underline{H})}\, \mathbf{\Sigma}\, |t-s|^{\mathrm{diag}(\underline{H})} \right)\mathbf{W}^\top. \quad (14)$$
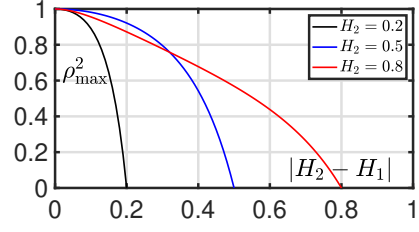

Fig. 1: **Interplay between the largest possible difference in selfsimilarity parameters and the largest (squared) correlation coefficients,** for different H_2 (bivariate setting).

In particular, the covariance matrix of mfBm simply reads $\mathbb{E}\left[B_{\mathbf{\Sigma},\mathbf{W},\underline{H}}(1)\, B_{\mathbf{\Sigma},\mathbf{W},\underline{H}}(1)^\top \right] = \mathbf{W}\mathbf{\Sigma}\mathbf{W}^\top$.

B. mfBm vs. ofBm

Comparing Eqs. (4) and (14), calculations not reported here yield the relations:

$$\mathbf{A}\mathbf{A}^* = \mathbf{W} \cdot (\mathbf{G} \odot \mathbf{\Sigma}) \cdot \mathbf{W}^\top \text{ and } \mathbf{H} = \mathbf{W}\,\mathrm{diag}(\underline{H})\,\mathbf{W}^{-1}. \quad (15)$$

In (15), $\mathbf{G}$ is a $M \times M$ matrix with entries $g_{m,m'} = \Gamma(H_m + H_{m'} + 1)\sin((H_m + H_{m'})\pi/2)/(2\pi)$, where Γ is the standard Gamma-Euler function and the symbol $\odot$ stands for the Hadamard matrix product. Eq. (15) shows that the mfBm $B_{\mathbf{\Sigma},\mathbf{W},\underline{H}}$ is a special case of the ofBm $\mathcal{B}_{\mathbf{H},\mathbf{A}}$, satisfying Assumptions OFBM1 to OFBM3. The mfBm $B_{\mathbf{\Sigma},\mathbf{W},\underline{H}}$ thus provides a model for multivariate selfsimilarity. It can be considered a practical counterpart of the formal ofBm model, yet better suited for *signal processing purposes* and applications to real-world data. It splits the abstract matrices $\mathbf{A}$ and $\mathbf{H}$ from the ofBm model into interpretable ingredients that are closer to applications: a vector of scaling exponents $\underline{H}$, a (premixing or intrinsic) covariance matrix $\mathbf{\Sigma}$, and a mixing matrix $\mathbf{W}$. In particular, Eq. (15) shows that mfBm $B_{\mathbf{\Sigma},\mathbf{W},\underline{H}}$ necessarily fulfills Assumption OFBM3 and thus implies by construction time-reversible statistics. An alternative permitting non time-reversible statistics had been considered in [20], yet not allowing for mixing hence less versatile for applications.

Furthermore, the mfBm model allows us to better grasp a key feature of multivariate selfsimilarity. Starting with the simple bivariate setting, $M = 2$, combining Assumption OFBM2 and Eq. (15) yields:

$$\rho_{12}^2 \leq \rho_{max}^2 := \frac{\Gamma(2H_1+1)\Gamma(2H_2+1)\sin(\pi H_1)\sin(\pi H_2)}{(\Gamma(H_1+H_2+1)\sin(\pi/2(H_1+H_2)))^2}. \quad (16)$$

This shows that the correlation between components and differences in selfsimilarity parameters cannot be chosen independently. This is illustrated in Fig. 1: The largest allowed (squared) cross-correlation amongst components decreases when the difference $|H_2 - H_1|$ increases, and conversely. These calculations extend to M-variate settings with larger (absolute values of) correlations amongst components restricting the largest achievable discrepancies $|H_m - H_{m'}|$ amongst entries of the vector of selfsimilarity parameters (see also [20]).

Finally, $B_{\mathbf{\Sigma},\underline{H}} \equiv B_{\mathbf{\Sigma},\mathbf{W}=\mathbb{I},\underline{H}}$ is obviously a special case of mfBm (cf., e.g., [19]–[21]). Yet, it displays no mixing across components, rendering the model less versatile for practical purposes.

C. Wavelet analysis of mfBm

For mfBm, Eqs. (8) and (9) take the specific forms

$$\left\{D_{B_{\boldsymbol{\Sigma},\mathbf{W},\underline{H}}}(2^j,k)\right\}_{k\in\mathbb{Z}} \overset{fdd}{=} \left\{\mathbf{W}\,2^{j(\underline{H}+\frac{1}{2})}D_{B_{\boldsymbol{\Sigma},\underline{H}}}(2^0,k)\right\}_{k\in\mathbb{Z}},$$
(17)

$$\mathbf{S}_{B_{\boldsymbol{\Sigma},\mathbf{W},\underline{H}}}(2^j) \overset{d}{=} \mathbf{W}\,2^{j(\underline{H}+\frac{1}{2})}\mathbf{S}_{B_{\boldsymbol{\Sigma},\underline{H}}}(2^0)\,2^{j(\underline{H}+\frac{1}{2})}\mathbf{W}^\top,$$
(18)

with $2^{j(\underline{H}+\frac{1}{2})} = \operatorname{diag}(2^{j(H_1+1/2)},\dots,2^{j(H_M+1/2)})$. This disentangles the contributions of the dilation vs. mixing operators to the shaping of the wavelet coefficients and spectrum.

Furthermore, Eqs. (10) and (11) obviously hold for mfBm. This permits to show, respectively, that both $\alpha_{k,k'}^{m,m'}$ and $\xi_m(2^0)$ depend on $\mathbf{W}$ and $\boldsymbol{\Sigma}$, but not on $\underline{H}$. An extension of the calculations in [25], [26] shows that $\xi_m(2^0)$ actually depends only on the entries $\left(\mathbf{S}_{B_{\boldsymbol{\Sigma},\mathbf{W},\underline{H}}}(2^0)\right)_{(m',m'')}$ for $m \le m', m'' \le M$ (see Remark A.1, Section A, supplementary material).

Finally and importantly, Eq. (11) demonstrates that the wavelet eigenvalue representations for both mfBm and ofBm disentangle the mixed power laws appearing in the wavelet spectrum $\mathbf{S}_{B_{\boldsymbol{\Sigma},\mathbf{W},\underline{H}}}(2^j)$. Thus, they restore the close relationship between power laws and scale-free dynamics, which had gotten altered by the mixing effect of the matrix $\mathbf{W}$. This motivates the estimation procedure $\hat{\underline{H}}^{\mathrm{M}}$ in Eq. (12), which thus applies to mfBm as well as to ofBm.

IV. SELFSIMILARITY PARAMETER VECTOR ESTIMATION

A. Repulsion bias-corrected multivariate estimation

1) Repulsion bias: $\hat{\underline{H}}^{\mathrm{M}}$ in Eq. (12) is based on eigenvalues computed from the wavelet spectra $\mathbf{S}_{\mathcal{B}_{\mathbf{H},\mathbf{A}}}(2^j)$, namely, on the empirical wavelet coefficient covariance matrices estimated at different scales. However, it is well known that, for finite-size sample estimation of covariance matrices, eigenvalues undergo the so-called *repulsion effect*: Estimated eigenvalues tend to be farther apart one from the others, compared to true eigenvalues, all the more as the sample size is decreasing compared to the number of components (cf. e.g., [29]).

In the wavelet spectrum representation, this *repulsion effect* is enhanced by the fact that empirical wavelet coefficient covariance matrices are estimated with sample size n_j that depends on analysis scales 2^j, essentially as $n_j = 2^{-j}N$. This induces scale-dependent repulsion biases in $\hat{\lambda}_m(2^j)$, and hence an overall bias in $\hat{\underline{H}}^{\mathrm{M}}$ [27]. The repulsion bias is illustrated in Fig. 2 (left), which displays $\log_2\hat{\lambda}_m(2^j)$ vs $\log_2 2^j = j$. It is based on a synthetic mfBm with $M = 6$ components of size $N = 2^{12}$, with $H_1 = \dots = H_6 = 0.6$, a randomly selected orthonormal mixing matrix $\mathbf{W} \ne \mathbb{I}$ and a correlation matrix $\boldsymbol{\Sigma}$ with off-diagonal coefficients set to 0.8. The five smallest curves should thus theoretically superimpose, whereas discrepancies of amplitudes in estimated eigenvalues, increasing with scales, can be observed.

2) Repulsion bias correction: Elaborating on [27], a novel estimation procedure for the estimation of $\underline{H}$ is devised and studied, hereinafter referred to as the *bias-corrected* estimator $\hat{\underline{H}}^{\mathrm{M,bc}} = (\hat{H}_1^{\mathrm{M,bc}},\dots,\hat{H}_M^{\mathrm{M,bc}})$. The main idea behind its construction lies in using the *same number* of wavelet coefficients in the wavelet spectrum estimation at each scale. As a result,
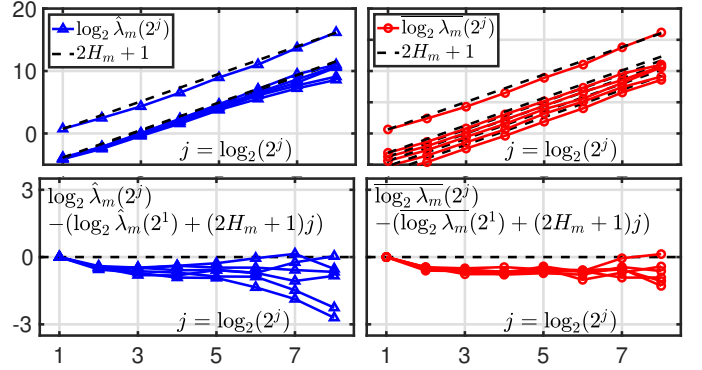

Fig. 2: **Repulsion bias in eigenvalue estimation.** Top: Logarithms of estimated eigenvalues $\log_2\hat{\lambda}_m(2^j)$ (left, blue) and $\overline{\log_2\lambda_m}(2^j)$ (right, red), as functions of the logarithms of the scales $\log_2 2^j = j$ for a synthetic mfBm with $M = 6$, $N = 2^{12}$, $\boldsymbol{\Sigma} \ne \mathbb{I}$, $\mathbf{W} \ne \mathbb{I}$ and $H_1 = \dots = H_6 = 0.6$, and $j_2 = 8$. Superimposed straight lines indicate the theoretical slopes $2H_m + 1$ (top, black). Bottom: Logarithms of estimated eigenvalues $\log_2\hat{\lambda}_m(2^j)$ (left, blue) and $\overline{\log_2\lambda_m}(2^j)$ (right, red) with theoretical linear behaviors subtracted.

the eigenvalues computed from the wavelet spectra at different scales 2^j undergo an *equivalent strength* repulsion effect, i.e., one that does not vary with the scale 2^j. To that end, let 2^{j_2} denote the largest analysis scale used in estimation (cf. Eq. (12)) and let n_{j_2} be the number of corresponding wavelet coefficients at that scale. At each scale $2^{j_1} \le 2^j \le 2^{j_2}$, a collection of wavelet spectra are computed from $\tau = 1,\dots,2^{j_2-j}$ consecutive non-overlapping windows of wavelet coefficients with same size n_{j_2}:

$$\mathbf{S}^{(\tau)}(2^j) := \frac{1}{n_{j_2}} \sum_{k=1+(\tau-1)n_{j_2}}^{\tau n_{j_2}} D_{\mathcal{B}_{\mathbf{H},\mathbf{A}}}(2^j,k)D_{\mathcal{B}_{\mathbf{H},\mathbf{A}}}(2^j,k)^T.$$
(19)

At each scale, the logarithms of the eigenvalues $\hat{\lambda}_m^{(\tau)}(2^j)$ computed from each wavelet spectrum $\mathbf{S}^{(\tau)}(2^j)$ are then averaged across windows:

$$\overline{\log_2\lambda_m}(2^j) := 2^{j-j_2} \sum_{\tau=1}^{2^{j_2-j}} \log_2\hat{\lambda}_m^{(\tau)}(2^j).$$
(20)

The impact of this bias correction procedure on the scaling of the eigenvalues is illustrated in Fig. 2 (right), which clearly shows that the theoretical slopes $2H_m + 1$ are much better reproduced by the $\overline{\log_2\lambda_m}(2^j)$ than by the classical $\log_2\hat{\lambda}_m(2^j)$. The bias-corrected multivariate estimates $\hat{H}_m^{\mathrm{M,bc}}$ are thus defined as linear regressions across (the logarithms of the) scales 2^j of the averaged log-eigenvalues $\overline{\log_2\lambda_m}(2^j)$:

$$\hat{H}_m^{\mathrm{M,bc}} = \frac{1}{2}\left(\sum_{j=j_1}^{j_2} w_j\overline{\log_2\lambda_m}(2^j) - 1\right).$$
(21)

The asymptotic estimation performance is studied theoretically in Section IV-B. The finite-size estimation performance is investigated numerically in Section V.

3) Nonmixing case: In the nonmixing case, i.e., when $\mathbf{W} = \mathbb{I}$, the selfsimilarity matrix reduces to $\mathbf{H} = \mathrm{diag}(\underline{H})$. While the multivariate estimators $\hat{H}_m^{\mathrm{M}}$ in Eq. (12) and $\hat{H}_m^{\mathrm{M,bc}}$ in Eq. (21) obviously apply to the nonmixing case, a natural estimation procedure, extending univariate analysis, consists in performing linear regressions on the logarithms of the diagonal coefficients $\mathbf{S}_{m,m}(2^j)$ only:

$$\hat{H}_m^{\mathrm{U}} = \frac{1}{2}\left(\sum_{j=j_1}^{j_2} w_j \log_2 \mathbf{S}_{m,m}(2^j) - 1\right). \qquad (22)$$

The resulting estimators $\hat{H}_m^{\mathrm{U}}$, intrinsically *univariate* in spirit, were intensively studied in [19] for nonmixing mfBm.

B. Theoretical study of asymptotic estimation performance

1) Asymptotic framework: The estimation performance is studied theoretically in the asymptotic double and joint limits of large sample sizes $N \to +\infty$, with linear regressions performed at coarse scales, in the range $(j_1(N), j_2(N)) \to +\infty$. Technically, the scaling range is defined as $(j_1(N) = j_1^0 + \log_2 a(N), j_2(N) = j_2^0 + \log_2 a(N))$. The pair (j_1^0, j_2^0) is an arbitrarily chosen range for N_0, and $a(N)$ must satisfy:

$$a(N)2^{j_2} \leq N, \quad \frac{a(N)}{N} + \frac{N}{a(N)^{2\varpi+1}} \longrightarrow 0, \text{ as } N \to \infty, \qquad (23)$$

$$\varpi = \min\left\{\min_{\{1 \leq i \leq M \mid H_i - H_{i-1} > 0\}}(H_i - H_{i-1}), \frac{H_1}{2} + \frac{1}{4}\right\} \qquad (24)$$

(where, for notational simplicity, $H_0 := 0$). These conditions essentially imply that $\log_2 a(N)$ increases as $\beta \log_2 N$, with $1/(2\varpi + 1) < \beta < 1$. They also imply that the number of scales involved in linear regressions does not vary with sample size N: $j_2(N) - j_1(N) + 1 = j_2^0 - j_1^0 + 1$.

2) Asymptotic consistency: Theorem IV.1 establishes the asymptotic consistency of the bias-corrected estimator $\underline{\hat{H}}^{\mathrm{M,bc}}$:

Theorem IV.1. *Under Assumptions (OFBM1−3) and for a mother wavelet as defined in Section II-B1, (with $\overset{\mathbb{P}}{\to}$ the convergence in probability):*

(i) *For any $j = j_1^0, \ldots, j_2^0$, $\tau = 1, \ldots, 2^{j_2^0 - j}$ and $m = 1, \ldots, M$, as $N \to \infty$,*

$$\frac{\hat{\lambda}_m^{(\tau)}(a(N)2^j)}{a(N)^{2H_m+1}} \overset{\mathbb{P}}{\to} \xi_m(2^j) = (2^j)^{2H_m+1}\xi_m(2^0) > 0, \qquad (25)$$

where $\hat{\lambda}_1^{(\tau)}(a(N)2^j), \ldots, \hat{\lambda}_M^{(\tau)}(a(N)2^j)$ are the eigenvalues of $\mathbf{S}^{(\tau)}(a(N)2^j)$. Likewise, as $N \to \infty$,

$$\frac{\lambda_m^{(\tau)}(a(N)2^j)}{a(N)^{2H_m+1}} \to \xi_m(2^j) = (2^j)^{2H_m+1}\xi_m(2^0) > 0, \qquad (26)$$

where $\lambda_1^{(\tau)}(a(N)2^j), \ldots, \lambda_M^{(\tau)}(a(N)2^j)$ are the eigenvalues of $\mathbb{E}\mathbf{S}^{(\tau)}(a(N)2^j)$.

(ii) *Let $a(N)$ be as in (23). Then, as $N \to \infty$,*

$$\hat{H}_m^{\mathrm{M,bc}} \overset{\mathbb{P}}{\to} H_m, \quad \forall m = 1, \ldots, M, \qquad (27)$$

Proof. For (*i*), the proof follows from adapting the technique for establishing Theorem 3.1 in [30], see Section A of the supplementary material for details.

For (*ii*), Eq. (25) leads to the following equivalence:

$$\hat{H}_m^{\mathrm{M,bc}} = \frac{1}{2}\sum_{j=j_1^0}^{j_2^0} \frac{w_j}{2^{j_2^0-j}} \sum_{\tau=1}^{2^{j_2^0-j}}\left(\log_2 \frac{\hat{\lambda}_m^{(\tau)}(a(N)2^j)}{a(N)^{2H_m+1}}\right. \qquad (28)$$
$$\left. - \log_2 a(N)^{2H_m+1}\right) - \frac{1}{2}$$

$$\overset{\mathbb{P}}{\sim} \frac{1}{2}\sum_{j=j_1^0}^{j_2^0} \frac{w_j}{2^{j_2^0-j}} \sum_{\tau=1}^{2^{j_2^0-j}}\left(\log_2 (2^j)^{2H_m+1}\xi_m(2^0)\right.$$
$$\left. - \log_2 a(N)^{2H_m+1}\right) - \frac{1}{2}$$

$$\overset{\mathbb{P}}{\sim} \frac{1}{2}\sum_{j=j_1^0}^{j_2^0} w_j\left((2H_m+1)j + \log_2 \frac{\xi_m(2^0)}{a(N)^{2H_m+1}}\right) - \frac{1}{2}.$$

Hence, by the definition of w_j, as $N \to \infty$,

$$\hat{H}_m^{\mathrm{M,bc}} \overset{\mathbb{P}}{\sim} \frac{1}{2}[(2H_m + 1)] - \frac{1}{2} = H_m. \qquad (29)$$

This establishes (27). $\qquad \square$

3) Simple eigenvalues condition: Our theoretical study of the asymptotic distribution relies on the condition that, asymptotically, all wavelet spectra eigenvalues are simple.
CONDITION (C0): $\forall m_1, m_2 \in \{1, \ldots, M\}$, with $m_1 \neq m_2$,

$$H_{m_1} = H_{m_2} \Rightarrow \xi_{m_1}(2^0) \neq \xi_{m_2}(2^0), \qquad (30)$$

which states that either the selfsimilarity parameters H_m are different, or when they are equal, $H_{m_1} = H_{m_2}$, $m_1 \neq m_2$, the constants $\xi_m(2^0)$ in Eq. (11) are different.

4) Asymptotic normality: Theorem IV.2 establishes the asymptotic normality of the bias-corrected estimator $\underline{\hat{H}}^{\mathrm{M,bc}}$:

Theorem IV.2. *Let $n_{a,j} = \frac{N}{a(N)2^j}$, $j = j_1^0, \ldots, j_2^0$, be the number of wavelet coefficients available at scale $a(N)2^j$. Under Assumptions (OFBM1−3), for a mother wavelet as defined in Section II-B1, and assuming that (C0) holds, with $a(N)$ defined in (23) and where $\overset{d}{\to}$ denotes the convergence in distribution:*

(i) *Then, as $N \to \infty$,*

$$\left\{\sqrt{n_{a,j_2^0}}\left(\log_2 \hat{\lambda}_m^{(\tau)}(a(N)2^j)\right.\right. \qquad (31)$$
$$\left.\left. - \log_2 \lambda_m^{(\tau)}(a(N)2^j)\right)\right\}_{\substack{j=j_1^0,\ldots,j_2^0 \\ m=1,\ldots,M \\ \tau=1,\ldots,2^{(j_2^0-j_1^0)-j}}}$$
$$\overset{d}{\to} \mathcal{N}(0, \mathbf{\Upsilon_B}),$$

where $\mathbf{\Upsilon_B}$ is a symmetric positive semidefinite matrix.

(ii) *Moreover, as $N \to \infty$,*

$$\left\{\sqrt{N/a(N)}\left(\hat{H}_m^{\mathrm{M,bc}} - H_m\right)\right\}_{m=1,\ldots,M} \overset{d}{\to} \mathcal{N}(0, \mathbf{\Sigma_B}), \qquad (32)$$

for some $\mathbf{\Sigma_B} \in \mathcal{S}_{\geq 0}(M, \mathbb{R})$.

Proof. The proof follows from adapting the proof of asymptotic normality for the non bias-corrected multivariate estimator $\underline{\hat{H}}^{\mathrm{M}}$ from Theorem 3.2 in [30], as detailed in Section B of the supplementary material. $\qquad \square$

V. EMPIRICAL FINITE-SIZE ESTIMATION PERFORMANCE

The goal of the present section is to complement the theoretical study of estimation performance in the limit of large scales and large sample sizes. We do so by investigating their finite sample size counterparts by means of Monte Carlo simulations, conducted on synthetic mfBm. Different parametric configurations $(\boldsymbol{\Sigma}, \mathbf{W}, \underline{H})$ are used so as to further assess their impact on estimation performance.

The performance of the proposed bias-corrected multivariate estimator $\hat{\underline{H}}^{\mathrm{M,bc}}$ (Eq. (21)) is compared, in terms of biases, variances, Mean Squared Error (MSE), and covariance structures, to those of the multivariate estimator $\hat{\underline{H}}^{\mathrm{M}}$ (Eq. (12)) and of the univariate estimator $\hat{\underline{H}}^{\mathrm{U}}$ (Eq. (22)).

A. Numerical simulation set-up

The Monte Carlo simulations involve $N_{\mathrm{MC}} = 1000$ independent realizations of synthetic mfBm. The increments $\{B_{\boldsymbol{\Sigma}, \mathbf{W}, \underline{H}}(t+1) - B_{\boldsymbol{\Sigma}, \mathbf{W}, \underline{H}}(t)\}_{t=1,\dots,N}$ of a mfBm, referred to as multivariate fractional Gaussian noise, form a multivariate stationary Gaussian process with covariance directly stemming from Eq. (14). Thus, it can be synthesized numerically by circulant embedding, cf. e.g., [31].

Results reported here are restricted to $M = 6$ for the sake of clarity ; additional results for $M = 20$, available upon request, led to identical conclusions. The range of sample sizes covered is $N \in \{2^{13}, \dots, 2^{18}\}$. The parameters $\boldsymbol{\Sigma}$, $\mathbf{W}$ and $\underline{H}$ are chosen to investigate a set of representative configurations:
Config1 ($\boldsymbol{\Sigma} \neq \mathbb{I}$, $\mathbf{W} \neq \mathbb{I}$, H_m all different) Generic situation involving mixing, nonzero correlations amongst components, as well as different selfsimilarity parameters. This type of situation calls for multivariate estimators ($\hat{\underline{H}}^{\mathrm{M,bc}}$, $\hat{\underline{H}}^{\mathrm{M}}$) ;
Config2 ($\boldsymbol{\Sigma} \neq \mathbb{I}$, $\mathbf{W} \neq \mathbb{I}$, H_m all equal) Generic situation with mixing and correlations amongst components. Yet, all selfsimilarity parameters are assumed equal. Thus, univariate estimation $\hat{\underline{H}}^{\mathrm{U}}$ could be used in this multivariate setting, as much as the $\hat{\underline{H}}^{\mathrm{M,bc}}$, $\hat{\underline{H}}^{\mathrm{M}}$;
Config3 ($\boldsymbol{\Sigma} \neq \mathbb{I}$, $\mathbf{W} = \mathbb{I}$, H_m all different) Nonmixing situation, ence, univariate $\hat{\underline{H}}^{\mathrm{U}}$ and multivariate estimation ($\hat{\underline{H}}^{\mathrm{M,bc}}$, $\hat{\underline{H}}^{\mathrm{M}}$) are all legitimate procedures ;
Config4 ($\boldsymbol{\Sigma} \neq \mathbb{I}$, $\mathbf{W} = \mathbb{I}$, H_m all equal) Nonmixing situations, yet, all selfsimilarity parameters are equal, and thus both univariate $\hat{\underline{H}}^{\mathrm{U}}$ and multivariate estimations ($\hat{\underline{H}}^{\mathrm{M,bc}}$, $\hat{\underline{H}}^{\mathrm{M}}$) should be appropriate procedures.
Config5 ($\boldsymbol{\Sigma} = \mathbb{I}$, $\mathbf{W} \neq \mathbb{I}$, H_m all different) Mixing situations with different selfsimilarity parameters, yet no correlation.

A generic covariance matrix $\boldsymbol{\Sigma}$ is chosen: $\boldsymbol{\Sigma}_{m,m'} = 0.7^{|m-m'|}$, $1 \leq m \leq m' \leq M$. A mixing matrix $\mathbf{W}$ is randomly drawn from the set of $M \times M$ real-valued nonorthogonal matrices and kept fixed for all experiments. The selfsimilarity parameter vector $\underline{H} = (H_1, \dots, H_6)$ is either set to $H_m = 0.7$, $\forall m$, or $\underline{H} = (0.4, 0.5, 0.6, 0.65, 0.7, 0.8)$. These different configurations are chosen so that CONDITION (C0) holds – this condition is required for asymptotic normality claims (cf. Section IV-B3). Wavelet analysis is performed using the least asymmetric Daubechies2 mother wavelet [28].
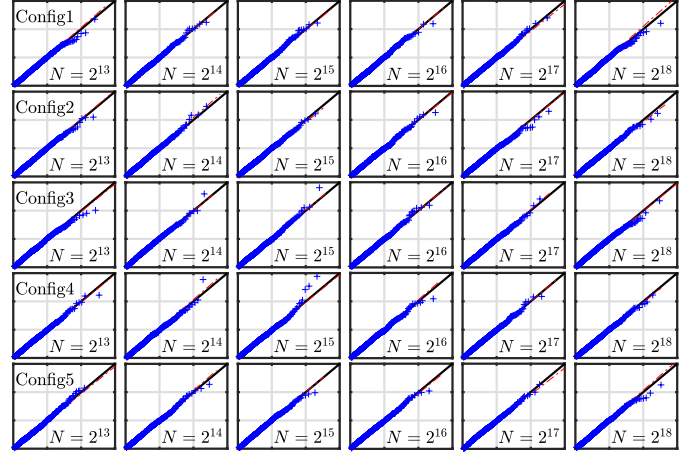

Fig. 3: **Finite-size joint normality for $\hat{\underline{H}}^{\mathrm{M,bc}}$.** Quantile-quantile plots of the empirical distribution of the Mahalanobis distance computed from the bias-corrected multivariate estimates $\hat{\underline{H}}^{\mathrm{M,bc}}$ against a χ^2 distribution with M degrees of freedom (expected theoretically under joint normality), for different sample sizes N and different configurations.

Identical conclusions can be obtained using other wavelets satisfying $N_\psi \geq 2$. Linear regressions are performed, with w_j as defined in [19], over scales that depend on N, as analyzed theoretically in Section IV-B1 (Eq. (23)): $2^{j_1} = a(N)2^6$ to $2^{j_2} = a(N)2^9$, where $a(N) = 2^{\lfloor \beta \log_2(N/N_0) \rfloor}$ with $\beta = 0.9$ and $N_0 = 2^{13}$.

B. Asymptotic normality assessment

The normality of $\hat{\underline{H}}^{\mathrm{M,bc}}$ is assessed by computing quantile-quantile plots for the squared Mahalanobis distance,

$$(\hat{\underline{H}}^{\mathrm{M,bc}} - \mathbb{E}\hat{\underline{H}}^{\mathrm{M,bc}})\mathrm{Var}(\hat{\underline{H}}^{\mathrm{M,bc}})^{-1}(\hat{\underline{H}}^{\mathrm{M,bc}} - \mathbb{E}\hat{\underline{H}}^{\mathrm{M,bc}})^\top,$$

where the ensemble average $\mathbb{E}$ is computed as the mean across independent realizations. Samples of this Mahalanobis distance are plotted against a χ^2 distribution with M degrees of freedom, since this distribution is expected to hold under the exact joint normality of $\hat{\underline{H}}^{\mathrm{M,bc}}$. Fig. 3 reports such quantile-quantile plots, for the different configurations and for different sample sizes. It shows very satisfactory straight lines, even at the smallest sample size $N = 2^{13}$. This suggests that the joint normality for $\hat{\underline{H}}^{\mathrm{M,bc}}$ holds very satisfactorily for finite sample size estimation, even for small sample sizes. It thus indicates fast convergence to asymptotic normality, proven theoretically in Theorem IV.2 (Eq. (32)).

C. Estimation performance: bias, covariance, MSE

For any estimate of the vector of selfsimilarity parameters $\hat{\underline{H}}$, matrices of bias, covariance and MSE are defined as

$$\mathrm{Bias}^2(\hat{\underline{H}}) = (\mathbb{E}\hat{\underline{H}} - \underline{H})(\mathbb{E}\hat{\underline{H}} - \underline{H})^\top, \tag{33}$$

$$\mathrm{Cov}(\hat{\underline{H}}) = \mathbb{E}\left[(\mathbb{E}\hat{\underline{H}} - \hat{\underline{H}})(\mathbb{E}\hat{\underline{H}} - \hat{\underline{H}})^\top\right], \tag{34}$$

$$\mathrm{MSE}(\hat{\underline{H}}) = \mathbb{E}\left[(\hat{\underline{H}} - \underline{H})(\hat{\underline{H}} - \underline{H})^\top\right]. \tag{35}$$
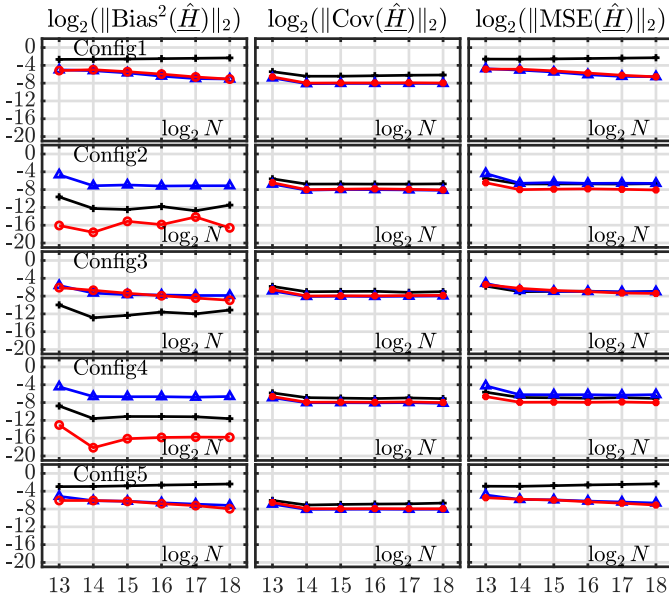
Fig. 4: **Estimation performance.** Logarithm of the spectral norms of the bias (left), covariance (middle) and MSE (right) for the proposed bias-corrected multivariate estimator $\underline{\hat{H}}^{\mathrm{M,bc}}$ (red), compared to those for the classical multivariate estimator $\underline{\hat{H}}^{\mathrm{M}}$ (blue) and for the univariate estimator $\underline{\hat{H}}^{\mathrm{U}}$ (black), as functions of the logarithm of the sample size N.

The estimation performance is quantified by the spectral norm of these matrices, i.e., by the largest absolute value of their eigenvalues [32].

Fig. 4 reports the (logarithms of the) spectral norms of the bias, covariance and MSE for the proposed bias-corrected multivariate estimator $\underline{\hat{H}}^{\mathrm{M,bc}}$, compared to those for the classical multivariate estimator $\underline{\hat{H}}^{\mathrm{M}}$ and for the univariate estimator $\underline{\hat{H}}^{\mathrm{U}}$. The results are shown as functions of the logarithm of the sample size N and for each of the different configurations.

For the generic case represented by **Config1** (Fig. 4, first row), the univariate estimator $\underline{\hat{H}}^{\mathrm{U}}$ suffers from significant bias stemming from the component mixing caused by $\mathbf{W}$. Thus, it is significantly outperformed by the multivariate estimators $\underline{\hat{H}}^{\mathrm{M}}$ and $\underline{\hat{H}}^{\mathrm{M,bc}}$. The latter show equivalent performance as the combination of $\mathbf{\Sigma}$, $\mathbf{W}$ and different H_m yield different eigenvalues for $\mathbb{E}\mathbf{S}(2^j)$ at all scales and thus small *repulsion bias* in their estimated values.

For **Config2** (Fig. 4, second row), with mixing yet equal H_m, the eigenvalues of $\mathbb{E}\mathbf{S}(2^j)$ lie much closer to each other at most scales and hence suffer from substantial repulsion biases in their estimation. This results in a significant bias for the classical multivariate estimator $\underline{\hat{H}}^{\mathrm{M}}$. Fig. 4, middle row, shows that the proposed bias-corrected multivariate estimator $\underline{\hat{H}}^{\mathrm{M,bc}}$ succeeds in significantly reducing bias without an increase in the variance. This leads to an overall much improved estimation performance as quantified by the MSE. While the univariate estimator $\underline{\hat{H}}^{\mathrm{U}}$ displays small bias, as expected in view of the equality of all selfsimilarity parameters, it also shows greater variance, thus, greater MSE, when compared to the proposed $\underline{\hat{H}}^{\mathrm{M,bc}}$.

For the nonmixing case, **Config3** (Fig. 4, third row), the univariate estimator $\underline{\hat{H}}^{\mathrm{U}}$ naturally shows much smaller biases compared to the multivariate estimators $\underline{\hat{H}}^{\mathrm{M}}$ and $\underline{\hat{H}}^{\mathrm{M,bc}}$. Yet, it also displays larger variances. Thus, as a result of bias-variance trade-off, the MSE of all three estimators are equivalent. Importantly, this shows that even for data corresponding to nonmixing situations, there is no cost in estimation performance associated with the use of multivariate estimators.

For the nonmixing case with equal H_m, **Config4** (Fig. 4, fourth row), the univariate estimator $\underline{\hat{H}}^{\mathrm{U}}$ and bias-corrected multivariate estimator $\underline{\hat{H}}^{\mathrm{M,bc}}$ perform similarly to **Config2**. This suggests that, for H_m all equal, the presence of mixing has no impact in terms of bias or variance. However, estimation performance of the multivariate estimator $\underline{\hat{H}}^{\mathrm{M}}$ is slightly more affected than in **Config2**, showing that the repulsion bias is more substantial in the absence of mixing.

Finally, for the mixing case with different H_m and no correlation, **Config5** (Fig. 4, fifth row), the performances are similar to those for **Config1** for all three estimators, indicating their robustness against component correlations $\mathbf{\Sigma}$.

All together, the results reported in Fig. 4 lead us to conclude that the proposed bias-corrected multivariate estimator $\underline{\hat{H}}^{\mathrm{M,bc}}$ displays the best estimation performance regardless of the actual configuration of the parameters. In practice it is not a priori known whether or not mixing is present, or whether selfsimilarity parameters are pairwise distinct or equal. We have seen that the estimator $\underline{\hat{H}}^{\mathrm{U}}$ fails when mixing is present. Also, we have found out that the estimator $\underline{\hat{H}}^{\mathrm{M}}$ fails when repulsion bias is present, i.e., when one needs to estimate identical selfsimilarity parameters. By contrast, $\underline{\hat{H}}^{\mathrm{M,bc}}$ displays neither one of these limitations – this versatility is clearly an important property that motivates its practical use.

D. Covariance structure of $\underline{\hat{H}}^{\mathrm{M,bc}}$

1) Correlation structure: We complement the estimation performance assessment by turning to the dependence structure amongst estimates – the latter constitutes a key feature in the practical use of the estimators of the selfsimilarity parameters. Asymptotic normality, as well as finite sample size approximate normality, hint on the importance of the correlation structure amongst estimated selfsimilarity parameters. Fig. 5 thus compares correlation coefficients between pairs $(\hat{H}_m, \hat{H}_{m'})$, $1 \leq m < m' \leq M$, as functions of the sample size N, for the three different estimators $\underline{\hat{H}}^{\mathrm{M,bc}}$, $\underline{\hat{H}}^{\mathrm{M}}$ and $\underline{\hat{H}}^{\mathrm{U}}$. The correlation structures are compared only for **Config1** and **Config3**, in upper and lower triangles respectively. The conclusions drawn about the other configurations are similar.

Fig. 5 clearly indicates that most pairs $(\hat{H}_m^{\mathrm{U}}, \hat{H}_{m'}^{\mathrm{U}})$, $1 \leq m < m' \leq M$, remain significantly correlated, even for large sample sizes. By contrast, for the multivariate estimators, the pairs $(\hat{H}_m^{\mathrm{M,bc}}, \hat{H}_{m'}^{\mathrm{M,bc}})$ and $(\hat{H}_m^{\mathrm{M}}, \hat{H}_{m'}^{\mathrm{M}})$, $1 \leq m < m' \leq M$ are quasi-decorrelated, for all sample sizes. These observations partly explain why the overall contribution of the variance to the MSE was larger for the univariate estimators when
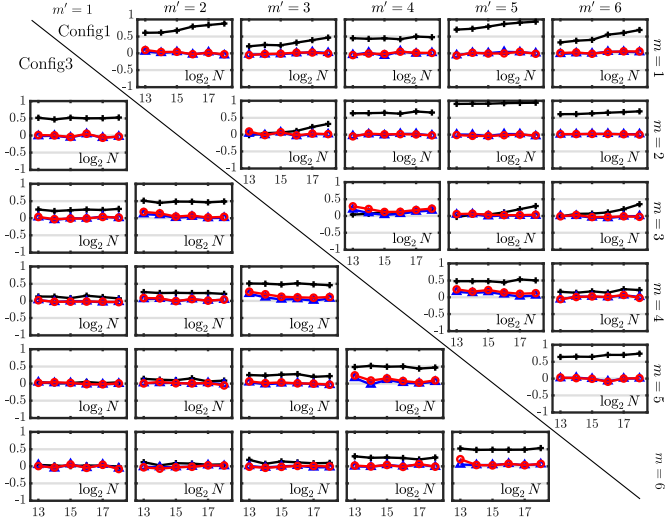
Fig. 5: **Correlation structures of $\underline{\hat{H}}^{\mathrm{M,bc}}$, $\underline{\hat{H}}^{\mathrm{M}}$ and $\underline{\hat{H}}^{\mathrm{U}}$.** Correlation coefficients between entries $\hat{H}_m$ and $\hat{H}_{m'}$, $1 \leq m \neq m' \leq M$, for $\underline{\hat{H}}^{\mathrm{M,bc}}$ (red), $\underline{\hat{H}}^{\mathrm{M}}$ (blue) and $\underline{\hat{H}}^{\mathrm{U}}$ (black), as functions of sample sizes N, for **Config1** (upper right triangle) and **Config3** (lower left triangle) .

compared to the multivariate estimators (cf. Fig. 4 above). These empirical observations are complemented and framed by the analytical calculations detailed in Sections C and D of the supplementary material. These show the following striking fact. Suppose CONDITION (C0) holds, and further assume that the weak correlation structure amongst wavelet coefficients of mfBm can be approximated into exact decorrelation. Then, the covariance matrices $\boldsymbol{\Sigma}_{\mathbf{B}}$ in Eq. (31) and $\boldsymbol{\Upsilon}_{\mathbf{B}}$ in Eq. (32) of Theorem IV.2 are well approximated, to the first order, by diagonal matrices. These empirical observations and analytical calculations reveal an important advantage of the mutivariate $\underline{\hat{H}}^{\mathrm{M,bc}}$ (and $\underline{\hat{H}}^{\mathrm{M}}$) against the univariate $\underline{\hat{H}}^{\mathrm{U}}$, even for non-mixing data: namely, that the former consist of asymptotically Gaussian, weakly dependent random vectors. This is a major practical feature, e.g., in the design of tests for the equality of selfsimilarity parameters, which is of significant interest in many applications [27], [33].

2) Variance: Besides approximate asymptotic decorrelation, the diagonal entries of the matrices $\boldsymbol{\Sigma}_{\mathbf{B}}$ and $\boldsymbol{\Upsilon}_{\mathbf{B}}$ remain to be discussed. The calculations reported in Remark A.2 in Section B, and in Sections C and D of the supplementary material show that these diagonal entries can be first-order approximated by a constant:

$$\boldsymbol{\Upsilon}_{\mathbf{B}} \simeq 2 \left(\log_2 e\right)^2 \mathbb{I} \quad \text{and} \quad \boldsymbol{\Sigma}_{\mathbf{B}} \simeq \frac{(\log_2 e)^2}{2} \sum_{j=j_1^0}^{j_2^0} w_j^2 2^j \, \mathbb{I}. \quad (36)$$

Specifically, this implies that, under the approximation of decorrelation between wavelet coefficients of mfBm and under (C0), the finite-size variances of $\hat{H}_m^{\mathrm{M,bc}}$ and $\hat{H}_m^{\mathrm{M}}$, for every $m \in \{1, \ldots, M\}$, can be approximated as:

$$\mathrm{Var}(\hat{H}_m^{\mathrm{M,bc}}) \approx \mathrm{Var}(\hat{H}_m^{\mathrm{M}}) \approx \frac{(\log_2 e)^2}{2} \sum_{j=j_1}^{j_2} \frac{w_j^2}{n_j} := V_N. \quad (37)$$
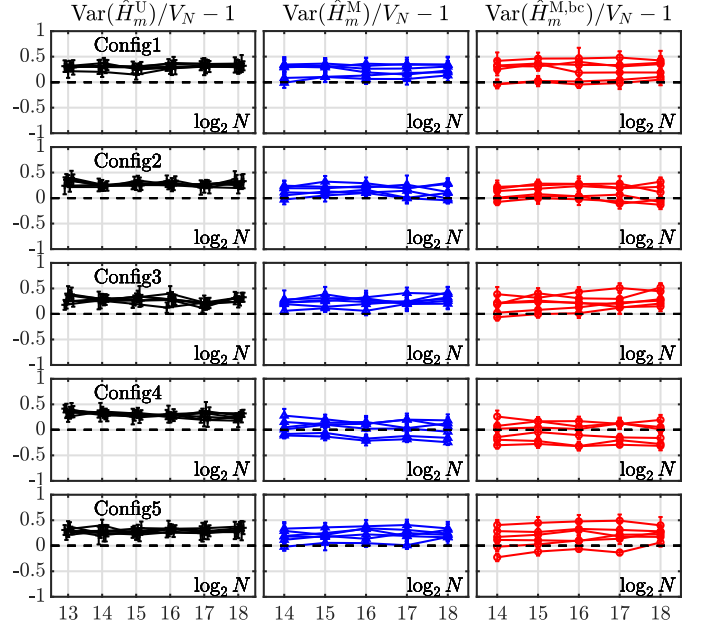

Fig. 6: **Approximation of the variance.** Relative differences between empirical variances of the (left, black) univariate estimator $\hat{H}_m^{\mathrm{U}}$, (middle, blue) multivariate estimator $\hat{H}_m^{\mathrm{M}}$ and (right, red) bias-corrected multivariate estimator $\hat{H}_m^{\mathrm{M,bc}}$ and the approximate variance V_N (Eq. (37)) with superimposed components $m = 1, \ldots, 6$ as functions of the sample size N.

When $N \to +\infty$, one obviously has $(N/a(N))V_N \to \frac{(\log_2 e)^2}{2} \sum_{j=j_1^0}^{j_2^0} w_j^2 2^j$. Thus, in conclusion, $\mathrm{Var}(\hat{H}_m^{\mathrm{M}})$ and $\mathrm{Var}(\hat{H}_m^{\mathrm{M,bc}})$ decrease with sample size N and depend on the range of regressions scales j_1 and j_2. However, to the first order, these depend on none of the model parameters $\boldsymbol{\Sigma}$, $\mathbf{W}$ and $\underline{H}$. This renders them comparable for all components. Interestingly, it was shown in [19] that, for the univariate $\hat{H}_m^{\mathrm{U}}$ in nonmixing settings, the variance can be approximated to the first order by the same closed-form relation: $\mathrm{Var}(\hat{H}_m^{\mathrm{U}}) \approx V_N$.

To further assess the practical relevance of the finite sample size approximation in Eq. (37), Fig. 6 reports the relative difference between the empirical variances of $\hat{H}_m^{\mathrm{U}}$, $\hat{H}_m^{\mathrm{M}}$ and $\hat{H}_m^{\mathrm{M,bc}}$ and V_N as a function of the logarithm of the sample size N for $m = 1, \ldots, 6$ and for the different configurations. Fig. 6 (left column), first, shows that, in all configurations, the variance of the univariate estimator $\hat{H}_m^{\mathrm{U}}$ is approximated by V_N with a 20% relative error resulting from the non-exact decorrelation of eigenvalues between scales. This first result extends the results reported in [19], limited to nonmixing cases. Fig. 6, second, confirms that the variances of the three estimators neither depend on $\underline{H}$ (**Config1** vs. **Config2**), nor on $\mathbf{W}$ (**Config1** vs. **Config3**), nor on $\boldsymbol{\Sigma}$ (**Config1** vs. **Config5**). Fig. 6, third, confirms that the variances of the multivariate estimators are, on average, lower than those of the univariate estimator. In summary, Fig. 6 shows that the closed-form approximation V_N in Eq. (37) is valid in all configurations for the three estimators, even for small sample sizes. This constitutes a valuable result for practical use in applications, e.g., when testing the equality of selfsimilarity parameters.

VI. APPLICATION TO EPILEPTIC SEIZURE PREDICTION

Selfsimilarity has been broadly used to model temporal dynamics in macroscopic brain activity. The purpose of the modeling effort is either to study healthy brain mechanisms (cf. e.g., [9], [10], [34]), or to diagnose brain disorders and pathologies. Notably, there is a particular interest in epileptic seizure prediction [35]. Yet, most studies have so far remained based on univariate analysis of scale-free dynamics, while brain dynamics is naturally conducive to multivariate analysis. The aim of this section is, thus, to comparatively assess the potential of multivariate and univariate selfsimilarity parameter estimations in the detection of ictal, preictal and interictal states against each other (i.e., time periods during seizure, prior to seizure onsets or distant from any seizure, respectively).

The multi-channel EEG recordings studied here were collected at the Boston Children's Hospital from pediatric patients with medically intractable seizures, from the CHB-MIT Scalp EEG database [36], available at https://physionet.org/content/chbmit/1.0.0/. EEG recordings stem from the International 10-20 system of EEG electrode positions and nomenclature, last for at least one hour, and are sampled at 256Hz. Multivariate analyses are performed for the 19 channels, within sliding 16 second-long wondows. Wavelet analysis is performed using the least asymmetric Daubechies2 mother wavelet.

Fig. 7 reports, for ictal, preictal and interictal windows, the logarithms of the diagonal entries of the wavelet spectrum $\mathbf{S}(2^j)$, $\log_2 \mathbf{S}_{m,m}(2^j)$, the log-eigenvalues $\log_2 \hat{\lambda}_m(2^j)$ and the proposed bias-corrected log-eigenvalues $\overline{\log_2 \lambda_m}(2^j)$, all of them as functions of scales 2^j for $m = 1, \ldots, 19$. Fig. 7 shows linear behaviors for these functions across fine scales $1 \leq 2^j \leq 2^7$, that hence indicate scale-free dynamics in these EEG time series, at frequencies above 1Hz, in agreement with the literature on epilepsy, cf. e.g., [37]. These scale-free dynamics hold for the three different states and both component-wise (univariate, $\log_2 \mathbf{S}_{m,m}(2^j)$) and jointly (multivariate, $\log_2 \hat{\lambda}_m(2^j)$ and $\overline{\log_2 \lambda_m}(2^j)$). The observed multivariate selfsimilarity constitutes per se a significant outcome of this work as selfsimilarity in EEG time series had so far only been observed component per component and not jointly. Combined with literature suggesting that differences between preictal and interictal dynamics are related to frequencies ranging from 10Hz to 85Hz [37], these observations suggest performing linear regressions for the estimation of $\underline{H}$ across scales ranging from $2^1 \leq 2^j \leq 2^4$.

Fig. 8 compares the (logarithms of) sorted Wilcoxon rank-sum test p-values $p_{\tau(m)}$ measuring the significance of the differences between the medians of ictal vs. preictal, ictal vs. interictal and preictal vs. interictal distributions, for the univariate estimator $\hat{\underline{H}}^{\mathrm{U}}$ (Eq. (22)), the multivariate estimator $\hat{\underline{H}}^{\mathrm{M}}$ (Eq. (12)) and the bias-corrected multivariate estimator $\hat{\underline{H}}^{\mathrm{M,bc}}$ (Eq. (21)). The p-values are ranked and compared to Benjamini-Hochberg multiple hypothesis thresholds [38], computed at a false discovery rate $\alpha = 0.05$. For all three estimators, the medians of the distributions of estimated selfsimilarity parameters clearly differ from ictal/preictal to interictal states but not from preictal to ictal states. This indicates that the scale-free structure of the temporal dynamics
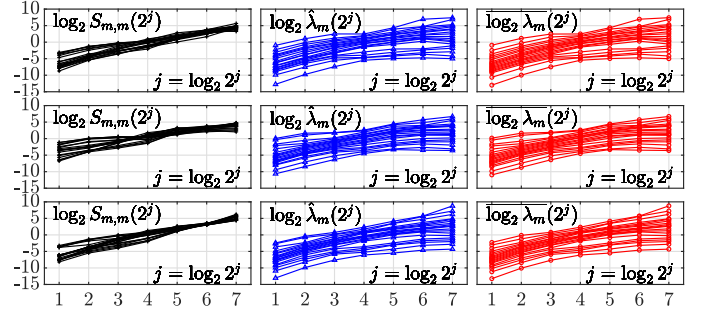

Fig. 7: **Multivariate scale-free dynamics in EEG time series, for ictal, preictal and interictal states.** Logarithms of the estimated diagonal entries $\log_2 \mathbf{S}_{m,m}(2^j)$ (left, black), estimated eigenvalues $\log_2 \hat{\lambda}_m(2^j)$ (middle, blue) and bias corrected eigenvalues $\overline{\log_2 \lambda_m}(2^j)$ (right, red) of the wavelet spectrum, for $m = 1, \ldots, 19$, for ictal (top), preictal (middle), interictal (bottom) windows.
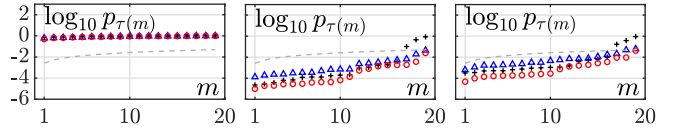

Fig. 8: **Differences in multivariate selfsimilarities.** Logarithms of the sorted p-values for Wilcoxon rank-sum tests between ictal / preictal distributions (left), ictal / interictal distributions (middle) and preictal / interictal distributions (right), for univariate estimates $\hat{H}_m^{\mathrm{U}}$ (black), multivariate estimates $\hat{H}_m^{\mathrm{M}}$ (blue) and bias corrected multivariate estimates $\hat{H}_m^{\mathrm{M,bc}}$ (red), vs. Benjamini-Hochberg multiple hypothesis testing thresholds (grey dashed line).

is modified by forthcoming epileptic seizures but not within periods evolving to a seizure. Moreover, the bias-corrected multivariate estimator $\hat{\underline{H}}^{\mathrm{M,bc}}$ shows stronger differences in selfsimilarity parameters between ictal or preictal states and interictal states, compared to either the univariate estimates $\hat{\underline{H}}^{\mathrm{U}}$ or the multivariate estimates $\hat{\underline{H}}^{\mathrm{M}}$. This illustrates the benefits of both taking into account cross-temporal dynamics in EEG time series analysis and correcting the repulsion bias in multivariate estimation for epilepsy seizure prediction.

These analyses are reported for a single patient only (Patient 22). Equivalent conclusions were drawn for all other patients. Of importance, the distributions of selfsimilarity parameters are found to differ significantly from one patient to another. This demands a single-patient procedure to detect ictal / preictal from interictal states. Interested readers are referred to the in-depth study carried out in [39].

VII. CONCLUSIONS AND PERSPECTIVES

In this work, we have, first, proposed a multivariate self-similarity model, defined as a linear mixture of correlated fractional Brownian motions, well-suited to adjust the versatility of real-world multivariate data, and formally related it to a former mathematical model. Second, we have turned the formal proposition of wavelet spectrum eigenvalue-based estimation for the vector of selfsimilarity parameters proposed

in [25], [26], into a versatile and reliable bias-corrected estimation procedure that can actually be used on real-world data. Its performance was studied mathematically in terms of asymptotics and by numerical simulations for finite-size samples. This permitted showing that the proposed multivariate estimators work in a broader context (mixing) than the univariate one and still performs better (decorrelation) in the nonmixing case, where the univariate estimator can be used. Third, its potential was illustrated at work on real multi-channel EEG data in a context of epileptic seizure prediction. A documented toolbox implementing estimation is publicly available at github.com/charlesglucas/ofbm_tools.

Future investigations include incorporating a multivariate time-scale bootstrap procedure so as to assess confidence levels in estimation from a single realization of data. Also, large-dimensional asymptotics will be studied both theoretically and numerically to account for situations where the number of components grows as fast as sample size, of major practical importance in real-world applications.

References

[1] U. Frisch, *Turbulence: the Legacy of A.N. Kolmogorov*. Cambridge University Press, 1995.
[2] R. L. Davis, H. M. Hodges, G. F. Smoot, P. J. Steinhardt, and M. S. Turner, "Cosmic microwave background probes models of inflation," *Physical Review Letters*, vol. 69, no. 13, p. 1856, 1992.
[3] P. Hubert, H. Bendjoudi, D. Schertzer, and S. Lovejoy, "Multifractal taming of extreme hydrometeorological events," *The Extremes of the Extremes (Proceedings of Reykjavik Symposium, July 2000)*, no. 271, pp. 51–56, 2002.
[4] D. Turcotte, "Implications of chaos, scale-invariance, and fractal statistics in geology," *Global and Planetary Change*, vol. 3, no. 3, pp. 301–308, 1990.
[5] M. W. Palmer, "Fractal geometry: a tool for describing spatial patterns of plant communities," *Vegetatio*, vol. 75, pp. 91–102, may 1988.
[6] R. Lopes and N. Betrouni, "Fractal and multifractal analysis: A review," *Medical image analysis*, vol. 13, no. 4, pp. 634–649, 2009.
[7] P. C. Ivanov, "Scale-invariant aspects of cardiac dynamics across sleep stages and circadian phases," *IEEE Engineering in Medicine and Biology Magazine*, vol. 26, no. 6, pp. 33–37, 2007.
[8] T. Nakamura, K. Kiyono, H. Wendt, P. Abry, and Y. Yamamoto, "Multiscale analysis of intensive longitudinal biomedical signals and its clinical applications," *Proceedings of the IEEE*, vol. 104, no. 2, pp. 242–261, 2016.
[9] B. J. He, "Scale-free brain activity: past, present, and future," *Trends in cognitive sciences*, vol. 18, no. 9, pp. 480–487, 2014.
[10] D. La Rocca, N. Zilber, P. Abry, V. van Wassenhove, and P. Ciuciu, "Self-similarity and multifractality in human brain activity: a wavelet-based analysis of scale-free brain dynamics," *Journal of Neuroscience Methods*, vol. 309, pp. 175–187, November 2018.
[11] M. Takahashi, "A fractal model of chromosomes and chromosomal dna replication," *Journal of Theoretical Biology*, vol. 141, no. 1, pp. 117–136, 1989.
[12] R. Fontugne, P. Abry, K. Fukuda, D. Veitch, K. Cho, P. Borgnat, and H. Wendt, "Scaling in internet traffic: a 14 year and 3 day longitudinal study, with multiscale analyses and random projections," *IEEE/ACM Transactions on Networking*, vol. 25, no. 4, pp. 2152–2165, 2017.
[13] B. B. Mandelbrot, "A multifractal walk down wall street," *Scientific American*, vol. 280, no. 2, pp. 70–73, 1999.
[14] J. Lengyel, S. Roux, F. Sémécurbe, S. Jaffard, and P. Abry, "Roughness and intermittency within metropolitan regions-Application in three French conurbations," *Environment and Planning B: Urban Analytics and City Science*, vol. 50, no. 3, pp. 600–620, 2023.
[15] P. Abry, S. G. Roux, H. Wendt, P. Messier, A. G. Klein, N. Tremblay, P. Borgnat, S. Jaffard, B. Vedel, J. Coddington *et al.*, "Multiscale Anisotropic Texture Analysis and Classification of Photographic Prints: Art scholarship meets image processing algorithms," *IEEE Signal Processing Magazine*, vol. 32, no. 4, pp. 18–27, 2015.

[16] G. Samorodnitsky and M. Taqqu, *Stable non-Gaussian Random Processes: Stochastic Models with Infinite Variance*. New York: Chapman and Hall, 1994.
[17] V. Pipiras and M. S. Taqqu, *Long-Range Dependence and Self-Similarity*. Cambridge university press, 2017, vol. 45.
[18] J.-M. Bardet, G. Lang, G. Oppenheim, A. Philippe, S. Stoev, and M. Taqqu, "Semi-parametric estimation of the long-range dependence parameter: A survey," in *Theory and Applications of Long-Range Dependence*, P. Doukhan, G. Oppenheim, and M. S. Taqqu, Eds. Boston: Birkhäuser, 2003, pp. 557–577.
[19] H. Wendt, G. Didier, S. Combrexelle, and P. Abry, "Multivariate Hadamard self-similarity: testing fractal connectivity," *Physica D: Nonlinear Phenomena*, vol. 356, pp. 1–36, October 2017.
[20] P.-O. Amblard and J.-F. Coeurjolly, "Identification of the multivariate fractional Brownian motion," *IEEE Transactions on Signal Processing*, vol. 59, no. 11, pp. 5152–5168, 2011.
[21] J.-F. Coeurjolly, P.-O. Amblard, and S. Achard, "Wavelet analysis of the multivariate fractional Brownian motion," *ESAIM: Probability and Statistics*, vol. 17, pp. 592–604, Aug. 2013.
[22] J. D. Mason and Y. Xiao, "Sample path properties of operator-self-similar Gaussian random fields," *Theory of Probability & Its Applications*, vol. 46, no. 1, pp. 58–78, 2002.
[23] G. Didier and V. Pipiras, "Integral representations and properties of operator fractional Brownian motions," *Bernoulli*, vol. 17, no. 1, pp. 1–33, 2011.
[24] ——, "Exponents, symmetry groups and classification of operator fractional Brownian motions," *Journal of Theoretical Probability*, vol. 25, no. 2, pp. 353–395, 2012.
[25] P. Abry and G. Didier, "Wavelet estimation for operator fractional Brownian motion," *Bernoulli*, vol. 24, no. 2, pp. 895–928, 2018.
[26] ——, "Wavelet eigenvalue regression for n-variate operator fractional Brownian motion," *Journal of Multivariate Analysis*, vol. 168, pp. 75–104, November 2018.
[27] C.-G. Lucas, P. Abry, H. Wendt, and G. Didier, "Bootstrap for testing the equality of selfsimilarity exponents across multivariate time series," in *European Signal Processing Conference (EUSIPCO)*, 2021, pp. 1960–1964.
[28] S. Mallat, *A Wavelet Tour of Signal Processing*. Elsevier, 1999.
[29] T. Tao, *Topics in Random Matrix Theory*. American Mathematical Society, 2012, vol. 132.
[30] P. Abry, B. C. Boniece, G. Didier, and H. Wendt, "On high-dimensional wavelet eigenanalysis," *arXiv preprint arXiv:2102.05761*, 2021.
[31] H. Helgason, V. Pipiras, and P. Abry, "Fast and exact synthesis of stationary multivariate gaussian time series using circulant embedding," *Signal Processing*, vol. 91, no. 5, pp. 1123–1133, 2011.
[32] C. D. Meyer and I. Stewart, *Matrix Analysis and Applied Linear Algebra*. SIAM, 2023.
[33] C.-G. Lucas, P. Abry, H. Wendt, and G. Didier, "Counting the number of different scaling exponents in multivariate scale-free dynamics: Clustering by bootstrap in the wavelet domain," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2022, pp. 5513–5517.
[34] P. Ciuciu, P. Abry, and B. J. He, "Interplay between functional connectivity and scale-free dynamics in intrinsic fMRI networks," *Neuroimage*, vol. 95, pp. 248–263, July 2014.
[35] O. D. Domingues, P. Ciuciu, D. La Rocca, P. Abry, and H. Wendt, "Multifractal analysis for cumulant-based epileptic seizure detection in eeg time series," in *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*, 2019, pp. 143–146.
[36] A. L. Goldberger, L. A. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng, and H. E. Stanley, "PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals," *Circulation*, vol. 101, no. 23, pp. e215–e220, 2000.
[37] K. Gadhoumi, J. Gotman, and J.-M. Lina, "Scale invariance properties of intracerebral EEG improve seizure prediction in mesial temporal lobe epilepsy," *PLoS one*, vol. 10, no. 4, p. e0121182, 2015.
[38] Y. Benjamini and Y. Hochberg, "Controlling the false discovery rate: a practical and powerful approach to multiple testing," *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 57, no. 1, pp. 289–300, 1995.
[39] C.-G. Lucas, P. Abry, H. Wendt, and G. Didier, "Epileptic seizure prediction from eigen-wavelet multivariate selfsimilarity analysis of multi-channel EEG signals," in *European Signal Processing Conference (EUSIPCO)*, 2023, pp. 970–974.
[40] J. R. Magnus, "On differentiating eigenvalues and eigenvectors," *Econometric Theory*, vol. 1, no. 2, pp. 179–191, 1985.

[41] P. Abry, B. C. Boniece, G. Didier, and H. Wendt, "Wavelet eigenvalue regression in high dimensions," *Statistical Inference for Stochastic Processes*, vol. 26, no. 1, pp. 1–32, 2023.

Supplementary Material

Fix $\ell \in \mathbb{N}$. Throughout the appendices, $\mathcal{S}(\ell, \mathbb{R})$ and $\mathcal{S}_{\geq 0}(\ell, \mathbb{R})$ denote, respectively, the space of symmetric and the cone of symmetric positive semidefinite $\ell \times \ell$ matrices. The notation $\text{vec}_{\mathcal{S}}$ denotes the projection on the free entries of a symmetric matrix. Also, π_{i_1, i_2} and $\pi_{i_1, i_2, \ldots, i_\ell}$, $\ell \in \mathbb{N}$, denote the projection of a matrix or vector, respectively, on its entries i_1, i_2 and $i_1, i_2, \ldots, i_\ell$. We express any vector $\mathbf{v}$ entry-wise as $\{\mathbf{v}_\ell\}_\ell$. Recall that $n_{a,j} = \frac{N}{a(N) 2^j}$, $j = j_1^0, \ldots, j_2^0$.

A. Proof of consistency

Proof. We only show (i). Fix $j \in \{j_1^0, \ldots, j_2^0\}$, $\tau \in \{1, \ldots, 2^{j_2^0 - j}\}$ and $m \in \{1, \ldots, M\}$. Suppose, for the moment, that the limit (25) holds. Recall that $p\text{-}\lim_{N \to \infty}$ denotes limit in probability. Writing $\widetilde{a}(N) = a(N) 2^j$,

$$
\begin{aligned}
\xi_m(2^j) &= p\text{-}\lim_{N \to \infty} \frac{\hat{\lambda}_m^{(\tau)}(a(N) 2^j)}{a(N)^{2H_m + 1}} \\
&= p\text{-}\lim_{N \to \infty} \frac{\hat{\lambda}_m^{(\tau)}(\widetilde{a}(N))}{(\widetilde{a}(N) 2^{-j})^{2H_m + 1}} \\
&= 2^{j(2H_m + 1)} \xi_m(1).
\end{aligned} \tag{38}
$$

This establishes the scaling relationship in (25).

We now need to show the convergence in (25). The argument follows from a simple adaptation of the technique for establishing Theorem 3.1 in [30], which originally pertains to a high-dimensional context.

In fact, in that theorem the measurements are given by a high-dimensional signal-plus-noise stochastic process of the form $Y(t) = \mathbf{P}(N) X(t) + Z(t)$, where, for $p \geq M$, $Y(t), Z(t) \in \mathbb{R}^p$, $\mathbf{P}(N) \in \mathcal{M}(p, M, \mathbb{R})$ and $X(t) \in \mathbb{R}^M$. $X(t)$ is the latent fractional stochastic process (the "signal"), $Z(t)$ is the "noise" component, and $\mathbf{P}(N)$ is a full rank matrix of coordinates. So, first assume the dimension p is fixed and set to M. Now suppose $X(t)$ is an ofBm under the conditions of Theorem IV.1. Next set the matrix $\mathbf{P}(N)$ to the $M \times M$ identity matrix. In addition, suppose the noise term $Z(t)$ is identically zero. In order to establish (25), it now suffices to follow the argument of the proof of Theorem 3.1 in [30]. The convergence (26) can be established by an analogous argument. This shows (i). $\square$

Remark A.1. *The framework for proving Theorem IV.1 provides a great deal of information about the asymptotic rescaled eigenvalues* $\xi_m(2^j)$, $m \in \{1, \ldots, M\}$, *as in* (25). *More specifically, for some function* φ *and some vector* $\widetilde{\mathbf{w}}$, *we can write*

$$
p\text{-}\lim_{N \to \infty} \frac{\hat{\lambda}_m^{(\tau)}(a(N) 2^j)}{a(N)^{2H_m + 1}} = \xi_m(2^j) = \varphi(\widetilde{\mathbf{w}}). \tag{39}
$$

In order to describe $\varphi(\widetilde{\mathbf{w}})$, *we need a few definitions and constructs. So, for the fixed* m, *define the sets of indices* $\mathcal{I}_- := \{\ell : H_\ell < H_m\}$ *and*

$$
\mathcal{I}_0 := \{\ell : H_\ell = H_m\} \text{ and } \mathcal{I}_+ := \{\ell : H_\ell > H_m\}. \tag{40}
$$

Note that $\mathcal{I}_-$ *and* $\mathcal{I}_+$ *are possibly empty. Also, let*

$$
r_1 := \text{card}(\mathcal{I}_-), \ r_2 := \text{card}(\mathcal{I}_0) \geq 1, \ r_3 := \text{card}(\mathcal{I}_+) \tag{41}
$$

be their respective cardinalities. For every $\tau = 1, \ldots, 2^{j_2^0 - j}$, *define the auxiliary wavelet random matrix*

$$
\widehat{\mathbf{B}}_a^{(\tau)}(2^j) := \mathbf{W}^{-1} a^{-\mathbf{H} - \frac{1}{2}\mathbb{I}} \mathbf{S}^{(\tau)}(a 2^j) a^{-\mathbf{H}^* - \frac{1}{2}\mathbb{I}} (\mathbf{W}^*)^{-1}, \tag{42}
$$

where $a \equiv a(N)$. *Also, let* $\mathbf{B}(2^j) = \mathbb{E} \mathbf{B}_a(2^j)$, *which neither depends on* τ *nor on* a. *Recast*

$$
\widehat{\mathbf{B}}_a^{(\tau)}(2^j) = (\widehat{\mathbf{B}}_{i\ell})_{i, \ell = 1,2,3}, \quad \mathbf{B}(2^j) = (\mathbf{B}_{i\ell})_{i, \ell = 1,2,3}. \tag{43}
$$

In (43), $\mathbf{B}_{i\ell}$ *and* $\widehat{\mathbf{B}}_{i\ell}$ *denote blocks of size* $r_i \times r_\ell$, *and the scale* 2^j *is omitted wherever doing so is unambiguous. Similarly, define*

$$
\mathbf{W} = \begin{pmatrix} \mathbf{W}_1 & \mathbf{W}_2 & \mathbf{W}_3 \end{pmatrix}, \tag{44}
$$

where, for $i = 1, 2, 3$, $\mathbf{W}_i \in \mathcal{M}(p, r_i, \mathbb{R})$, *and each column vector is denoted by* $\mathbf{w}_\ell$, $\ell = 1, \ldots, r$. *Let* $\mathbf{\Pi}_3 \in \mathcal{M}(r, \mathbb{R})$ *be the projection matrix onto* $\text{Ker}\{\mathbf{W}_3^*\}$. *Also, let* $\mathbf{M} = \mathbf{B}_{22} - \mathbf{B}_{23} \mathbf{B}_{33}^{-1} \mathbf{B}_{32} \in \mathcal{S}(r_2, \mathbb{R})$ *and* $\mathbf{\Lambda} = \mathbf{\Pi}_3^* \mathbf{W}_2 \mathbf{M} \mathbf{W}_2^* \mathbf{\Pi}_3 \in \mathcal{S}_{\geq 0}(r, \mathbb{R})$. *For unit vectors* $\mathbf{u} \in \mathbb{R}^r$, *define the function*

$$
\varphi(\mathbf{u}) := \mathbf{u}^* \mathbf{\Lambda} \mathbf{u}. \tag{45}
$$

Then, for φ *as in* (45), *the proof of Proposition 5.1 and Lemma B.1 in [30] imply that there exists a possibly random vector*

$$
\widetilde{\mathbf{w}} = \widetilde{\mathbf{w}}(\omega) \in \text{span}\{\mathbf{\Lambda}\} = \text{span}\{\mathbf{W}_2, \mathbf{W}_3\} \cap \text{Ker}\{\mathbf{W}_3^*\} \tag{46}
$$

based on which we can express the limit (39) *as* $\varphi(\widetilde{\mathbf{w}})$ (***n.b.:*** $\xi_m(2^j)$ *is* deterministic). *In particular, relations* (45) *and* (46) *imply that* $\varphi(\mathbf{w})$ – *and, hence,* $\xi_m(2^j)$ – *only depends on the matrices* $\mathbf{B}_{22}$, $\mathbf{B}_{23}$, $\mathbf{B}_{33}$, $\mathbf{W}_2$ *and* $\mathbf{W}_3$. *Consequently, the asymptotic rescaled function in* (25) *only depends on parameters associated with the index sets* $\mathcal{I}_0$ *and* $\mathcal{I}_+$ *in* (40).

B. Proof of asymptotic normality

In the proof of Theorem IV.2, we write $a \equiv a(N)$ whenever notationally convenient. Also, we make use of the following notation. Let $\{X_N\}_{N \in \mathbb{N}}$, $\{Y_N\}_{N \in \mathbb{N}}$ be two sequences of random variables. We write $X_N \overset{d}{\sim} Y_N$, $N \to \infty$, when the two sequences have the same limit in distribution, which is assumed to be well defined.

Proof. First, we show (i). Fix $j \in \{j_1^0, \ldots, j_2^0\}$. The proof of (31) is by means of an adaptation of the proof of Theorem 3.2 in [30]. In fact, let $\widehat{\mathbf{B}}_a(2^j)$ and $\mathbf{B}(2^j)$ be as in (43). Under conditions (OFBM1-3), as a consequence of Theorem 3.1 in [25],

$$
\left\{ \sqrt{n_{a, j_2^0}} \text{vec}_{\mathcal{S}} \left(\widehat{\mathbf{B}}_a^{(\tau)}(2^j) - \mathbf{B}(2^j) \right) \right\}_{\tau = 1, \ldots, 2^{j_2^0 - j}} \overset{d}{\to} \mathcal{N}(0, \mathbf{\Sigma_B}) \tag{47}
$$

as $n \to \infty$ for some $\mathbf{\Sigma_B} \in \mathcal{S}_{\geq 0} \left(2^{j_2^0 - j} \cdot M(M+1)/2, \mathbb{R} \right)$. In particular,

$$
\widehat{\mathbf{B}}_a^{(\tau)}(2^j) - \mathbf{B}(2^j) \overset{\mathbb{P}}{\to} 0, \quad \tau = 1, \ldots, 2^{j_2^0 - j}. \tag{48}
$$

So, fix $m \in \{1, \ldots, M\}$. Recall that, for any matrix $\mathbb{M} \in \mathcal{S}(M, \mathbb{R})$, the differential of a *simple* eigenvalue $\lambda_m(\mathbb{M})$ exists in a connected vicinity of $\mathbb{M}$ and is given by

$$
d\lambda_m(\mathbb{M}) = \mathbf{u}_m^* \{d\mathbb{M}\} \mathbf{u}_m, \tag{49}
$$

where $\mathbf{u}_m$ is a unit eigenvector of $\mathbb{M}$ associated with $\lambda_m(\mathbb{M})$ (see [40], p. 182, Theorem 1). Now, for every $\mathbf{B} \in \mathcal{S}_{\geq 0}(M, \mathbb{R})$, define the function $f_{N,m}$ by

$$f_{N,m}(\mathbf{B}) = \log_2 \lambda_m\left(\frac{\mathbf{W}a^{\mathrm{diag}(\underline{H})}\mathbf{B}a^{\mathrm{diag}(\underline{H})}\mathbf{W}^*}{a^{2H_m}}\right). \quad (50)$$

For any fixed $\delta > 0$, let $\mathcal{O}_\delta = \{\mathbf{B} \in \mathcal{S}_{\geq 0}(M, \mathbb{R}) : \|\mathbf{B} - \mathbf{B}(2^j)\| < \delta\}$. By Proposition 3.1 in [25], $\det \mathbf{B}(2^j) > 0$. Thus, an application of the proof of Theorem 3.2 in [30] implies that there exists some $\delta_0 > 0$ such that, for large enough N and for any $\mathbf{B} \in \mathcal{O}_{\delta_0}$, the mean value theorem-type relation

$$f_{N,m}(\mathbf{B}) - f_{N,m}(\mathbf{B}(2^j))$$
$$= \sum_{1 \leq i_1 \leq i_2 \leq M} \frac{\partial}{\partial b_{i_1,i_2}} f_{N,m}(\breve{\mathbf{B}}) \cdot \pi_{i_1,i_2}(\mathbf{B} - \mathbf{B}(2^j)) \quad (51)$$

holds for some matrix $\breve{\mathbf{B}} \in \mathcal{S}_{>0}(M, \mathbb{R})$ lying in a segment connecting $\mathbf{B}$ and $\mathbf{B}(2^j)$ across $\mathcal{S}_{>0}(M, \mathbb{R})$. Define the event $A_N = \{\omega : \widehat{\mathbf{B}}_a^{(\tau)}(2^j) \in \mathcal{O}_{\delta_0}, \tau = 1, \ldots, 2^{j_2^0-j}\}$. By (48), $\mathbb{P}(A_N) \to 1$, $N \to \infty$. Hence, by (51), with probability going to 1, for large enough N, for $\tau = 1, \ldots, 2^{j_2^0-j}$ and $m = 1, \ldots, M$, expansion (51) holds with $\mathbf{B} = \widehat{\mathbf{B}}_a^{(\tau)}(2^j)$, $\breve{\mathbf{B}} = \breve{\mathbf{B}}_a^{(\tau)}(2^j) \in \mathcal{S}_{>0}(M, \mathbb{R})$. In addition, we can express, for any $i_1, i_2 \in \{1, \ldots, M\}$,

$$\frac{\partial}{\partial b_{i_1,i_2}} f_{N,m}(\breve{\mathbf{B}}_a^{(\tau)}(2^j))$$
$$= \frac{1}{\ln 2} \frac{a^{2H_m+1}}{\lambda_m(\breve{\mathbf{S}}^{(\tau)}(a2^j))} \frac{\partial}{\partial b_{i_1,i_2}} \lambda_m\left(\frac{\breve{\mathbf{S}}^{(\tau)}(a2^j)}{a^{2H_m+1}}\right), \quad (52)$$

where $\breve{\mathbf{S}}^{(\tau)}(a2^j) := \mathbf{W}a^{\mathrm{diag}(\underline{H})}\breve{\mathbf{B}}_a^{(\tau)}(2^j)a^{\mathrm{diag}(\underline{H})}\mathbf{W}^*$. Expressions (51) and (52) imply that, as $N \to \infty$,

$$\ln 2 \sqrt{n_{a,j_2^0}}\left(f_{N,m}(\widehat{\mathbf{B}}_a^{(\tau)}(2^j)) - f_{N,m}(\mathbf{B}_a(2^j))\right)$$
$$\stackrel{d}{\sim} \frac{\sqrt{n_{a,j_2^0}}}{\xi_m(2^j)} \sum_{i_1,i_2=r_1+1}^{M} \langle\mathbf{w}_{i_1}(N), \mathbf{u}_m(N)\rangle a^{H_{i_1}-H_m}$$
$$\cdot \langle\mathbf{w}_{i_2}(N), \mathbf{u}_m(N)\rangle a^{H_{i_2}-H_m} \pi_{i_1,i_2}(\widehat{\mathbf{B}}_a^{(\tau)}(2^j) - \mathbf{B}_a(2^j))$$
$$\stackrel{d}{\sim} \frac{\sqrt{n_{a,j_2^0}}}{\xi_m(2^j)}\left\{\sum_{i_1,i_2=r_1+1}^{r_1+r_2} + 2\sum_{i_1=r_1+1}^{r_1+r_2}\sum_{i_2=r_1+r_2+1}^{M}\right.$$
$$+ \left.\sum_{i_1,i_2=r_1+r_2+1}^{M}\right\}\langle\mathbf{w}_{i_1}(N), \mathbf{u}_m(N)\rangle a^{H_{i_1}-H_m}$$
$$\cdot\langle\mathbf{w}_{i_2}(N), \mathbf{u}_m(N)\rangle a^{H_{i_2}-H_m} \pi_{i_1,i_2}(\widehat{\mathbf{B}}_a^{(\tau)}(2^j) - \mathbf{B}(2^j)). \quad (53)$$

However, by Proposition B.1 in [30], under condition (30) we can suppose that there exists a deterministic unit vector $\mathbf{u}_m$ such that

$$\mathbf{u}_m(N) \stackrel{\mathbb{P}}{\to} \mathbf{u}_m. \quad (54)$$

So, let

$$\boldsymbol{\gamma}_m = \{\gamma_{m,i}\}_{i=1,\ldots,r} := p\text{-}\lim_{N\to\infty} \mathbf{W}^*\mathbf{u}_m(N) \in \mathbb{R}^r. \quad (55)$$

Also, by Lemma B.8 in [30], we can assume the existence of the limiting vector

$$\mathbf{x}_{*,m} = \{\mathbf{x}_{*,m,i}\}_{i\in\mathcal{I}_+} :=$$
$$p\text{-}\lim_{N\to\infty}\left\{\langle\mathbf{w}_i, \mathbf{u}_m(N)\rangle \cdot a(N)^{H_i-H_m}\right\}_{i\in\mathcal{I}_+} \in \mathbb{R}^{r_3}. \quad (56)$$

Therefore, (53) has the same limit in distribution as

$$\frac{\sqrt{n_{a,j_2^0}}}{\xi_m(2^j)}\left\{\sum_{i_1,i_2=r_1+1}^{r_1+r_2} \gamma_{m,i_1} \cdot \gamma_{m,i_2}\right.$$
$$+ 2\sum_{i_1=r_1+1}^{r_1+r_2}\sum_{i_2=r_1+r_2+1}^{M} \gamma_{m,i_1} \cdot \mathbf{x}_{*,m,i_2}$$
$$+ \left.\sum_{i_1,i_2=r_1+r_2+1}^{M} \mathbf{x}_{*,m,i_1}\cdot\mathbf{x}_{*,m,i_2}\right\} \cdot \pi_{i_1,i_2}(\widehat{\mathbf{B}}_a^{(\tau)}(2^j) - \mathbf{B}(2^j)). \quad (57)$$

Now note that (57) further holds for all τ, m and j, and also that these expansions start from the same sequence of asymptotically Gaussian random matrices $\widehat{\mathbf{B}}_a^{(\tau)}(2^j)$ (see (47)). Hence, the jointly asymptotically normal limit (31) holds. Thus, (i) is proven.

Statement (32) can be proven by naturally adapting the proof of Theorem 3.2 in [41], originally developed for the high-dimensional context (see the related comments in the proof of Theorem IV.1 in this paper). In particular, the definitions (23) and (24) are applied in a mathematically sharp way. This establishes (ii). $\quad\square$

Remark A.2. *For a single j, we can show that the limit in distribution* (31) *of wavelet log-eigenvalues is scale-free. Consider the vector $\boldsymbol{\gamma}_m$ as in* (55) *as well as the vector*

$$\boldsymbol{\varrho}_m := \mathbf{B}_{33}^{-1}(1)\mathbf{B}_{32}(1)\pi_{r_1+1,\ldots,r_1+r_2}(\boldsymbol{\gamma}_m) \in \mathbb{R}^{r_3}. \quad (58)$$

Assume for the moment that

$$\sqrt{n_{a,j_2^0}}\left\{\log_2 \hat{\lambda}_m^{(\tau)}(a(N)2^j) - \log_2 \lambda_m^{(\tau)}(a(N)2^j)\right\}$$
$$\stackrel{d}{\sim} \frac{1}{\ln 2} \frac{\sqrt{n_{a,j_2^0}}}{\xi_m(1)}\left\{\sum_{i_1,i_2=r_1+1}^{r_1+r_2} \gamma_{m,i_1} \cdot \gamma_{m,i_2}\right.$$
$$+ 2\sum_{i_1=r_1+1}^{r_1+r_2}\sum_{i_2=r_1+r_2+1}^{M} \gamma_{m,i_1} \cdot \varrho_{m,i_2}$$
$$+ \left.\sum_{i_1,i_2=r_1+r_2+1}^{M} \varrho_{m,i_1} \cdot \varrho_{m,i_2}\right\} \cdot \pi_{i_1,i_2}(\widehat{\mathbf{B}}_a^{(\tau)}(1) - \mathbf{B}(1)). \quad (59)$$

In light of (47), *it is clear that* (59) *converges to a Gaussian law that does not depend on j, as claimed.*

To establish (59), *recast entry-wise* $\mathbf{B}(2^j) =: \left(b_{\ell_1,\ell_2}(2^j)\right)_{\ell_1,\ell_2=1,\ldots,r}$ *(cf.* (43)). *By Proposition 3.1, (P5), in [25],*

$$\mathbf{B}(2^j) = \left((2^j)^{H_{\ell_1}+H_{\ell_2}+1} b_{\ell_1,\ell_2}(1)\right)_{\ell_1,\ell_2=1,\ldots,r}. \quad (60)$$

Bearing in mind (54) *and* (58)*, by the proofs of Lemma B.8 and Proposition 5.1 in* [30]*, and by* (60)*, we can recast* (56) *in the form*

$$
\begin{aligned}
\mathbf{x}_{m,*} &= -\mathbf{B}_{33}^{-1}(2^j)\mathbf{B}_{32}(2^j)\mathbf{W}_2^*\mathbf{u}_m \\
&= -(2^j)^{H_m}\mathrm{diag}\big((2^j)^{-H_{r_1+r_2+1}},\dots,(2^j)^{-H_r}\big) \\
&\quad \mathbf{B}_{33}^{-1}(1)\mathbf{B}_{32}(1)\mathbf{W}_2^*\mathbf{u}_m \\
&= -(2^j)^{H_m}(2^j)^{-\mathrm{diag}(H_{r_1+r_2+1},\dots,H_r)}\boldsymbol{\varrho}_m.
\end{aligned}
\tag{61}
$$

Rewrite, entry-wise,

$$
\widehat{\mathbf{B}}_a^{(\tau)}(2^j) = \big(\widehat{b}_{\ell_1,\ell_2}^{(\tau)}(2^j)\big)_{\ell_1,\ell_2=1,\dots,r}.
\tag{62}
$$

By an adaptation of the proof of Proposition 3.1, (P6), in [25] *we obtain*

$$
\widehat{\mathbf{B}}_a^{(\tau)}(2^j) \stackrel{d}{=} \big((2^j)^{H_{\ell_1}+H_{\ell_2}+1}\widehat{b}_{\ell_1,\ell_2}^{(\tau)}(1)\big)_{\ell_1,\ell_2=1,\dots,r}.
\tag{63}
$$

So, by (53)*,* (60)*,* (61) *and* (63)*, as* $N \to \infty$*,*

$$
\begin{aligned}
&\sqrt{n_{a,j_2^0}}\big\{\log_2\hat{\lambda}_m^{(\tau)}(a(N)2^j) - \log_2\lambda_m^{(\tau)}(a(N)2^j)\big\} \\
&\stackrel{d}{\sim} \frac{1}{\ln 2}\frac{\sqrt{n_{a,j_2^0}}}{\xi_m(2^j)}\Bigg\{\sum_{i_1=r_1+1}^{r_1+r_2}\sum_{i_2=r_1+1}^{r_1+r_2}\gamma_{m,i_1}\gamma_{m,i_2}(2^j)^{2H_m+1} \\
&\quad + 2\sum_{i_1=r_1+1}^{r_1+r_2}\sum_{i_2=r_1+r_2+1}^{M}\gamma_{m,i_1}[(2^j)^{H_m}(2^j)^{-H_{i_2}}] \\
&\qquad \boldsymbol{\varrho}_{m,i_2}(2^j)^{H_m+H_{i_2}+1} \\
&\quad + \sum_{i_1=r_1+r_2+1}^{M}\sum_{i_2=r_1+r_2+1}^{M}[(2^j)^{H_m}(2^j)^{-H_{i_1}}]\boldsymbol{\varrho}_{m,i_1} \\
&\qquad [(2^j)^{H_m}(2^j)^{-H_{i_2}}]\boldsymbol{\varrho}_{m,i_2}(2^j)^{H_{i_1}+H_{i_2}+1}\Bigg\} \\
&\quad \cdot \pi_{i_1,i_2}\big(\widehat{\mathbf{B}}_a^{(\tau)}(1) - \mathbf{B}(1)\big).
\end{aligned}
$$

In view of the scaling relation in (25)*, relation* (59) *holds, as claimed.*

C. Approximations of log-eigenvalue covariances

Fix $j \in \{j_1^0,\dots,j_2^0\}$. In this appendix, we use the bivariate context to provide some justification for the very useful approximations

$$
\mathrm{corr}(\log_2\hat{\lambda}_m(a(N)2^j), \log_2\hat{\lambda}_{m'}(a(N)2^j)) \approx 0,
\tag{64}
$$

$$
\mathrm{corr}(\log_2\hat{\lambda}_m^{(\tau)}(a(N)2^j), \log_2\hat{\lambda}_{m'}^{(\tau)}(a(N)2^j)) \approx 0,
\tag{65}
$$

as well as

$$
\mathrm{Var}\big(\log_2\hat{\lambda}_m(a(N)2^j)\big) \approx 2(\log_2 e)^2/n_{a,j},
\tag{66}
$$

$$
\mathrm{Var}\big(\log_2\hat{\lambda}_m^{(\tau)}(a(N)2^j)\big) \approx 2(\log_2 e)^2/n_{a,j_2^0}.
\tag{67}
$$

So, consider the case $M = 2$ and $0 < H_1 < H_2 < 1$. Since the arguments for showing all relations (64)-(67) are similar, calculations are detailed for the bias-corrected eigenvalues $\hat{\lambda}_m^{(\tau)}$ only. Let $\widehat{\mathbf{B}}_a(2^j)$ be as in (62) and further write entry-wise $\mathbf{B}(1) = \mathbb{E}\widehat{\mathbf{B}}_a(1) = \{b_{\ell_1\ell_2}(1)\}_{\ell_1,\ell_2=1,2}$ (cf. (43)). Assume

$$
b_{12}^2(1) > 0.
\tag{68}
$$

In view of (59), it can be shown that we can write the expansions

$$
\begin{aligned}
&\sqrt{n_{a,j_2^0}}\big(\ln\hat{\lambda}_1^{(\tau)}(a(N)2^j) - \ln\lambda_1^{(\tau)}(a(N)2^j)\big) \\
&\stackrel{d}{\sim} \frac{b_{11}(1)b_{22}(1)}{\det\mathbf{B}(1)}\sqrt{n_{a,j_2^0}}\cdot\frac{(\widehat{b}_{11}(1)-b_{11}(1))}{b_{11}(1)} \\
&\quad + (-2)\frac{b_{12}^2(1)}{\det\mathbf{B}(1)}\sqrt{n_{a,j_2^0}}\cdot\frac{(\widehat{b}_{12}(1)-b_{12}(1))}{b_{12}(1)} \\
&\quad + \frac{b_{12}^2(1)}{\det\mathbf{B}(1)}\sqrt{n_{a,j_2^0}}\cdot\frac{(\widehat{b}_{22}(1)-b_{22}(1))}{b_{22}(1)}
\end{aligned}
\tag{69}
$$

and

$$
\begin{aligned}
&\sqrt{n_{a,j_2^0}}\big(\ln\hat{\lambda}_2^{(\tau)}(a(N)2^j) - \ln\lambda_2^{(\tau)}(a(N)2^j)\big) \\
&\qquad \stackrel{d}{\sim} \sqrt{n_{a,j_2^0}}\cdot\frac{(\widehat{b}_{22}(1)-b_{22}(1))}{b_{22}(1)}
\end{aligned}
\tag{70}
$$

(see relations (4.33) and (4.35) in [25]). So, for the sake of computation, we normalize $b_{11}(1) = 1 = b_{22}(1)$. Also, for notational simplicity, we write $X_k = \pi_1(\mathbf{W}^{-1}D(1,k))$, $Y_k = \pi_2(\mathbf{W}^{-1}D(1,k))$, $\forall k \in \mathbb{Z}$, where π_1 and π_2 denote the projection onto the ℓ-coordinate. For ofBm, the celebrated property of approximate decorrelation in the wavelet domain is given by Proposition 3.2 in [25]. It implies that, for $k \neq k'$,

$$
\mathbb{E}X_kX_{k'} \approx 0, \quad \mathbb{E}Y_kY_{k'} \approx 0, \quad \mathbb{E}X_kY_{k'} \approx 0.
\tag{71}
$$

In view of (71), of the stationarity of fixed-scale wavelet coefficients (see [25], Proposition 3.1, $(P2)$) and of the Isserlis theorem, we obtain the approximations

$$
\mathrm{Cov}\big(\widehat{b}_{11}(1),\widehat{b}_{22}(1)\big) \approx \frac{2b_{12}^2(1)}{n_{a,j_2^0}},
\tag{72}
$$

$$
\mathrm{Cov}\big(\widehat{b}_{12}(1),\widehat{b}_{22}(1)\big) \approx \frac{2b_{12}(1)}{n_{a,j_2^0}}, \quad \mathrm{Var}\big(\widehat{b}_{22}(1)\big) \approx \frac{2}{n_{a,j_2^0}}.
\tag{73}
$$

As a consequence of the expansions (69) and (70), as well as of the approximations (72) and (73),

$$
\begin{aligned}
&\mathrm{Cov}\Big(\sqrt{n_{a,j_2^0}}\big(\ln\hat{\lambda}_1^{(\tau)}(a(N)2^j) - \ln\lambda_1^{(\tau)}(a(N)2^j)\big), \\
&\qquad\qquad \sqrt{n_{a,j_2^0}}\big(\ln\hat{\lambda}_2^{(\tau)}(a(N)2^j) - \ln\lambda_2^{(\tau)}(a(N)2^j)\big)\Big) \\
&\approx \frac{n_{a,j_2^0}}{\det\mathbf{B}(1)}\Big\{\frac{2b_{12}^2(1)}{n_{a,j_2^0}} - 2\cdot\frac{b_{12}^2(1)}{b_{12}(1)}\frac{2b_{12}(1)}{n_{a,j_2^0}} + \frac{2b_{12}^2(1)}{n_{a,j_2^0}}\Big\} = 0.
\end{aligned}
\tag{74}
$$

Namely, $\log_2\hat{\lambda}_1^{(\tau)}(a(N)2^j)$ and $\log_2\hat{\lambda}_2^{(\tau)}(a(N)2^j)$ are approximately decorrelated over large scales for large values of n_{a,j_2^0}.

Alternatively, suppose condition (68) is dropped, namely, assume $b_{12}^2(1) = 0$. Then, after appropriately eliminating null denominators, the expansions (69) and (70) and the approximation (72) immediately show that the log-eigenvalues are approximately decorrelated.

Next, we develop approximations for the variances. Similarly to (72) and (73), we obtain

$$
\mathrm{Cov}\big(\widehat{b}_{11}(1),\widehat{b}_{12}(1)\big) \approx \frac{2b_{12}(1)}{n_{a,j_2^0}},
\tag{75}
$$

$$
\mathrm{Var}\big(\widehat{b}_{11}(1)\big) \approx \frac{2}{n_{a,j_2^0}}, \quad \mathrm{Var}\big(\widehat{b}_{12}(1)\big) \approx \frac{b_{12}^2(1)}{n_{a,j_2^0}} + \frac{1}{n_{a,j_2^0}}.
\tag{76}
$$

Hence,

$$\text{Var}\left(\sqrt{n_{a,j_2^0}}\left(\ln\hat{\lambda}_1^{(\tau)}(a(N)2^j) - \ln\lambda_1^{(\tau)}(a(N)2^j)\right)\right)$$

$$\approx \frac{n_{a,j_2^0}}{\det\mathbf{B}(1)^2}\left[\text{Var}\,\hat{b}_{11}(1) + 4b_{12}^2(1)\text{Var}\,\hat{b}_{12}(1)\right.$$

$$+ b_{12}^4(1)\text{Var}\,\hat{b}_{22}(1) - 4b_{12}(1)\text{Cov}(\hat{b}_{11}(1),\hat{b}_{12}(1))$$

$$\left. + 2b_{12}(1)^2\text{Cov}(\hat{b}_{11}(1),\hat{b}_{22}(1)) - 4b_{12}^3(1)\text{Cov}(\hat{b}_{12}(1),\hat{b}_{22}(1))\right]$$

$$= \frac{n_{a,j_2^0}}{(1-b_{12}^2(1))^2}\left[\frac{2}{n_{a,j_2^0}} + 4b_{12}^2(1)\left(\frac{b_{12}^2(1)}{n_{a,j_2^0}} + \frac{1}{n_{a,j_2^0}}\right)\right.$$

$$+ b_{12}^4(1)\frac{2}{n_{a,j_2^0}} - 4b_{12}(1)\frac{2b_{12}(1)}{n_{a,j_2^0}} + 2b_{12}^2(1)\frac{2b_{12}^2(1)}{n_{a,j_2^0}}$$

$$\left. - 4b_{12}^3(1)\frac{2b_{12}(1)}{n_{a,j_2^0}}\right] = 2$$

$$(77)$$

and

$$\text{Var}\left(\sqrt{n_{a,j_2^0}}\left(\ln\hat{\lambda}_2^{(\tau)}(a(N)2^j) - \ln\lambda_2^{(\tau)}(a(N)2^j)\right)\right)$$

$$\approx n_{a,j_2^0}\cdot\frac{2}{n_{a,j_2^0}} = 2. \tag{78}$$

Analogously, we can show that

$$\text{Var}\left(\sqrt{n_{a,j}}\left(\ln\hat{\lambda}_2(a(N)2^j) - \ln\lambda_2(a(N)2^j)\right)\right) \approx 2. \tag{79}$$

Remark A.3. *Relations (64)-(67) provide approximations for the second moments of wavelet log-eigenvalues that depend neither on model parameters nor on the scale. Beyond these approximations, Remark A.2 shows rigorously and more broadly that the limit in distribution (31) of wavelet log-eigenvalues for a single j is scale-free.*

D. Additional covariance approximations for multivariate estimation

By (64), pairs of multivariate estimates $\hat{H}_m^{\text{M}}$ and $\hat{H}_{m'}^{\text{M}}$, $m \neq m'$, are approximately decorrelated. On the other hand, by (65), the same holds for pairs of their bias-corrected counterparts $\hat{H}_m^{\text{M,bc}}$ and $\hat{H}_{m'}^{\text{M,bc}}$. In summary, for $m \neq m'$,

$$\text{corr}(\hat{H}_m^{\text{M}}, \hat{H}_{m'}^{\text{M}}) \approx 0, \quad \text{corr}(\hat{H}_m^{\text{M,bc}}, \hat{H}_{m'}^{\text{M,bc}}) \approx 0.$$

On the other hand, by (65), we can further approximate

$$\text{Var}(2\hat{H}_m^{\text{M,bc}}) = \text{Var}\left(\sum_{j=j_1^0}^{j_2^0} w_j\overline{\log_2\lambda_m}(a(N)2^j)\right)$$

$$\approx \sum_{j=j_1^0}^{j_2^0} w_j^2\,\text{Var}\left(\frac{1}{2^{j_2^0-j}}\sum_{\tau=1}^{2^{j_2^0-j}}\log_2\hat{\lambda}_m^{(\tau)}(a(N)2^j)\right)$$

$$\approx \sum_{j=j_1^0}^{j_2^0}\frac{w_j^2}{2^{j_2^0-j}}\text{Var}\left(\log_2\hat{\lambda}_m^{(1)}(a(N)2^j)\right)$$

$$\approx 2(\log_2 e)^2\sum_{j=j_1^0}^{j_2^0}2^{j-j_2^0}\frac{w_j^2}{n_{a,j_2^0}} = 2(\log_2 e)^2\sum_{j=j_1^0}^{j_2^0}\frac{w_j^2}{n_{a,j}}, \tag{80}$$

where the last line results from $n_{a,j}2^j = n_{a,j_2^0}2^{j_2^0}$.

Similarly, by (64),

$$\text{Var}(2\hat{H}_m^{\text{M}}) \approx 2(\log_2 e)^2\sum_{j=j_1^0}^{j_2^0}\frac{w_j^2}{n_{a,j}}. \tag{81}$$