Detecting and Estimating Multivariate Self-Similar Sources in High-Dimensional Noisy Mixtures

Patrice Abry, Herwig Wendt, Gustavo Didier

DETECTING AND ESTIMATING MULTIVARIATE SELF-SIMILAR SOURCES IN HIGH-DIMENSIONAL NOISY MIXTURES

Patrice Abry[1], *Herwig Wendt*[2], *Gustavo Didier*[3]

[1] Univ Lyon, Ens de Lyon, Univ Claude Bernard, CNRS, Laboratoire de Physique, Lyon, France.
[2] IRIT, CNRS (UMR 5505) , Université de Toulouse, France.
[3] Math. Dept., Tulane University, New Orleans, USA.

ABSTRACT

Nowadays, because of the massive and systematic deployment of sensors, systems are routinely monitored via a large collection of time series. However, the actual number of sources driving the temporal dynamics of these time series is often far smaller than the number of observed components. Independently, self-similarity has proven to be a relevant model for temporal dynamics in numerous applications. The present work aims to devise a procedure for identifying the number of multivariate self-similar mixed components and entangled in a large number of noisy observations. It relies on the analysis of the evolution across scales of the eigenstructure of multivariate wavelet representations of data, to which model order selection strategies are applied and compared. Monte Carlo simulations show that the proposed procedure permits identifying the number of multivariate self-similar mixed components and to accurately estimate the corresponding self-similarity exponents, even at low signal to noise ratio and for a very large number of actually observed mixed and noisy time series.

***Index Terms*—** multivariate self-similarity, operator fractional Brownian motion, wavelet spectrum, model order selection

1. INTRODUCTION

Context: multiple sensors versus few sources. Recent technological developments have permitted the massive production of low cost sensors with low cost deployment, as well as the easy storage and processing of large amounts of data, the so-called *data deluge*. Therefore, it is very common that one same system is monitored by a large number of sensors producing multivariate and dependent data, i.e., a large number of time series recorded together, whose joint temporal dynamics contain information about the system under scrutiny. This is the case in numerous applications that are very different in nature, such as macroscopic brain activity, where the number of observed time series ranges from hundreds (MEG data) to several tens of thousands (fMRI data), cf., e.g., [1, 2], or in meteorology and climate studies, where the use of large numbers of measured components has become standard (e.g., [3]). However, the number of actual mechanisms, or sources, driving the spatio-temporal dynamics of one single system is usually far smaller than the number of recorded time series (sparsity). Therefore, it has become a research trend in data analysis to develop methods for identifying the actual number of sources driving the dynamics of a given system, as well as to extract the information provided by these sources from the usually noisy observations.

Work supported by Grant ANR-16-CE33-0020 MultiFracs. G.D. was partially supported by the ARO grant W911NF-14-1-0475.

Context: multivariate self-similarity. Scale-free dynamics has proven to be a useful concept to model real-world data produced from numerous applications in several branches of science and technology (e.g., [4, 5] and references therein). In essence, the scale-free paradigm assumes that temporal dynamics are not governed by one or a few characteristic time scales but, instead, by a large continuum of time scales that jointly contribute to the temporal dependencies. Self-similarity provides a formal, versatile model for scale-free dynamics [6]. In a stochastic processes framework, fractional Brownian motion (fBm) [7] has been widely and successfully used to model real-world data characterized by scale-free dynamics. While the modeling of the latter in applications has remained so far based on the univariate fBm model, a multivariate extension called operator fractional Brownian motion (ofBm) was recently proposed [8–11] to permit their modeling in a multivariate time series setting [12]. The availability and potential of this multivariate self-similarity model for describing real-world data immediately raises a first yet critical question: *how many different self-similar components actually exist amongst the possibly very large number of time series recorded by sensors?*

Related work. The general issues of identifying the number of actual sources in potentially high-dimensional multivariate signals, and of characterizing them from noisy observations, has continuously been addressed in statistical signal processing over the years. Often embedded in the general themes of dimensionality reduction, source separation [13], or model order selection and system identification [14–17], numerous technically distinct solutions were proposed in theory and widely used in applications. Examples include principal component analysis (PCA), factor analysis, sparse graphical Gaussian models, etc. (cf., e.g., [18]). Independently, a multivariate wavelet variance eigenvalue-based representation of multivariate self-similarity was recently put forward and studied [12]. Notably, it was shown to lead to efficient and robust estimation of self-similarity exponents in intricate multivariate settings [19]. Notwithstanding significant success in applications, none of the classical dimension reduction tools takes explicitly into account multivariate scale-free dynamics (see [20, 21] for exceptions). Thus, there is a great paucity of methodological and practical data modeling tools that are both inherently *multivariate* and *scale-free analytical*.

Goal, contributions and outline. The present work is a first contribution aiming at identifying and characterizing multivariate self-similar components embedded in a large set of noisy observations. It relies on ofBm, whose definition and properties are recalled in Section 2.1, as well as on the recently proposed multivariate eigenvalue wavelet representation, detailed in Section 2.2. The model, comprising a high-dimensional mixture of low-dimensional multivariate self-similar sources in additive noise, is described in Section 3.1.

Section 3.2 contains the proposed model order selection procedure. It relies on information theoretic tools constructed and studied in different contexts [15–17] and adapted here to the multivariate and multiscale eigenvalue wavelet representation of the noisy observations. Using large size Monte Carlo experiments, the performance of the devised model order selection methods is assessed in computational practice for various instances of ofBm (combinations of scaling exponents, mixing matrices and covariance), number of actual self-similar components versus observed time series, noise level and sample size. The method's performance, depicted and analyzed in Section 4, is shown to be promising in its ability to identify the number of ofBm components amidst noisy observations and to estimate the corresponding multiple self-similarity exponents.

2. OPERATOR FRACTIONAL BROWNIAN MOTION AND WAVELET ESTIMATION OF SCALING EXPONENTS

2.1. Operator fractional Brownian motion

The present work makes uses of ofBm for modeling multivariate self-similar structures in real-world data, whose definition and properties are briefly recalled here (the interested reader is referred to [10] for the most general definition and properties of ofBm). Let $\underline{B}_{\underline{H},\Sigma}(t) = \big(B_{H_1}(t),\ldots,B_{H_M}(t)\big)_{t\in\mathbb{R}}$ denote a collection of M possibly correlated fBm components defined by their individual self-similarity exponents $\underline{H} = (H_1,\ldots,H_M)$, $0 < H_1 \leq \ldots \leq H_M < 1$. Let Σ be a pointwise covariance matrix with entries $(\Sigma)_{m,m'} = \sigma_m\sigma_{m'}\rho_{m,m'}$, where σ_m^2 are the variances of the components and $\rho_{m,m'}$ their pairwise Pearson correlation coefficients. We define ofBm as the stochastic process $\underline{B}_{P,\underline{H},\Sigma}(t) \triangleq P\underline{B}_{\underline{H},\Sigma}(t)$, where P is a real-valued, $M \times M$ invertible matrix that mixes the components (changes the scaling coordinates) of $\underline{B}_{\underline{H},\Sigma}(t)$. In this case, ofBm consists of a multivariate Gaussian self-similar process with stationary increments. Moreover, it satisfies the (operator) self-similarity relation

$$\{\underline{B}_{P,\underline{H},\Sigma}(t)\}_{t\in\mathbb{R}} \stackrel{\text{fdd}}{=} \{a^{\underline{\underline{H}}}\underline{B}_{P,\underline{H},\Sigma}(t/a))\}_{t\in\mathbb{R}}, \tag{1}$$

$\forall a > 0$, with matrix exponent $\underline{\underline{H}} = P\text{diag}(\underline{H})P^{-1}$, termed Hurst matrix parameter, $a^{\underline{\underline{H}}} \triangleq \sum_{k=0}^{+\infty}\log^k(a)\underline{\underline{H}}^k/k!$, where $\stackrel{\text{fdd}}{=}$ stands for the equality of finite dimensional distributions. It is well documented that the scaling exponents $\underline{H}$ (Hurst eigenvalues) and the covariance matrix Σ cannot be chosen independently [10, 22]. For simplicity, hereafter we denote

$$Y(t) \triangleq \underline{B}_{P,\underline{H},\Sigma}(t).$$

2.2. Wavelet based scaling exponent estimation

Multivariate wavelet transform. The multivariate DWT of the multivariate process $\{Y(t)\}_{t\in\mathbb{R}}$ is defined as $D_Y(2^j,k) \triangleq (D_{Y_1}(2^j,k),\ldots,D_{Y_M}(2^j,k))$, $\forall k \in \mathbb{Z}$, $\forall j \in \mathbb{Z}$, and $m \in \{1,\ldots,M\}$, with $D_{Y_m}(2^j,k) \triangleq \langle 2^{-j/2}\psi(2^{-j}t-k)|Y_m(t)\rangle$, where ψ_0 denotes the mother wavelet. For a detailed introduction to wavelet transforms, interested readers are referred to, e.g., [23].

Multivariate self-similarity in the wavelet domain. It can be shown that the wavelet coefficients $\{D_Y(2^j,k)\}_{k\in Z}$ satisfy the (operator) self-similarity relation [12, 19]

$$\{D_Y(2^j,k)\}_{k\in\mathbb{Z}} \stackrel{\text{fdd}}{=} \{2^{j(\underline{\underline{H}}+\frac{1}{2}I)}D_Y(1,k)\}_{k\in\mathbb{Z}}, \tag{2}$$

for every fixed octave j. When $P = I$ (no mixing), the (canonical) components of ofBm are generally correlated fBm processes. In this case, M-variate (operator) self-similarity in the wavelet domain boils down to M entrywise self-similarity relations [22]

$$\begin{aligned}&\{D_{Y_1}(2^j,k),\ldots,D_{Y_M}(2^j,k)\}_{k\in\mathbb{Z}}\\ &\stackrel{\text{fdd}}{=} \{2^{j(H_1+\frac{1}{2})}D_{Y_1}(1,k),\ldots,2^{j(H_M+\frac{1}{2})}D_{Y_M}(1,k)\}_{k\in\mathbb{Z}}.\end{aligned} \tag{3}$$

Estimation of $\underline{H}$. Extending univariate wavelet analysis, it is natural to consider the empirical wavelet spectrum (variance), which is given by the $M \times M$ matrices

$$S_Y(2^j) = \frac{1}{n_j}\sum_{k=1}^{n_j} D_Y(2^j,k)D_Y(2^j,k)^*, \quad n_j = \frac{N}{2^j},$$

where N is the data sample size. Proceeding as in the univariate setting leads to the estimators [22]

$$\hat{H}_m^{(U)} = \left(\sum_{j=j_1}^{j_2} w_j \log_2 S_{Y,(m,m)}(2^j)\right)\Big/2 - \frac{1}{2}, \forall m. \tag{4}$$

Starting from (3), namely, when there is no mixing ($P = I$), $\hat{H}_m^{(U)}$ is expected to be a good estimator of H_m. However, starting from (2), i.e., a general mixing matrix P (non-canonical coordinates) and Hurst matrix parameter $\underline{H}$, it is clear that this is no longer the case (cf., [12, 19] and Section 4.1 for further discussions).

These observations lead to the definition of an estimator for $\underline{H}$ that is relevant in the general setting of non-diagonal P using the eigenvalues $\Lambda_Y(2^j) = \{\lambda_1(2^j),\ldots,\lambda_M(2^j)\}$ of $S_Y(2^j)$. The estimators $\underline{\hat{H}} = (\hat{H}_1,\ldots,\hat{H}_M)$ for $(H_1,\ldots,H_M)$ are defined by means of weighted linear regressions across scales $2^{j_1} \leq a \leq 2^{j_2}$

$$\hat{H}_m = \left(\sum_{j=j_1}^{j_2} w_j \log_2 \lambda_m(2^j)\right)\Big/2 - \frac{1}{2}, \forall m. \tag{5}$$

It is shown both theoretically and in practice that $\underline{\hat{H}}$ benefits from satisfactory performance (consistency, asymptotic normality), see [12, 19].

3. MODEL ORDER SELECTION FOR HIGH-DIMENSIONAL NOISY MIXTURES

3.1. High-dimensional noisy observations

To model the embedding of ofBm Y in a large set of noisy observations Z, let now W denote a matrix of size $L \times M$, $L \geq M$, with full rank M, and let $\mathcal{N}(t) = \big(\mathcal{N}_1(t),\ldots,\mathcal{N}_L(t)\big)_{t\in\mathbb{R}}$ be a collection of L independent vectors, each consisting of i.i.d. standardized Gaussian variables. The L observed noisy time series are modeled as

$$Z(t) \triangleq WY(t) + \sigma_{\mathcal{N}}\mathcal{N}(t), \tag{6}$$

where $\sigma_{\mathcal{N}}^2$ controls the variance of the noise. Adhering to the convention that W and P are normalized such that WP has rows with unit norm, the Signal-to-Noise Ratio (SNR) is defined as

$$\text{SNR} \triangleq \sum_{m=1}^{M}\sigma_m\Big/(L\sigma_{\mathcal{N}}) \tag{7}$$

where σ_m, $m = 1,\ldots,M$, is defined in Section 2.1.

3.2. Model order selection procedure

The proposed procedure relies on the eigenstructure of the empirical wavelet spectra $S_Z(2^j)$ of the multivariate wavelet representation of the L-variate observations Z. It is obvious that, in the absence of noise (i.e., SNR $= +\infty$), $S_Z(2^j)$ has rank M for all scales 2^j (as long as $n_j > M$). Conversely, when noise is added (SNR $< +\infty$), $S_Z(2^j)$ has rank L. Hence, heuristically speaking, when the SNR is sufficiently large, $S_Z(j)$ possess $L - M$ small eigenvalues corresponding to noise, and M (large) eigenvalues that correspond to the hidden ofBm components. Then, M can be estimated by using classical eigenvalue-based model order selection criteria, $\hat{M} = \phi_N((\lambda_1, \ldots, \lambda_L))$, such as Akaike's Information Criterion (AIC) or Minimum Description Length (MDL), see, e.g., [15–17] for details and more references.

However, in the context of multivariate self-similarity, wavelet spectra matrices $S_Z(2^j)$ are available corresponding to each octave and must be combined. To this end, we study two strategies, which we refer to as *majority vote* (MV) and *rescaled-average* (RS). As a common preprocessing step, the wavelet spectra $S_Z(2^j)$ of the L-variate observations Z and their eigenvalues $\Lambda_Z(2^j)$ of $S_Z(2^j)$ are computed, estimates $\hat{H}_l$ are obtained using (5), and the *rescaled eigenvalues* $(\overline{\Lambda}_Z(2^j))_l \triangleq 2^{-2j\hat{H}_l}(\Lambda_Z(2^j))_l$, $l = 1, \ldots, L$ are computed. The rationale behind $\overline{\Lambda}_Z(2^j)$ is the use the multivariate self-similarity of the sources for reducing the variance of their eigenvalues at each scale and for enabling averaging across scales.

For some ϕ_N, the model order selection strategies are defined as follows.

MV: A model order estimate $\hat{M}(j)$ is computed for $\overline{\Lambda}_Z(2^j)$

$$\hat{M}(j) = \phi_{n_j}(\overline{\Lambda}_Z(2^j))$$

for each scale j. Then the scale-wise estimates $\hat{M}_j$ are combined by majority vote (where $\mathbf{1}()$ is the indicator funtion)

$$\hat{M}_{MV} = \arg\max_{m=0,\ldots,L} \sum_{j=j_1}^{j_2} \mathbf{1}(\hat{M}(j) - m).$$

RS: The rescaled eigenvalues $\overline{\Lambda}_Z(2^j)$ are averaged across scales

$$\overline{\Lambda}_Z = \sum_{j=j_1}^{j_2} \overline{\Lambda}_Z(2^j).$$

A model order estimate $\hat{M}$ is computed from the averages

$$\hat{M}_{RS} = \phi_N(\overline{\Lambda}_Z).$$

For both strategies, the AIC classical model order selection is used

$$\phi_N(\Lambda) = \arg\min_k -N(\alpha - k)\log\big(g(k)/a(k)\big) + k(2\alpha - k), \quad (8)$$

where $\alpha = \min(L, N)$, $a(k)$, $g(k)$ are the arithmetic and geometric mean of the k smallest values of Λ, respectively.

Once the number of ofBm components $\hat{M}$ is estimated, a scaling exponents vector $(\hat{H}_1, \ldots, \hat{H}_{\hat{M}})$ can be computed using the estimation procedure described in (5), applied to the $\hat{M}$ largest values of $\Lambda_Z(2^j)$, where $\Lambda_Z(2^j)$ denotes the eigenvalues of $S_Z(2^j)$.

4. PERFORMANCE ASSESSMENT

We apply the ofBm model order selection and parameter estimation procedure to 1000 independent realizations of ofBm with $M = 4$ components of length $N = 2^{14}$ using various SNR values, as defined in (7). We consider 4 different sizes for the mixing matrix W, $L \in (10, 20, 50, 100)$; W is drawn at random for each realization of ofBm and normalized such that WP has unit norm rows. The ofBm parameters are set to $\underline{H} = (0.2, 0.5, 0.7, 0.9)$ and $\Sigma = \text{Toeplitz}(1, 0.2, 0.2, 0.3)$. We use the Daubechies wavelet with $N_\psi = 2$, and pick $(j_1, j_2) = (4, 9)$ in (5).

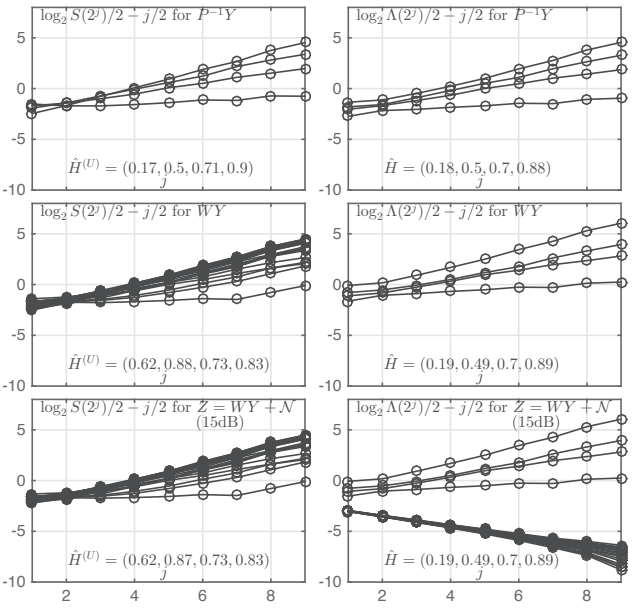

Fig. 1. **Illustration of estimation procedure.** Univariate analysis (left column) and multivariate analysis (right column) for single realization of ofBm with $M = 4$ components ($\underline{H} = (0.2, 0.5, 0.7, 0.9)$). Top row: without mixing and noise; Second row: with mixing and embedding in $L = 20$-dimensional noise; Bottom row: mixed $L = 20$ noisy time series (15dB SNR).

4.1. Univariate vs multivariate estimation

Fig. 1 plots univariate structure functions (i.e., the wavelet spectra $(S(2^j))_{mm}$; left column) and the multivariate estimates $\Lambda(2^j)$ (right column) as a function of the octave j for (from top to bottom): a single realization of unmixed ofBm $P^{-1}Y$ with $M = 4$ components without noise; mixture WY of dimension $L = 20$; mixture with additive white Gaussian noise Z (with SNR of 15dB). The results show that for unmixed noise-free ofBm in canonical coordinates, $P^{-1}Y$, the univariate and multivariate estimates $(S(2^j))_{mm}$ and $(\Lambda(2^j))_m$ and the estimated exponents H_m are essentially identical. However, for the L-dimensional mixture WY, the univariate estimates fail to provide relevant results. In contrast, mixing (change of coordinates) does not affect the multivariate procedure. For the latter, the analysis of the $M = 4$ largest eigenvalues yields estimates that are very similar to those for unmixed ofBm $P^{-1}Y$ (see [20] for similar findings; note that $L - M = 16$ eigenvalues $\lambda_m(2^j)$ are not visible in the plot because they equal zero). Finally, when noise is added to the mixture WY, the univariate estimates $S(2^j)$ do not allow to conclude on the composition of the observed data Z (i.e., mixture of M multivariate self-similar components + noise) and are essentially identical to the noise-free mixture case WY. In stark contradistinction, the presence of noise appears in the form of distinct eigenvalues for the multivariate estimates $\Lambda(2^j)$. The behavior of the latter is visually quite different (smaller values that are consistent across noise components) from that of the $M = 4$ largest eigenvalues, which can be identified with the hidden Hurst eigenvalue-driven scaling components of the ofBm. This enables us to unveil the existence of $M = 4$ components in the mixture. What is more, the estimation of the corresponding exponents H_m is not affected by the presence of noise.

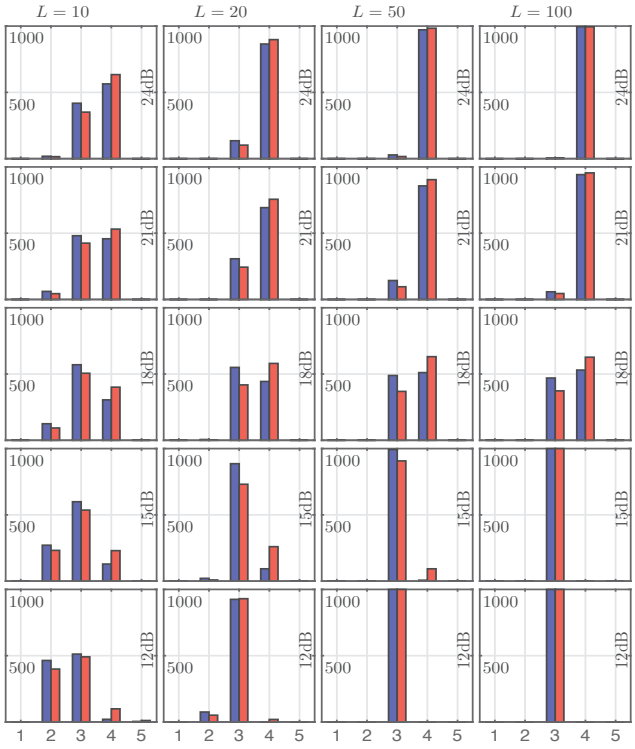

Fig. 2. **Histograms of selected model orders $\hat{M}$** for different SNR (top to bottom) and number of components L (left to right). The blue and red bars correspond to multiscale model selection strategies MV and RS, respectively (cf., Section 3.2). The ground truth is $M = 4$.

4.2. Model order selection performance

Fig. 2 plots the histograms of the model orders that are selected for the noisy mixtures of ofBm using the MV and RS procedures and (8) (blue and red bars), for various SNR values (increasing from bottom to top) and mixture dimensions L (from left to right), respectively. The results lead to the following conclusions. First, the model order selection procedure is overall satisfactorily accurate as long as the SNR is sufficiently large. For instance, a large majority of decisions are correct for 21dB SNR and $L \geq 20$. Second, except for very severe SNR, the procedure does not detect more components than there actually are in the mixture, regardless of the SNR value and L. As a result, the decisions are slightly conservative (i.e., $\hat{M} \leq M$) on average. This can be heuristically interpreted as follows. For small SNR, the ofBm components with smallest H_m are confounded with the noise, hence leading to underestimation of M. By contrast, for large SNR, the gap between noise and ofBm components becomes large, leading to two distinct groups of eigenvalues and no ambiguity in the composition of the mixture.

Further, the MV and RS strategies for combining information from different scales lead to comparable results, with slightly better performance for RS; for instance, for 18dB SNR and mixture dimension $L = 50$, RS yields 58% correct decisions, while MV underestimates the number of components in more than 45 out of 100 cases. Finally, it is interesting to note that the performance of the proposed approach increases for large mixture dimension L ("blessing of dimensionality"). Indeed, correct decisions are given with higher probability for large L. Moreover, it is observed that for small L, the estimates $\hat{M}$ are spread out between different values, while for large L, the procedure always tends to prefer one single value $\hat{M}$ for a given SNR value.

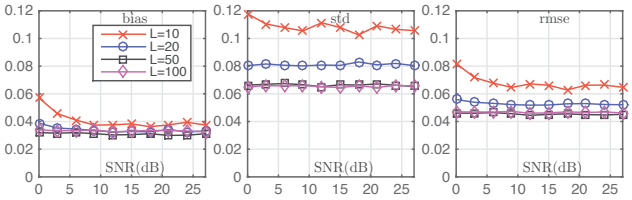

Fig. 3. **Estimation performance for ofBm parameters** from $M = 4$ largest eigenvalues $\Lambda(2^j)$ for different L as a function of SNR: bias, standard deviations, and rmse (from left to right).

4.3. Estimation performance

Fig. 3 reports results on the average estimation performance, evaluated over 100 independent realizations, for the exponents H_m corresponding to the $M = 4$ largest eigenvalues $\Lambda(2^j)$ of Z, as a function of SNR and mixture dimension L: bias (left), standard deviations (std, center) and root-mean squared errors (rmse, right); results are given as the square root of the averaged (over the $M = 4$ components) squares of the quantities and lead to the following complementary conclusions. First, estimates $\hat{H}_m$ are also more accurate for large L, which mirrors the results of Section 4.2: the standard deviations of estimates for $L = 10$ are up to twice as large as those yielded for $L = 100$. This has never been reported before and indicates that the multivariate estimation procedure benefits from extra robustness when the ofBm components are embedded in a high-dimensional ($L \gg M$) mixture (cf. [24] for a related analysis in a different context). Second, for a small SNR and L, estimation accuracy is limited by biased estimates (because ofBm components are drowned in noise). Finally, for large SNR, estimation variances become independent of the noise level because they converge to the variance of $\hat{H}$ for the noise-free case (essentially controlled by the effective sample size, i.e., N, j_1 and j_2).

5. CONCLUSIONS

In this work, we propose a method for estimating the number of multivariate self-similar components (sources) M in $L \geq M$ noisy time series. The method relies on the ofBm model for multivariate self-similarity, on a multivariate estimation method for the Hurst eigenvalues $\underline{H}$ of ofBm, and on the use of classical information theoretic criteria applied to the multiscale eigenstructure of the wavelet spectra of the noisy observations. To the best of our knowledge, this work reports for the first time *i)* that the estimation of $\underline{H}$ when $L > M$ based on multivariate procedures has satisfactory performance, and *ii)* an operational procedure for the estimation of M in multivariate self-similarity with considerable accuracy, including in large dimensional situations. In the model selection procedure, wavelet spectra are combined across scales using averages or majority votes; alternative strategies will be studied in future work; similarly, different model selection criteria could be employed, e.g., using more accurate models for the eigenvalue distributions - classical AIC was used to provide a proof of concept. In future work, the proposed tools will be used in the modeling of multivariate scale-free dynamics in data from macroscopic brain activity (M/EEG, fMRI).

6. REFERENCES

[1] P. Ciuciu, G. Varoquaux, P. Abry, S. Sadaghiani, and A. Kleinschmidt, "Scale-free and multifractal time dynamics of fMRI signals during rest and task," *Frontiers in physiology*, vol. 3, 2012.

[2] N. Zilber, P. Ciuciu, P. Abry, and V. Van Wassenhove, "Modulation of scale-free properties of brain activity in MEG," in *IEEE International Symposium on Biomedical Imaging (ISBI)*. IEEE, 2012, pp. 1531–1534.

[3] F. A. Isotta, C. Frei, V. Weilguni, M. Perčec Tadić, P. Lassegues, B. Rudolf, V. Pavan, C. Cacciamani, G. Antolini, S. M. Ratto, and M. Munari, "The climate of daily precipitation in the alps: development and analysis of a high-resolution grid dataset from pan-alpine rain-gauge data," *International Journal of Climatology*, vol. 34, no. 5, pp. 1657–1675, 2014.

[4] D. Veitch and P. Abry, "A wavelet-based joint estimator of the parameters of long-range dependence," *IEEE Transactions on Information Theory*, vol. 45, no. 3, pp. 878–897, 1999.

[5] H. Wendt, P. Abry, and S. Jaffard, "Bootstrap for empirical multifractal analysis," *IEEE Signal Processing Magazine*, vol. 24, no. 4, pp. 38–48, 2007.

[6] G. Samorodnitsky and M. Taqqu, *Stable non-Gaussian random processes*, Chapman and Hall, New York, 1994.

[7] B. B. Mandelbrot and J. W. van Ness, "Fractional Brownian motion, fractional noises and applications," *SIAM Reviews*, vol. 10, pp. 422–437, 1968.

[8] M. Maejima and J. D. Mason, "Operator-self-similar stable processes," *Stochastic Processes and their Applications*, vol. 54, no. 1, pp. 139–163, 1994.

[9] 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.

[10] G. Didier and V. Pipiras, "Integral representations and properties of operator fractional Brownian motions," *Bernoulli*, vol. 17, no. 1, pp. 1–33, 2011.

[11] G. Didier and V. Pipiras, "Exponents, symmetry groups and classification of operator fractional Brownian motions," *Journal of Theoretical Probability*, vol. 25, pp. 353–395, 2012.

[12] P. Abry and G. Didier, "Wavelet estimation for operator fractional Brownian motion," *Bernoulli*, vol. 24, no. 2, pp. 895–928, 2018.

[13] P. Comon and C. Jutten, *Handbook of Blind Source Separation: Independent Component Analysis and Applications*, Academic Press, 2010.

[14] T. Söderström and P. Stoica, *System identification*, Prentice-Hall, Inc., 1988.

[15] P. Stoica and Y. Selen, "Model-order selection: a review of information criterion rules," *IEEE Signal Processing Magazine*, vol. 21, no. 4, pp. 36–47, 2004.

[16] A. P. Liavas and P. A. Regalia, "On the behavior of information theoretic criteria for model order selection," *IEEE Transactions on Signal Processing*, vol. 49, no. 8, pp. 1689–1695, 2001.

[17] J. P. C. L. Da Costa, A. Thakre, F. Roemer, and M. Haardt, "Comparison of model order selection techniques for high-resolution parameter estimation algorithms," in *Proc. 54th International Scientific Colloquium (IWK'09), Ilmenau, Germany*, 2009.

[18] R. A. Johnson and D. W. Wichern, *Applied multivariate statistical analysis*, vol. 4, Prentice-Hall New Jersey, 2014.

[19] P. Abry and G. Didier, "Wavelet eigenvalue regression for n-variate operator fractional Brownian motion," *arXiv preprint arXiv:1708.03359*, 2017.

[20] G. Didier, H. Helgason, and P. Abry, "Demixing multivariate-operator selfsimilar processes," in *IEEE International Conference on Acoustics, Speech and Signal, Processing (ICASSP), Brisbane, Australia*, 2015, pp. 1–5.

[21] P. Abry, G. Didier, and H. Li, "Two-step wavelet-based estimation for Gaussian mixed fractional processes," *arXiv preprint 1607.05167*, 2018.

[22] 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.

[23] S. Mallat, *A Wavelet Tour of Signal Processing*, Academic Press, San Diego, CA, 1998.

[24] C. Lam and Q. Yao, "Factor modeling for high-dimensional time series: inference for the number of factors," *The Annals of Statistics*, vol. 40, no. 2, pp. 694–726, 2012.