COUNTING THE NUMBER OF DIFFERENT SCALING EXPONENTS IN MULTIVARIATE SCALE-FREE DYNAMICS: CLUSTERING BY BOOTSTRAP IN THE WAVELET DOMAIN

Charles-Gérard Lucas[1] *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, Univ. Toulouse, CNRS, Toulouse, France.
[3] Math. Dept., Tulane University, New Orleans, USA.

ABSTRACT

Multivariate selfsimilarity has become a classical tool to analyze collections of time series recorded jointly on one same system. Often, it amounts to estimating as many scaling exponents as time series. However, this leaves open the important question how many such scaling exponents are actually different. Elaborating on earlier work aiming to test the hypothesis that all exponents are equal, we intend here to count the number of different scaling exponents from a single finite size multivariate time series. To this end, we devise an original clustering procedure that combines a wavelet domain block multivariate bootstrap scheme with a test strategy for a reduced set of multiple hypotheses on the pairwise equality of scaling exponents that are relevant to clustering. Monte Carlo simulations, making use of synthetic multivariate selfsimilar processes, assess the relevance and performance of the proposed procedure under different scenarios and demonstrate that the proposed method yields practically satisfactory cluster number and size estimations.

***Index Terms*—** Multivariate selfsimilarity, multivariate scaling exponents, multivariate wavelet transform, bootstrap.

1. INTRODUCTION

Context. Modern real-world applications often entail the joint analysis of multivariate times series, collected from one same system by a (possibly large) number of sensors. Scale-free dynamics has proven a fruitful paradigm to monitor the status of the system for many applications very different in nature (cf. e.g., [1]). Multivariate scale-free analysis amounts to estimating the scaling exponents of the data, and it is common practice to estimate as many exponents as there are components in the collected data [2,3]. However, to correctly interpret the status of a system, knowledge of how many of the estimated scaling exponents actually correspond to different values, and estimating how many components are controlled by each of these distinct scaling exponent, is crucial. This is e.g., the case in neurosciences with the analysis of infraslow brain activity (cf. [4]). These two issues have been rarely addressed in the context of multivariate scale-free dynamics and form the core of this work.

Related works. Selfsimilarity and fractional Brownian motion have long been recognized as relevant models for the scale-free temporal dynamics of single time series [5–7]. In practice, the estimation of the selfsimilarity exponent is thus central and can be efficiently and robustly performed by means of wavelet transforms [8, 9]. The multivariate time series encountered in many modern applications call both for multivariate selfsimilarity models, such as the recently proposed operator fractional Brownian motion (ofBm) [10–13], and for multivariate wavelet representation based estimation procedures, such as those developed in [2, 3, 14, 15]. These estimation procedures output as many selfsimilarity parameter estimates as there are time series in the data. The actual and relevant use of such collections of estimates for understanding the system under study requires, as a first step, determining how many of the estimates are actually different. In earlier works [15, 16], we proposed a wavelet-domain block-bootstrap strategy for deciding whether all selfsimilarity exponents are equal or not. Yet, this leaves untouched the critical issue of counting the number of equal selfsimilarity exponents.

Goals, contributions and outline. The present work aims to devise a clustering procedure that, from a single observation of finite size multivariate data, is able to count the number of actually distinct selfsimilarity exponents. To that end, the definition of multivariate selfsimilarity is briefly recalled in Section 2 together with the eigen-wavelet-based selfsimilarity parameter estimation procedure. The proposed clustering procedure, the key contribution of the present work, is then detailed in Section 3: Building up on [15, 16], the clustering strategy consists in an original multiple hypothesis test procedure for the equality of the reduced set of pairs of selfsimilarity exponents which, after ordering from small to large values, are neighbors. By construction, the so-obtained test decisions translate directly into clustering results. To estimate the null distributions of the test statistics, we also propose an original wavelet domain time-scale-component-block bootstrap scheme. Monte Carlo experiments using ofBm, reported in Section 4, are used to assess the relevance and performance of the proposed clustering procedure. Results demonstrate that the proposed bootstrap procedure accurately approximates the null distributions, and that our procedure yields satisfactory clustering performance under several scenarios, with different sample sizes or number of components with identical selfsimilarity exponents, both in terms of number of exponents that are actually different, and in terms of the number of components each of them control. OfBm synthesis, selfsimilarity exponent estimation and test procedures are implemented by the authors and will be made available upon publication.

2. MULTIVARIATE SELFSIMILARITY

2.1. Model: Operator fractional Brownian motion

Fractional Brownian motion (fBm) is the only Gaussian, self-similar stochastic process with stationary increments. It is classically used as a paradigm for (univariate) selfsimilarity [5–7]. In the present work, we use a collection, $X \triangleq \{X_{H_1}(t), \ldots, X_{H_M}(t)\}_{t\in\mathbb{R}}$, of M fBm, each with possibly different selfsimilarity exponents $\underline{H} = (H_1, \ldots, H_M)$, $0 < H_1 \leq \ldots \leq H_M < 1$, made dependent via a $M \times M$ covariance matrix Σ_X. To model M-variate selfsimilarity, these components are linearly mixed via a $M \times M$ real-valued and

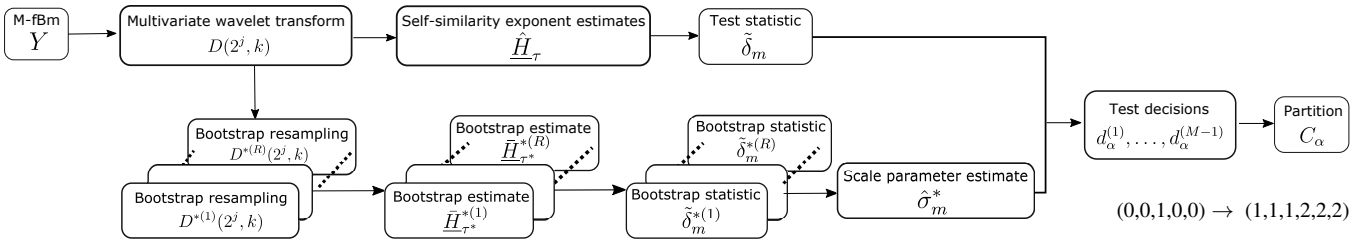


Fig. 1: **Flowchart of the clustering strategy.**

invertible mixing matrix W, referred to as M-fBm:

$$Y \triangleq \{Y_1^{\underline{H},\Sigma_X,W}(t),\ldots,Y_M^{\underline{H},\Sigma_X,W}(t)\}_{t\in\mathbb{R}} \triangleq W\,\{X_{H_1}(t),\ldots,X_{H_M}(t)\}_{t\in\mathbb{R}} = WX. \quad (1)$$

M-fBm is a specific case of the more general and inspiring *operator fractional Brownian motion*, constructed in [10–13], which indicates that $\underline{H}$ and Σ_X are dependent parameters that cannot be chosen independently [12].

2.2. Estimation: Wavelet analysis

The key step in selfsimilarity analysis lies in the estimation of the $\underline{H} = (H_1,\ldots,H_M)$ from the M-variate time series Y. An estimation procedure based on the multivariate discrete wavelet transform (DWT) was proposed in [2,3]. Let $D_{Y_m}(2^j,k) = \langle 2^{-j/2}\psi_0(2^{-j}t-k)|Y_m(t)\rangle$, $\forall k\in\mathbb{Z}$, $\forall j\in\{j_1,\ldots,j_2\}$ denote the univariate DWT coefficients, with ψ_0 the mother wavelet [17, 18]. Multivariate DWT is defined by stacking of the univariate DWT, $D_Y(2^j,k) = (D_{Y_1}(2^j,k),\ldots,D_{Y_M}(2^j,k))$, $\forall k\in\mathbb{Z}$, $\forall j\in\{j_1,\ldots,j_2\}$, with $\forall m\in\{1,\ldots,M\}$. The behaviors along scales 2^j of the eigenvalues of the $M\times M$ multivariate wavelet spectrum

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

computed at scale 2^j, provide relevant estimation of $\underline{H}$.

Yet, this approach was shown to suffer from the so-called *repulsion effect* in computing equal eigenvalues, thus inducing biased estimates for $\underline{H}$ [15]. To limit the resulting bias, it was proposed to compute modified wavelet spectra with equal number of wavelet coefficients contributing at each scale:

$$S^{(w)}(2^j) \triangleq \frac{1}{n_{j_2}}\sum_{k=1+(w-1)n_{j_2}}^{wn_{j_2}} D_Y(2^j,k)D_Y(2^j,k)^*, \quad (3)$$

for $w=1,\ldots,2^{j-j_2}$. The eigenvalues $\{\lambda_1^{(w)}(2^j),\ldots,\lambda_M^{(w)}(2^j)\}$ are computed independently at each scale 2^j and for each non-overlapping window w, hence affected by a repulsion effect of same intensity across all scales. Then, estimators $\hat{H}_1,\ldots,\hat{H}_M$ are defined as linear regressions, against scales $2^{j_1},\ldots,2^{j_2}$, of the logarithms $\bar{\lambda}_m(2^j) \triangleq 2^{j_2-j}\sum_{w=1}^{2^{j-j_2}}\log_2(\lambda_m^{(w)}(2^j))$ averaged across windows w:

$$\hat{H}_m = \left(\sum_{j=j_1}^{j_2} v_j\bar{\lambda}_m(2^j)\right)/2 - \frac{1}{2}, \quad \forall m=1,\ldots,M, \quad (4)$$

with v_j classical regression weights verifying $\sum_j jv_j = 1$ and $\sum_j v_j = 0$ (cf. [9]). The relevance of this estimation procedure applied to M-fBm was assessed in [15].

3. MULTIVARIATE TIME SCALE BLOCK-BOOTSTRAP BASED CLUSTERING

3.1. Testing multiple ordered pairwise hypotheses

For notational simplicity, assume that the vector $\underline{H}$ is sorted, $\forall m = 1,\ldots,M-1, H_{m+1}\geq H_m$. Our strategy for estimating the groups of equal exponents among $\underline{H}$ consists in testing the equality of the $M-1$ pairs of components that are neighbors in $\underline{H}$, with correction for multiple test decisions. The positions of the detected non-equalities among these $M-1$ pairs of ordered exponents, if any, then defines the boundaries for the clusters for the distinct exponents.

The null hypotheses for these $M-1$ pairwise tests are thus defined as

$$\mathcal{H}_0^{(m)} : H_{m+1} = H_m, \quad m=1,\ldots,M-1. \quad (5)$$

To construct a test for each hypothesis, we first compute, from a single observation of finite size M-variate data, the vector of M estimates $\underline{\hat{H}}$ as described in Section 2.2, and sort its entries in ascending order, $\underline{\hat{H}}_\tau = (\hat{H}_{\tau(1)},\ldots,\hat{H}_{\tau(M)})$ with $\hat{H}_{\tau(m+1)}\geq\hat{H}_{\tau(m)}$, $m=1,\ldots,M-1$. We then compute the $M-1$ test statistics defined as

$$\tilde{\delta}_m = \hat{H}_{\tau(m+1)} - \hat{H}_{\tau(m)}. \quad (6)$$

It can been shown that the entries of $\underline{\hat{H}}$ are asymptotically Gaussian and weakly dependent, which naturally leads to postulate that under $\mathcal{H}_0^{(m)}$, $\tilde{\delta}_m$ follows a half-normal distribution, $\{\sqrt{2}/\sigma_m\sqrt{\pi}\}\exp(-\delta_m^2/2\sigma_m^2)$. Our numerical simulations validate that this is a reasonable approximation, see Section 4.2 below. The tests for (5) can thus be computed as

$$\text{reject } \mathcal{H}_0^{(m)} \text{ if } \tilde{\delta}_m > \gamma_m, \quad m=1,\ldots,M-1, \quad (7)$$

where each γ_m is a rejection threshold. However, since the scale parameters σ_m are unknown and a priori depend on $\underline{H}$, Σ_X and W, the values for γ_m are also unknown. In order to perform the tests, we therefore propose to estimate σ_m using a bootstrap procedure [19, 20] that is described next.

3.2. Multivariate wavelet domain bootstrap for $\tilde{\delta}_m$

Following schemes proposed in earlier works [15, 16], collections of bootstrap resamples of the multivariate wavelet coefficients $D(2^j,k)$, $k=1,\ldots,n_j$ are computed by a multivariate wavelet-domain block-bootstrap procedure that preserves their scale-time multivariate (cross-)dependence structure (while component-wise resampling would, to the contrary, destroy the cross-component dependencies [16, 19]).

Technically, at each scale 2^j, R block bootstrap resamples $D_j^{*(r)} = (D^{*(r)}(2^j,1),\ldots,D^{*(r)}(2^j,n_j))$, $r=1,\ldots,R$ are drawn with replacement from $(D(2^j,k),\ldots,D(2^j,k+L_B-1))$,

$k = 1, \ldots, n_j$. The overlapping blocks have size L_B in time, and range across all scales and all components jointly. From each bootstrap sample $D_j^{*(r)}$, bootstrap estimates $S^{*(r,w)}(2^j)$ and $\hat{H}_m^{*(r)}$ are computed using Eqs. (3-4).

In order to obtain an approximation of the null distribution of $\tilde{\delta}_m$, under any possible hypothesis (that is, for $\mathcal{H}_0^{(m)}$ true or not true), the following further steps are required. First, the average value is subtracted from each of the bootstrap resample components, $\bar{H}_m^{*(r)} = \hat{H}_m^{*(r)} - \langle \hat{H}_m^* \rangle,\ m = 1, \ldots, M,\ r = 1, \ldots, R$. Second, the components $\bar{H}_m^{*(r)}$ are ordered for each bootstrap resample individually, $\underline{\bar{H}}_{\tau^*}^{*(r)} = (\bar{H}_{\tau^*(r,1)}^{*(r)}, \ldots, \bar{H}_{\tau^*(r,M)}^{*(r)})$, with $\bar{H}_{\tau^*(r,m+1)}^{*(r)} \geq \bar{H}_{\tau^*(r,m)}^{*(r)},\ m = 1, \ldots, M-1,\ r = 1, \ldots, R$. Then, the bootstrap test statistics $\tilde{\delta}_m^{*(r)} = \bar{H}_{\tau^*(r,m+1)}^{*(r)} - \bar{H}_{\tau^*(r,m)}^{*(r)},\ r = 1, \ldots, R$, are computed and their bootstrap variance Var^* can be used to obtain estimates for the scale parameters σ_m under the null hypotheses

$$\hat{\sigma}_m^{*2} = \mathrm{Var}^*(\tilde{\delta}_m^*) / \left(1 - \frac{2}{\pi}\right). \tag{8}$$

Finally, the bootstrap estimated p-values associated with the rejection of $\mathcal{H}_0^{(m)}$, $m = 1, \ldots, M-1$, are computed as

$$p_m^* = 1 - F\left(\frac{\tilde{\delta}_m}{\hat{\sigma}_m^*}\right), \tag{9}$$

where F is the cumulative density function of the standardized half-normal distribution.

3.3. Multiple hypotheses test decisions and clustering

The bootstrap p-values (9) can now be used to perform the tests (7). Specifically, to correct for multiple hypotheses, a decision $d^{(m)}$ to reject or not independently each of the null hypothesis $\mathcal{H}_0^{(m)}$ at a preset level of false discovery rate (FDR) α is taken using the Benjamini-Hochberg correction procedure [21],

$$d_\alpha^{(m)} = 1 : \quad p_m^* < \frac{\alpha}{M-1}\pi(m), \tag{10}$$

where $\pi(\cdot)$ denotes the permutation that orders the p-values, $p_{\pi(1)} < \ldots < p_{\pi(M-1)}$.

Finally, the vector of $M-1$ decisions is an indicator vector for the cluster boundaries for $\underline{H}$: if there are P rejected hypotheses $\mathcal{H}_0^{(m)}$, there are $P+1$ distinct selfsimilarity exponents and clusters. The detected cluster labels for H_m are defined as

$$C_\alpha(m) = \sum_{m'=1}^{m} D_\alpha(m') \tag{11}$$

where $D_\alpha = (1, d_\alpha^{(1)}, \ldots, d_\alpha^{(M-1)})$.

The whole clustering strategy described in Section 3 is summarized in Fig. 1.

4. PERFORMANCE ASSESSMENT

4.1. Numerical experiments set-up

Monte Carlo experiments. To assess the relevance of the proposed multivariate wavelet domain block bootstrap-assisted clustering procedure defined in Section 3, and to quantify its performance, Monte Carlo simulations are conducted, making use of $N_{MC} = 1000$ independent copies of synthetic $M = 6$-variate M-fBm of sample size $N = 2^{16}$. Four scenarios are considered:
— Scenario1 corresponds to $H_1 = \ldots = H_M = 0.8$ (1 cluster).
— Scenario2 consists of 2 clusters of similarity exponents of size 3 with values 0.6 and 0.8, such that $\mathcal{H}_0^{(3)}$ is not true.
— Scenario3 consists of 3 clusters of equal size 2 with values 0.4, 0.6 and 0.8, such that $\mathcal{H}_0^{(2)}$ and $\mathcal{H}_0^{(4)}$ are not true.
— Scenario4 consists of 3 clusters of different sizes 1, 3 and 2 with values 0.4, 0.6 and 0.8, such that $\mathcal{H}_0^{(1)}$ and $\mathcal{H}_0^{(4)}$ are not true.

The covariance matrix Σ_X is chosen such that all diagonal entries are set to 1 and all non-diagonal entries are set to $r = 0.5$. The $M \times M$ invertible matrix W is randomly selected and kept fixed for all experiments. Wavelet analysis is performed with the least asymmetric Daubechies3 wavelet across analysis scales $2^{j_1} = 2^8$ to $2^{j_2} = 2^{11}$. $R = 500$ block-bootstrap resamples are drawn from overlapping blocks of size $L_B = 6$ (corresponding to the size of the Daubechies3 mother-wavelet time support).

Clustering performance assessment. The adjusted random index (ARI) and Normalized Mutual Information (NMI) [22] are used to quantify clustering performance, i.e., how well components with same H are grouped together and components with different H are separated: ARI measures the number of pairs of elements that are correctly grouped or separated, whereas NMI measures the joint entropy of the distributions of estimated and correct clustering.

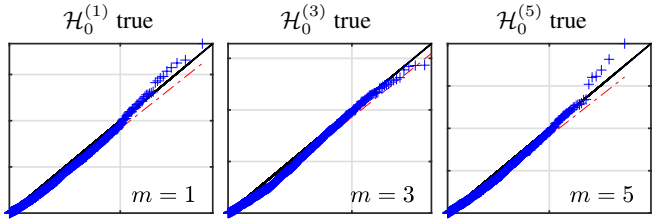


Fig. 2: **Distributions of the test statistics $\tilde{\delta}_m$.** Standardized Monte Carlo statistics $\tilde{\delta}_m/\sigma_m$ against a standardized half-normal distribution under $H_1 = \ldots = H_6 = 0.8$ (Scenario1).

4.2. Empirical distributions of $\tilde{\delta}_m$ and $\tilde{\delta}_m^*$

Distribution of $\tilde{\delta}_m$ under $\mathcal{H}_0^{(m)}$. Fig. 2 plots quantile-quantile plots of the standardized half-normal distribution against the empirical (Monte Carlo) distributions of the normalized test statistics $\tilde{\delta}_m/\hat{\sigma}_m$ under $\mathcal{H}_0^{(m)}$. Here, $\hat{\sigma}_m$ is the Monte Carlo estimate of the standard deviation of $\tilde{\delta}_m$. It demonstrates that under the null hypothesis, the distribution of the test statistic $\tilde{\delta}$ can indeed be very well approximated by a half-normal distribution.

Distributions of $\tilde{\delta}_m^*$ under $\mathcal{H}_0^{(m)}$ and $\mathcal{H}_1^{(m)}$. Fig. 3 depicts quantile-quantile plots of the distribution of $\tilde{\delta}_m^*/\hat{\sigma}_m^*$ for several m against a standardized half-normal distribution for Scenario2: in this case, the null hypothesis $\mathcal{H}_0^{(m)}$ is true for $m = 1$ and $m = 5$, but it is not true for $m = 3$. The results clearly indicate that in all cases ($\mathcal{H}_0^{(m)}$ true and $\mathcal{H}_0^{(m)}$ not true), the bootstrap distributions are in good agreement with the half-normal distribution.

Bootstrap estimation of the scale parameter σ_m. Finally, Table 1 reports the Monte Carlo estimates $\hat{\sigma}_m$ for the scale parameter σ_m and the Monte Carlo averages and standard deviations for the corresponding bootstrap scale parameter estimates $\hat{\sigma}_m^*$ under Scenario1 ($\mathcal{H}_0^{(m)}$ is true for any m). The results confirm that the bootstrap estimates $\hat{\sigma}_m^*$ are in excellent agreement with $\hat{\sigma}_m$. Together with the above observations on the shape of the respective distri-

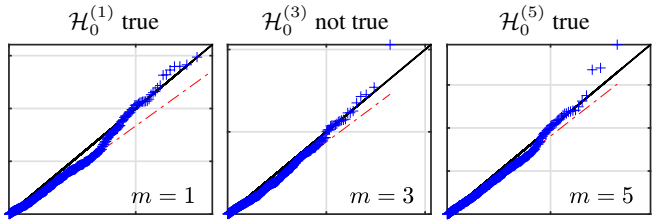


Fig. 3: **Distributions of bootstrap statistics $\tilde{\delta}^*_m$.** Quantile-quantile-plots of bootstrap standardized $\tilde{\delta}^*_m / \hat{\sigma}^*_m$ for an arbitrarily chosen realization against a standardized half-normal distribution under Scenario2 ($\mathcal{H}_0^{(m)}$ is not true for $m = 3$ and is true for $m \neq 3$, and $M = 6$); black solid lines correspond with identical quantiles, dashed red lines join the first and third quantiles of the distributions.

Table 1: Monte Carlo estimates of scale parameters σ_m and bootstrap scale parameter estimates $\hat{\sigma}^*_m$ (Monte Carlo average and standard deviation) for $H_1 = \ldots = H_M = 0.8$ (Scenario1).

	$m=1$	$m=2$	$m=3$	$m=4$	$m=5$
$\hat{\sigma}_m \times 10^2$	1.65	1.16	1.01	1.06	1.49
$\hat{\sigma}^*_m \times 10^2$	1.65 $\pm$0.13	1.10 $\pm$0.07	0.99 $\pm$0.06	1.07 $\pm$0.06	1.51 $\pm$0.09

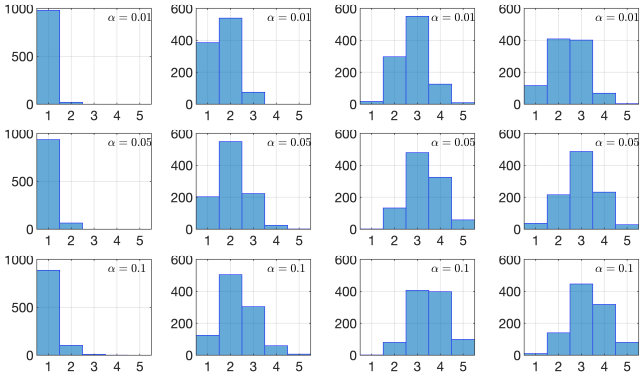
butions, this numerically validates the assumptions underlying the proposed bootstrap based test procedure.

4.3. Performance of the clustering strategy

We finally assess the clustering performance for the four scenarios defined above. Fig. 4 reports the histograms of the estimated number of clusters for the different scenarios and three different levels of false discovery rate $\alpha = (0.01,\ 0.05,\ 0.10)$. For Scenario1, for which all $M - 1 = 5$ null hypotheses are true, we obtain actual false discovery rates of $(0.02,\ 0.06,\ 0.11)$, respectively, which is in good agreement with the preset values. This further corroborates the above numerical analysis on the relevance of our test and detection procedure for null hypotheses. For Scenario2, Scenario3 and Scenario4, for which more than one cluster are to be detected, our proposed procedure detects the correct number of clusters in a majority of the cases. This demonstrates, first, that our proposed test has practically reasonable power (aka detection probability), and second, that our clustering strategy leads to satisfactory results concerning the number of detected groups of equal H_m. As expected, the rate of cluster number over-estimation (under-estimation) increases (decreases) with increasing preset false discovery rate α, respectively.

Table 2 further quantifies this clustering performance analysis and shows average values for NMI and ARI with 95% confidence interval (obtained as averages across realizations) for all four scenarios and false discovery rate $\alpha = 0.05$. The results indicate that the proposed procedure overall leads to practically satisfactory performance, with NMI values of up to 0.87, and ARI values of up to 0.94. It is interesting to note that Scenario4 is more difficult to cluster than Scenario3, which contains the same number of clusters and values for H_m but equi-sized clusters.

Overall, these results unambiguously validate the relevance practically satisfactory performance of the proposed procedure.



(a) Scenario1 (b) Scenario2 (c) Scenario3 (d) Scenario4

Fig. 4: **Estimated number of clusters.** Histograms of the estimated numbers of clusters for (a) 1 cluster, (b) 2 clusters and (c-d) 3 clusters with $M = 6$.

Table 2: Clustering performance with 95% confidence interval for significance level $\alpha = 0.05$.

	Scenario1	Scenario2	Scenario3	Scenario4
NMI	n/a	0.66 ± 0.02	0.87 ± 0.01	0.79 ± 0.01
ARI	0.94 ± 0.02	0.60 ± 0.03	0.68 ± 0.02	0.59 ± 0.02

5. CONCLUSION AND PERSPECTIVES

The present work constitutes a first attempt to count the number of scaling exponents, estimated from multivariate data, that are actually different. It relies on combining a wavelet eigenvalues-based estimation procedure, with a multivariate wavelet domain block bootstrap scheme, and false discovery rate corrections for multiple hypotheses testing. Monte Carlo experiments show that the bootstrap procedure satisfactorily reproduces the real distributions of the test statistics and that the overall procedure yields very satisfactory performance in estimating the actual number of scaling exponents and in grouping identical ones together, under several scenarios with one, two or three clusters, with clusters of possibly unbalanced size.

Matlab routines permitting the estimation of the multivariate selfsimilarity exponents as well as their clustering are publicly available[1] as part of our reproducible research effort.

Future investigations will include exploring whether the use of $M(M-1)/2$ non-sorted hypotheses further improves performance compared to using $M - 1$ sorted pair hypotheses. Large dimension issues (i.e., situations where the number of components grows at fixed rate with the sample size) will also be investigated.

6. ACKNOWLEDGMENT

Supported by PhD Grant DGA/AID (no 01D20019023), ANR-16-CE33-0020 MultiFracs, ANR-18-CE45-0007 MUTATION. G. Didier's long term visits to ENS Lyon were sponsored by ENS Lyon, the CNRS and the Simons Foundation collaboration grant #714014.

[1] https://github.com/charlesglucas/ofbm_tools

7. REFERENCES

[1] P. Abry, S. Wendt, H.and Jaffard, and G. Didier, "Multivariate scale-free temporal dynamics: From spectral (fourier) to fractal (wavelet) analysis," *Comptes Rendus Physique*, vol. 20, no. 5, pp. 489–501, 2019.

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

[3] P. Abry and G. Didier, "Wavelet eigenvalue regression for n-variate operator fractional Brownian motion," *Journal of Multivariate Analysis*, vol. 168, pp. 75–104, November 2018.

[4] Daria La Rocca, Herwig Wendt, Virginie van Wassenhove, Philippe Ciuciu, and Patrice Abry, "Revisiting functional connectivity for infraslow scale-free brain dynamics using complex wavelets," *Frontiers in Physiology*, vol. 11, pp. 1651, 2021.

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

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

[7] V. Pipiras and M. S. Taqqu, *Long-Range Dependence and Self-Similarity*, vol. 45, Cambridge University Press, 2017.

[8] P. Flandrin, "Wavelet analysis and synthesis of fractional Brownian motion," *IEEE Trans. Info. Theory*, vol. 38, no. 2, pp. 910–917, 1992.

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

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

[11] 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.

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

[13] 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.

[14] H. Wendt, P. Abry, and G. Didier, "Bootstrap-based bias reduction for the estimation of the self-similarity exponents of multivariate time series," in *IEEE Int. Conf. Acoust., Speech, and Signal Proces. (ICASSP)*, Brighton, UK, May 2019.

[15] C.-G. Lucas, P. Abry, H. Wendt, and G. Didier, "Bootstrap for testing the equality of selfsimilarity exponents across multivariate time series," in *Proc. European Signal Processing Conference (EUSIPCO)*, Dublin, Ireland, August 2021.

[16] H. Wendt, P. Abry, and G. Didier, "Wavelet domain bootstrap for testing the equality of bivariate self-similarity exponents," in *Proc. IEEE Workshop Statistical Signal Proces. (SSP)*, Freiburg, Germany, June 2018.

[17] I. Daubechies, *Ten Lectures on Wavelets*, vol. 61, SIAM, 1992.

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

[19] S. N. Lahiri, *Resampling Methods for Dependent Data*, Springer, New York, 2003.

[20] A. M. Zoubir and D. R. Iskander, *Bootstrap techniques for signal processing*, Cambridge University Press, 2004.

[21] 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.

[22] N. X. Vinh, J. Epps, and J. Bailey, "Information theoretic measures for clusterings comparison: Variants, properties, normalization and correction for chance," *The Journal of Machine Learning Research*, vol. 11, pp. 2837–2854, 2010.