Multiscale approximate eigenvectors for multivariate self-similarity estimation

Herwig Wendt
CNRS, IRIT, University of Toulouse
Toulouse, France, herwig.wendt@irit.fr

Gustavo Didier
Mathematics Department, Tulane University
New Orleans, LA, USA, gdidier@tulane.edu

Marcus Carlsson, Erik Troedsson
Centre for Mathematical Sciences, Lund University
Lund, Sweden, firstname.lastname@math.lund.se

Patrice Abry
CNRS, ENS de Lyon, Laboratoire de Physique
Lyon, France, patrice.abry@ens-lyon.fr

***Abstract*—Many real-world systems generate multivariate time series data, often exhibiting self-similarity and scale-invariance across different modalities. The estimation of Hurst exponents in such settings is crucial for analyzing long-range dependencies. Yet, traditional eigenanalysis-based methods suffer from scale-dependent distortions, particularly in the presence of scaling amplitude discrepancies. In this work, we propose a novel multiscale eigenanalysis approach that leverages joint diagonalization of wavelet random matrices to improve estimation accuracy. By approximating a common eigenvector basis across multiple scales, our method mitigates the limitations of scale-wise eigenvalue regressions and provides robust estimation of multivariate self-similarity parameters. We demonstrate the effectiveness of our approach through extensive Monte Carlo simulations, showcasing improved performance over traditional methods in both orthogonal and non-orthogonal mixing scenarios. These findings establish joint eigenvector-based wavelet analysis as a powerful tool for multivariate self-similarity estimation.**

***Index Terms*—multivariate self-similarity, Hurst exponent, random matrices, joint diagonalization, wavelets**

I. Introduction

Context: scale invariance. *Scale invariance* is a fundamental property encountered in a wide range of signals from various fields, including physics and engineering [1], [2]. A signal X is called scale-invariant, or *fractal*, when its temporal dynamics lack a characteristic scale. Unlike traditional statistical mechanics, where modeling typically focuses on fixed scales, scale-invariant signals are described using *scaling exponents*, which relate the behavior of the system across multiple scales. A cornerstone model for scale invariance, based on the functional central limit theorem, is *fractional Brownian motion* (fBm) [3]. FBm is the only Gaussian, stationary-increment, self-similar process. The latter property means that its finite-dimensional distributions (f.d.d.) are scale-invariant, i.e., $\{B_H(t)\}_{t\in\mathbb{R}} \overset{\text{f.d.d.}}{=} \{a^H B_H(t/a)\}_{t\in\mathbb{R}}$, where $a > 0$ and the scaling exponent $0 < H < 1$ is known as the *Hurst* parameter. Estimating H plays a crucial role in various signal processing tasks such as diagnosis, classification, and detection. The *wavelet transform* provides a robust analytical framework for many well-established methods of estimating H [4].

Multivariate self-similarity. While most applications have relied on the univariate fBm model, the increasing prevalence of large-scale data collected from multiple sensors has led to the need for modeling multivariate self-similarity. For instance, in neuroscience, macroscopic brain activity time series can range from hundreds (e.g., MEG data) to tens of thousands (e.g., fMRI data) [2]. In econometrics, detecting scaling laws in multivariate fractional time series is crucial for identifying meaningful long-run relationships, such as *cointegration* [5]. In these complex data sets, the presence of multiple scaling laws – i.e., distinct *Hurst exponents* – implies distinct large-scale behavior along possibly non-canonical coordinate axes. Ignoring the interplay of these distinct scaling laws may lead to arbitrarily large *estimation biases* [6], [7].

Eigenanalysis of wavelet random matrices. *Random matrices* have emerged as a key analytical tool in both mathematical physics [8] and high-dimensional statistics and machine learning [9]. In particular, *wavelet random matrices* (WRMs) offer an asymptotically universal and robust framework for studying multivariate and high-dimensional stochastic dynamics over large scales. For many (Gaussian or non-Gaussian) multivariate fractal systems, the multiple scaling laws hidden in the data eventually emerge as power laws in the eigenspectrum of WRMs, driven by the Hurst exponents. To the best of our knowledge, WRMs were first introduced and studied by the authors and collaborators in a series of papers [10], [11].

Challenge: incongruent scaling amplitudes. Despite their strong asymptotic properties, WRM-based eigenanalysis faces challenges when working with finite samples. A key issue is the potential for large discrepancies in the magnitudes of scaling *amplitudes*. After log-linearization in the eigendomain, this can lead to affine scaling laws that *cross over*, resulting in biased slope (Hurst exponent) estimators. Inspired by blind source separation problems [12], one potential solution for systems with instantaneous correlations is to jointly diagonalize WRMs from two different scales exactly, and estimate the mixing matrix that "scrambles" the multiple power laws [13]. While this method has shown good performance for several

Work partially supported by Swedish Research Council Grant 2022-04917. G.D.s long term visits to ENS de Lyon were supported by the school, the CNRS and the Simons Foundation collaboration grant #714014.

instances of scaling systems, it is limited by relying on only two scales of analysis, which may result in poor Hurst exponent estimates, especially when scaling laws cross over.

Goals and contributions. The main goal of this work is to propose an efficient estimation method for instantaneously correlated, multivariate self-similar systems that addresses the crossover phenomenon. We review the operator fractional Brownian motion (ofBm) model of multivariate self-similarity in Section II, and recap the wavelet eigenvalue-based methodology for estimating the M Hurst exponents of ofBm. We also summarize the asymptotic theory for WRM eigenvectors developed in [10], which has been largely unexplored for inferential purposes. In Section III, we present as our first main contribution a novel estimation procedure for the vector of Hurst exponents based on WRM diagonalization across *multiple* scales, leveraging a recently proposed algorithm [14]. This method introduces a completely new approach within the framework of WRMs. In Section IV, we provide as a second contribution an extensive numerical study of the finite-sample performance of the proposed method, demonstrating its efficiency for realistic sample sizes and significant improvement over standard eigenanalysis-based methods [6]. Finally, in Section V, we conclude with a discussion of the implications of our work and potential future research directions.

II. Multivariate self-similarity

A. Canonical model

Operator fractional Brownian motion (ofBm) is a canonical model for multidimensional scale-invariant structures in real-world data. We briefly recall its definition and some properties [15]. 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 Hurst exponents $\underline{H} = (H_1,\ldots,H_M)$, $0 < H_1 \le \ldots \le H_M < 1$. Let Σ be a pointwise covariance matrix with entries $(\Sigma)_{\ell,\ell'} = \sigma_\ell\sigma_{\ell'}\rho_{\ell,\ell'}$, where σ_ℓ^2 are the variances of the components and $\rho_{\ell,\ell'}$ their (pairwise) correlation coefficients. OfBm is defined as the Gaussian, stationary-increment stochastic process $\underline{B}_{P,\underline{H},\Sigma}(t) = P\underline{B}_{\underline{H},\Sigma}(t)$, where P is a real-valued, $M\times M$ invertible matrix that mixes together the components (changes the scaling coordinates) of $\underline{B}_{\underline{H},\Sigma}(t)$. Moreover, it satisfies the (operator) self-similarity relation

$$\{\underline{B}_{P,\underline{H},\Sigma}(t)\}_{t\in\mathbb{R}} \overset{\text{f.d.d.}}{=} \{a^{\mathbf{H}}\underline{B}_{P,\underline{H},\Sigma}(t/a))\}_{t\in\mathbb{R}}, \qquad (1)$$

$\forall a > 0$. In (1), the matrix Hurst exponent is given by $\mathbf{H} = P\mathrm{diag}(\underline{H})P^{-1}$, and $a^{\mathbf{H}} = \sum_{k=0}^{\infty}\log^k(a)\mathbf{H}^k/k!$. When the demixed process $P^{-1}\underline{B}_{P,\underline{H},\Sigma}(t) = \underline{B}_{\underline{H},\Sigma}(t)$ has uncorrelated entries, (1) is *instantaneously correlated*. Below, we only consider instantaneously correlated instances of ofBm ($\Sigma = I$).

B. Wavelet analysis of ofBm

Multivariate wavelet transform. Let ψ be a mother wavelet, i.e., a real-valued function such that $\int_{\mathbb{R}}\psi^2(t)dt = 1$. For all $k,j\in\mathbb{N}$, the multivariate discrete wavelet transform of $\{\underline{B}_{P,\underline{H},\Sigma}(t)\}_{t\in\mathbb{R}}$ is defined as $D(2^j,k) = (D_1(2^j,k),\ldots,D_M(2^j,k))$, where $D_\ell(2^j,k) = \langle 2^{-j/2}\psi(2^{-j}t-k)|\underline{B}_{P,\underline{H},\Sigma,\ell}(t)\rangle \in \mathbb{R}$ for $\ell\in\{1,\ldots,M\}$. For a detailed introduction to wavelet transforms, see [16]. It can be shown that the wavelet coefficients $\{D(2^j,k)\}_{k\in\mathbb{Z}}$ satisfy, for every fixed octave j, the operator self-similarity relation $\{D(2^j,k)\}_{k\in\mathbb{N}} \overset{\text{f.d.d.}}{=} \{2^{j(\mathbf{H}+\frac{1}{2}I)}D(1,k)\}_{k\in\mathbb{N}}$ [6], [7].

Eigenanalysis. Starting from the measurements (1), the *wavelet random matrix* (WRM) at octave $j\in(j_1,\ldots,j_2)$ is given by the $M\times M$ symmetric matrices

$$S(2^j) = \frac{1}{n_j}\sum_{k=1}^{n_j} D(2^j,k)D(2^j,k)^T, \qquad (2)$$

where n is the time series (sample) size and $n_j \simeq n/2^j$ is the number of wavelet coefficients available at scale 2^j. It was shown in [6], [7] that, in general, estimation based on the *entrywise* multiscale behavior of $S(2^j)$ is arbitrarily biased.

So, let $\lambda_1(2^j),\ldots,\lambda_M(2^j)$ be the *eigenvalues* of the random matrix $S(2^j)$ as in (2). Notably, it was shown in [6], [7] that, in the large-sample and large-scale limits $n,j\to\infty$,

$$\lambda_m(2^j) \approx \xi_m\cdot 2^{j(2H_m+1)},$$

for *scaling amplitude* constants $\xi_m > 0$. Consequently, the vector of Hurst exponents $\underline{H}$ can be efficiently estimated as follows. Consider the eigen-decompositions of the random matrices (2), i.e.,

$$S(2^j) = U_j\Lambda(2^j)U_j^T, \quad j\in(j_1,\ldots,j_2). \qquad (3)$$

In (3), $\Lambda(2^j) = \mathrm{diag}(\lambda_1(2^j),\ldots,\lambda_M(2^j))$ displays the eigenvalues of $S(2^j)$ along the main diagonal. Also, the orthogonal matrix $U_j\in\mathbb{R}^{M\times M}$ contains the eigenvectors $\mathbf{u}_{m,n}(2^j)$, $m=1,\ldots,M$, of $S(2^j)$ as columns. The wavelet estimator $(\widehat{H}_1,\ldots,\widehat{H}_M)$ for $\underline{H}$ is defined as log-eigenvalue regressions

$$\widehat{H}_m = \frac{1}{2}\left(\sum_{j=j_1}^{j_2} w_j\log_2\lambda_m(2^j) - 1\right),\ m=1,\ldots,M, \quad (4)$$

where the weights w_j satisfy $\sum_j jw_j = 1$ and $\sum_j w_j = 0$ (cf. [4]). Under mild assumptions, it was shown in [6], [7] that $(\widehat{H}_1,\ldots,\widehat{H}_M) \overset{\mathbb{P}}{\to} (H_1,\ldots,H_M)$ as $n,j\to\infty$, with asymptotically Gaussian fluctuations.

Eigenvectors. In [10], the asymptotic behavior of wavelet eigenvectors is further characterized. Suppose the Hurst exponents in (1) are pairwise distinct. Recast the mixing matrix $P = (P_1,\ldots,P_M)$ from (1) in terms of its column vectors. In the easily interpretable case where the matrix P is orthogonal, the eigenvectors of (3) asymptotically align with the columns of P, up to a sign. Specifically, we have

$$\min_{j=j_1,\ldots,j_2}\min_{m=1,\ldots,M}\big|\langle\mathbf{u}_{m,n}(2^j),P_m\rangle\big| = 1 - O_{\mathbb{P}}\left(\frac{1}{2^{j_1\delta_H}}\right) \quad (5)$$

as $n,j\to\infty$. In (5), $\delta_H = \min_{1\le m<m'\le M}(H_{m'}-H_m)$ and $O_{\mathbb{P}}$ means bounded in probability.

In the general case where P is not orthogonal, let us set $M=3$ for the sake of expositional clarity and fix any j. Then, as $n,j\to\infty$, and at the same convergence rate as in (5), up to a sign, the top eigenvector $\mathbf{u}_{3,n}(2^j)$ and P_3 eventually

align (cf. (5)), and the next eigenvector $\mathbf{u}_{2,n}(2^j)$ approaches span$\{P_2, P_3\}$ and converges. Consequently, the remaining eigenvector $\mathbf{u}_{1,n}(2^j)$ is pushed away from span$\{P_2, P_3\}$ and eventually converges. The case for general M can be inductively inferred from this description. Since, for any fixed m, the eigenvectors $\mathbf{u}_{m,n}(2^j)$, for $j \in (j_1, \ldots, j_2)$, exhibit the same asymptotic behavior, the WRMs in (3) for a range of large scales can be *approximately jointly diagonalized* by the eigenvectors $\mathbf{u}_{m,n}(2^{j_0})$, for $m = 1, \ldots, M$, for any large j_0.

III. Multiscale Eigenvector Based Estimator

A. Approximate joint eigenanalysis of ofBm

In this section, we introduce a new estimator for $(H_1, \ldots, H_M)$, building on the fact that the eigenvectors U_j of $S(2^j)$ converge to a common limiting matrix U. In a finite-sample setting, instead of scale-wise wavelet eigenanalysis as in (3), our key idea is to *jointly diagonalize* the set of matrices

$$\mathbf{S} = \{S(2^{J_1}), \ldots, S(2^{J_2})\} \tag{6}$$

approximately, using a single nonsingular matrix $\check{U}$. This leads to the approximate factorization:

$$S(2^j) \approx \check{U}\check{\Lambda}(2^j)\check{U}^T, \quad j \in (J_1, \ldots, J_2), \tag{7}$$

where $\check{\Lambda}(2^j)$ is an approximately diagonal matrix whose diagonal entries correspond to the approximate eigenvalues

$$(\check{\lambda}_1(2^j), \ldots, \check{\lambda}_M(2^j)) = (\check{\Lambda}_{11}(2^j), \ldots, \check{\Lambda}_{MM}(2^j)). \tag{8}$$

To construct the wavelet estimator $(\check{H}_1, \ldots, \check{H}_M)$ for $(H_1, \ldots, H_M)$, we replace the scale-wise eigenvalues $\lambda_m(2^j)$ in (4) with their approximations $\check{\lambda}_m(2^j)$.

B. Motivation for Joint Diagonalization

In light of the assumption that $2^{j_1}/n = o(1)$ [10], relation (5) demonstrates that the convergence of the M eigenvectors to an orthogonal matrix U is generally much *slower* than the growth of n. As a result, $\check{U}$ as in (7) may not be a close approximation of U. Yet, we can reasonably assume that the eigenvectors U_j at different scales, $j \in (J_1, \ldots, J_2)$, satisfy

$$U_j \approx \check{U}, \quad \text{even if } \check{U} \neq U. \tag{9}$$

This assumption provides a key advantage: the approximate diagonalization in (7) preserves the *ordering of eigenvalues across scales* by associating them with shared approximate eigenvectors. In contrast, the usual scale-wise approach (3) orders the eigenvalues at each scale independently by magnitude. This can lead to *inconsistent* eigenvalue ordering across scales, causing *permutations* known as *crossover* that introduce biases when used in (4). This is illustrated in Fig. 1 (top row).

C. Joint Diagonalization of WRMs

Joint diagonalization as an optimization problem. A key remaining step is to identify a suitable approximate diagonalizer $\check{U}$. The problem of jointly diagonalizing a set of matrices has been extensively studied, and several algorithms have been proposed for its solution [17]–[19]. Most commonly, these methods formalize (7) by defining $\check{U}$ as the minimizer

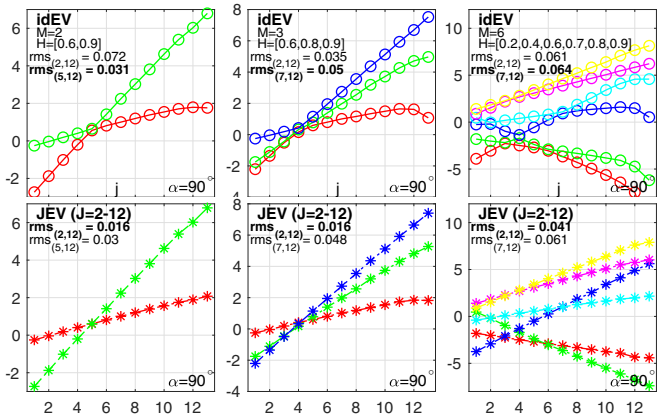

Fig. 1. **Log-scale plots for orthogonal mixing matrix and distinct H.** idEV (top) and JEV (bottom) for (from left to right) $M = 2, 3, 6$. Rms values are averaged over components and obtained for the indicated scaling range optimal for idEV and $(j_1, j_2) = (2, 12)$ for JEV, respectively.

of a functional $f_{\mathbf{S}}(U)$, defined for nonsingular matrices U and penalizing the magnitude of the off-diagonal elements in the approximately diagonalized matrices $\check{\Lambda}(2^j) = \check{U}^{-1}S(2^j)\check{U}$. This leads to the optimization problem

$$\check{U} = \text{argmin}_U f_{\mathbf{S}}(U), \; f_{\mathbf{S}}(U) = \frac{1}{2}\sum_{j=J_1}^{J_2}\sum_{i\neq j}\left|(U^{-1}S(2^j)U)_{ij}\right|^2. \tag{10}$$

This objective function is highly non-convex and has been studied in detail in [20], where explicit and efficient expressions for its gradient, $\nabla f_{\mathbf{S}}|_U$, and Hessian operator, $H_{\mathbf{S}}|_U(Z)$, have been derived. These expressions allow for effective numerical optimization. A recent method proposed in [14] employs a conjugate gradient (CG) algorithm with a multiplicative change of basis to solve (10). This approach significantly outperforms previous algorithms in terms of accuracy.

Special Case: Unitary Joint Diagonalization. Since the matrices $S(2^j)$ are symmetric, we impose the constraint that U be an orthogonal matrix, i.e., $U^{-1} = U^T$. To adapt the algorithm from [14] to this special case, we enforce the unitary constraint by projecting onto the set of orthogonal matrices at each iteration. This projection is performed using the singular value decomposition (SVD): given any matrix A, the closest unitary matrix B is given by $B = VW^T$, where $V\Sigma W^T$ is the SVD of A [20], [21]. This projection step ensures that U remains unitary throughout the optimization process.

Conjugate Gradient Algorithm for Estimating $\check{U}$. The CG algorithm adapted from [14] for estimating $\check{U}$ consists of the following steps:

Initialization ($i = 0$): Set $\mathbf{S}_0 = \mathbf{S}$ and $U_0 = U_{init}$. Compute the initial search direction $R_0 = -\nabla f_{\mathbf{S}_0}|_I$ and perform steps 3 to 7 (gradient descent step).

1) Set $\tilde{R}_{i-1} = (I + \lambda_{i-1}R_{i-1})^{-1}R_{i-1}$ given R_{i-1}, $\mathbf{S}_i$.
2) Compute the gradient $\nabla f_{\mathbf{S}_i}|_I$ and update the search direction using Daniel's rule for nonlinear conjugate gradient, $R_i = -\nabla f_{\mathbf{S}_i}|_I + \beta_{i-1}\tilde{R}_{i-1}$, where $\beta_{i-1} = \frac{\langle \nabla f_{\mathbf{S}_i}|_I, H_{\mathbf{S}_i}|_I(\tilde{R}_{i-1})\rangle}{\langle \tilde{R}_{i-1}, H_{\mathbf{S}_i}|_I(\tilde{R}_{i-1})\rangle}$ [22].

3) Compute the Newton step-size $\lambda_i = -\frac{\langle \nabla f_{\mathbf{S}_i|_I, R_i} \rangle}{\langle R_i, H_{\mathbf{S}|_I(R_i)} \rangle}$.
4) Compute the SVD $(I + \lambda_i R_i) = V_i \Sigma W_i^T$
5) Update the matrix set $\mathbf{S}_{i+1} = (V_i W_i^T)^T \mathbf{S}_i (V_i W_i^T)$.
6) Update the eigenvector estimate $U_{i+1} = U_i (V_i W_i^T)$.
7) If not converged, set $i = i + 1$ and go back to 1).

IV. Numerical results

Monte Carlo Simulation. We apply the log-regression estimator (4) using both the scale-wise eigenanalysis (3) and the proposed approximate multiscale eigenanalysis (8), denoted *idEV* and *JEV*, respectively. These methods are evaluated over a large number of realizations, $N_{MC} = 100$, of an ofBm process with $n = 2^{16}$ and $M = 2, 3, 6, 32$. For better visibility in log-log plots and without loss of generality, we analyze the increment process of ofBm. The mixing matrix $P = P_\alpha$ is constructed as an orthonormal matrix drawn at random and modified such that the last column forms an angle α with the second-to-last column while remaining orthogonal to all other columns. Specifically, when $\alpha = 90°$, P_α remains fully orthonormal, whereas for $\alpha \to 0°$, it degenerates to rank $M-1$. For the analysis, we employ a Daubechies wavelet with two vanishing moments. For idEV, a case-wise optimal set of scales (j_1, j_2) is selected. For JEV, unless stated otherwise, the scales are fixed at $(j_1, j_2) = (2, 12)$ and $(J_1, J_2) = (2, 12)$. The CG algorithm from Sec. III-C is executed for 100 iterations using normalized matrices $\tilde{S}(2^j) = S(2^j)/\|S(2^j)\|$. The initialization is set as $U_{init} = \text{eig}\left(\sum_{j=J_1}^{J_2} \tilde{S}(2^j)\right)$. Performance is quantified using the mean, standard deviation (std), and root mean squared error (rms) over all realizations.

Performance for orthogonal mixing matrices. Fig. 1 presents average log-log plots for $\lambda(2^j)$ (idEV, top) and $\check{\lambda}(2^j)$ (JEV, bottom) for $M = 2, 3, 6$, where the mixing matrix P is orthogonal ($\alpha = 90°$) and all values of H are distinct. In this case, the eigenvectors U_j asymptotically align with P. Clearly, idEV exhibits scaling laws with crossovers, leading to poor estimates for H with large rms values when all scales are used in (4). While this issue can be mitigated by selecting a larger value for j_1 (with the optimal value highlighted in bold in Fig. 1), the challenge becomes increasingly difficult as the number of components M grows. In such cases, multiple crossovers may occur, often requiring manual inspection that is tedious, error-prone, and difficult to automate. In contrast, our proposed JEV method effectively resolves the crossover issue and produces nearly perfect scaling laws. Consequently, almost all scales can be used for estimating H, significantly improving overall performance (in terms of rms values) without requiring manual tuning of (j_1, j_2). Tab. I further details the performance in terms of mean, standard deviation (std), and rms for each H individually. Unlike idEV, which relies on manually tuned scales (j_1, j_2), JEV, with fixed $(j_1, j_2) = (2, 12)$, provides nearly bias-free estimates and outperforms idEV by a factor of 2 to 5 in rms values.

Robustness to identical H ($\delta_H = 0$). Fig. 2 (idEV (left), JEV with $J_1 = 2$ (center) and $J_1 = 3$ (right), respectively) presents results for $M = 6$ with an orthogonal mixing matrix

	$M=2$		$M=3$			$M=6$					
H idEV	0.6	0.9	0.6	0.8	0.9	0.2	0.4	0.6	0.7	0.8	0.9
mean	0.59	0.89	0.57	0.78	0.90	0.13	0.35	0.56	0.70	0.80	0.86
std	2.99	2.86	4.22	4.74	4.13	3.64	4.48	5.32	4.40	5.54	4.57
rms	3.19	3.03	5.38	5.29	4.16	8.23	6.81	6.69	4.42	5.56	6.29
mean	0.60	0.90	0.59	0.80	0.90	0.18	0.39	0.61	0.70	0.80	0.89
std	1.64	1.62	1.49	1.72	1.51	1.54	1.66	6.48	1.40	1.64	6.35
rms	1.66	1.63	1.58	1.74	1.51	2.72	1.89	6.53	1.41	1.65	6.45

TABLE I

Performance for orthogonal mixing matrix and distinct H. Mean, std and rms for idEV (top) and JEV (bottom) for $M = 2, 3, 6$ (from left to right), obtained with optimal (for overall rms) scales (j_1, j_2) for idEV (given in Fig. 1) and $(j_1, j_2) = (2, 12)$ for JEV, respectively. Best results marked in color.

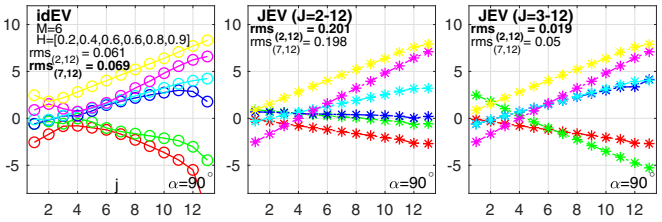

Fig. 2. **Log-scale plots for orthogonal mixing matrix and identical H.** idEV (left) and JEV with $J_1 = 2$ (center) and $J_1 = 3$ (right). Rms values are averaged over components and obtained for the indicated scaling range optimal for idEV and $(j_1, j_2) = (2, 12)$ for JEV, respectively.

in the extreme case where two Hurst exponents are identical, and convergence of eigenvectors is not guaranteed (cf. (5)). For idEV, the results are similar to those observed for $M = 6$ previously. More interestingly, the center plot illustrates that JEV can indeed fail when exponents are identical, as predicted by (5). However, the plot on the right demonstrates that this issue can be effectively mitigated by selecting a larger scale for J_1, thereby restoring excellent estimation performance.

Robustness to non-orthogonal mixing matrices. Fig. 3 presents average log-log plots for $\lambda(2^j)$ (idEV, left) and $\check{\lambda}(2^j)$ (JEV, center) for $M = 3$ across multiple instances of mixing matrices P_α with $\alpha = 61°, 42°, 23°$. Similar to the orthogonal mixing case ($\alpha = 90°$), scaling laws for idEV are significantly affected for all three mixing matrices, leading to poor estimates for H and nearly identical rms values. For JEV, performance remains comparable to the orthogonal case for the two smaller modes ($H = 0.6, 0.8$) as α decreases. However, estimation for the largest mode ($H = 0.9$) deteriorates, requiring the use of large scales in (4). Moreover, for $\alpha = 42°$ and $\alpha = 23°$, setting $J_1 = 10$ is required to obtain accurate results, indicating that the eigenvector $\mathbf{u}_1(2^j)$ varies significantly across scales. Eventually, for $\alpha = 23°$, the log-eigenvalues for JEV associated with $H = 0.8, 0.9$ collapse with those for idEV. However, the JEV log-eigenvalues for $H = 0.6$ scale nearly perfectly across all scales. This behavior is further examined in Tab. II, which summarizes JEV performance as α and J_1 vary. The results indicate that J_1 must be increased as α decreases. The performance for the largest modes ($H = 0.8, 0.9$) eventually deteriorates. Remarkably, estimation for the smallest mode ($H = 0.6$) remains nearly as accurate as in the orthogonal case, regardless of the precise value of α.

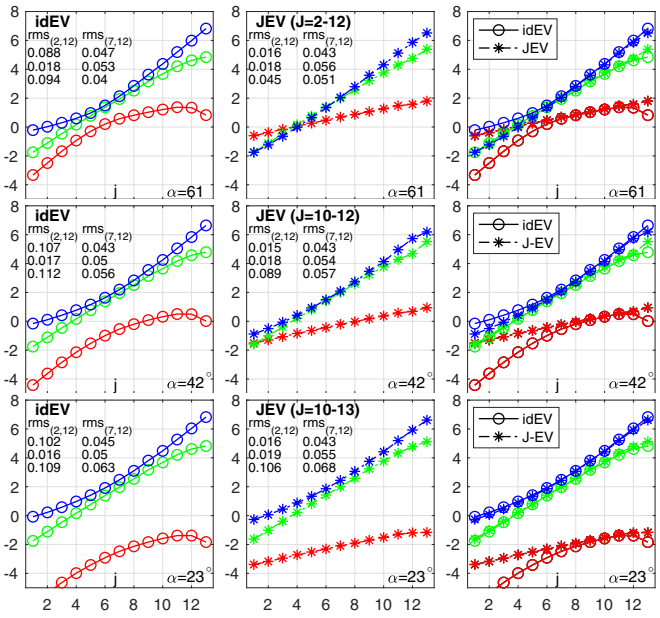

Fig. 3. **Log-scale plots for non-orthogonal mixing matrix and distiinct H:** idEV (left), JEV (center) and overlay (right) for mixing matrix W_α with $\alpha = 61°, 42°, 23°$ (from top to bottom); rms values are given for each H separately, with optimal (for overall rms) scales (j_1, j_2) for idEV and $(j_1, j_2) = (2, 12)$ for JEV, respectively.

mean									
	$H = 0.6$			$H = 0.8$			$H = 0.9$		
J_1	2	6	10	2	6	10	2	6	10
$\alpha = 61$	0.60	0.61	0.60	0.80	0.80	0.80	0.86	0.86	0.87
$\alpha = 42$	0.69	0.62	0.60	0.79	0.78	0.79	0.83	0.84	0.84
$\alpha = 23$	0.70	0.63	0.60	0.79	0.77	0.77	0.84	0.84	0.84
std $\times 10^2$									
	2	6	10	2	6	10	2	6	10
$\alpha = 61$	1.17	1.13	1.14	1.19	1.18	1.22	1.62	1.60	1.66
$\alpha = 42$	1.30	1.18	1.17	2.51	0.99	0.98	2.64	2.68	2.71
$\alpha = 23$	1.08	1.15	1.18	3.25	1.08	1.07	3.71	3.82	3.89
rms $\times 10^2$									
	2	6	10	2	6	10	2	6	10
$\alpha = 61$	1.18	1.24	1.14	1.25	1.24	1.32	3.89	4.37	3.84
$\alpha = 42$	8.73	2.43	1.20	2.79	2.12	1.50	7.85	7.03	6.54
$\alpha = 23$	9.65	2.83	1.23	3.33	3.12	2.88	6.90	6.83	6.76

TABLE II

Performance for nonorthogonal mixing matrix: ESTIMATION FOR H FOR JEV WITH OPTIMAL (FOR OVERALL RMS) SCALES (j_1, j_2) SELECTED FOR EACH H INDIVIDUALLY, RESPECTIVELY, AND DIFFERENT VALUES FOR J_1.

Overall, these findings demonstrate that the proposed JEV method offers significant advantages over idEV not only for orthogonal mixing matrices but also under more challenging scenarios, ensuring robust and improved performance.

V. CONCLUSIONS

In this work, we introduced a novel multiscale eigenanalysis approach for estimating Hurst exponents in multivariate self-similar processes. By leveraging joint diagonalization of wavelet random matrices, our method overcomes limitations of traditional scale-wise eigenanalysis, particularly in the presence of scaling amplitude discrepancies. This approach provides a robust estimation framework that preserves eigenvalue ordering across scales, reducing estimation biases caused by crossover effects. Through extensive Monte Carlo simulations, we demonstrated the superiority of our method in both orthogonal and non-orthogonal mixing scenarios. Our findings highlight the potential of joint eigenvector-based wavelet analysis as a powerful tool for analyzing self-similar dynamics in multivariate time series. Future research includes extending this methodology to broader classes of mixing matrices and to non-instantaneous correlations, as well as studying its applicability to real-world datasets in neuroscience.

REFERENCES

[1] D. Sornette, *Critical phenomena in natural sciences: chaos, fractals, selforganization and disorder: concepts and tools*, Springer, 2006.

[2] D. La Rocca, H. Wendt, V. van Wassenhove, P. Ciuciu, and P. Abry, "Revisiting functional connectivity for infraslow scale-free brain dynamics using complex wavelets," *Front. Physiol.*, vol. 11, pp. 1651, 2021.

[3] P. Flandrin, "Wavelet analysis and synthesis of fractional Brownian motion," *IEEE Trans. Info. Theory*, vol. 38, pp. 910 – 917, 1992.

[4] 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.

[5] J. Hualde and P.M. Robinson, "Semiparametric inference in multivariate fractionally cointegrated systems," *J. Econometrics*, vol. 157, no. 2, pp. 492–511, 2010.

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

[7] P. Abry and G. Didier, "Wavelet eigenvalue regression for n-variate operator fractional Brownian motion," *J. Multivar. Anal.*, vol. 168, pp. 75–104, 2018.

[8] G. W. Anderson, A. Guionnet, and O. Zeitouni, *An Introduction to Random Matrices*, Number 118. Cambridge University Press, 2010.

[9] M. J. Wainwright, *High-Dimensional Statistics: a Non-Asymptotic Viewpoint*, vol. 48, Cambridge University Press, 2019.

[10] P. Abry, B. C. Boniece, G. Didier, and H. Wendt, "On high-dimensional wavelet eigenanalysis," *Ann. Appl. Probab.*, vol. 34, no. 6, pp. 5287–5350, 2024.

[11] P. Abry, G. Didier, O. Orejola, and H. Wendt, "On the empirical spectral distribution of large wavelet random matrices based on mixed-Gaussian fractional measurements in moderately high dimensions," *Electronic Journal of Probability*, vol. 30, pp. 1–48, 2025.

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

[13] P. Abry, G. Didier, and H. Li, "Two-step wavelet-based estimation for gaussian mixed fractional processes," *Statistical Inference for Stochastic Processes*, vol. 22, pp. 157–185, 2019.

[14] E. Troedsson, M. Carlsson, and H. Wendt, "On gradient based descent algorithms for joint diagonalization of matrices," in *European Signal Processing Conference (EUSIPCO)*, Lyon, France, August 2024.

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

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

[17] A. Mesloub, A. Belouchrani, and K. Abed-Meraim, "Efficient and stable joint eigenvalue decomposition based on generalized givens rotations," in *Proc. European Signal Process. Conf. (EUSIPCO)*, Rome, Italy, 2018.

[18] R. Iferroudjene, K. A. Meraim, and A. Belouchrani, "A new Jacobi-like method for joint diagonalization of arbitrary non-defective matrices," *Applied Math. and Comput.*, vol. 211, no. 2, pp. 363–373, 2009.

[19] R. André, X. Luciani, and E. Moreau, "Joint eigenvalue decomposition algorithms based on first-order Taylor expansion," *IEEE T. Signal Process.*, vol. 68, pp. 1716–1727, 2020.

[20] E. Troedsson, D. Falkowski, C.-F. Lidgren, H. Wendt, and M. Carlsson, "On joint eigen-decomposition of matrices," *Arxiv:2409.10292*, 2025.

[21] J. B. Keller, "Closest unitary, orthogonal and Hermitian operators to a given operator," *Math. Magazine*, vol. 48, no. 4, pp. 192–197, 1975.

[22] J. W. Daniel, "The conjugate gradient method for linear and nonlinear operator equations," *SIAM J. Num. Anal.*, vol. 4, no. 1, pp. 10–26, 1967.