Wavelet-based demixing for multivariate fractional Brownian motion

Kinan Abbas[1], Herwig Wendt[2], Gustavo Didier[3], Patrice Abry[1]
[1]*CNRS, ENS de Lyon, Laboratoire de Physique,* Lyon, France. {kinan.abbas, patrice.abry}@ens-lyon.fr
[2]*CNRS, IRIT, University of Toulouse*, Toulouse, France. herwig.wendt@irit.fr
[3]*Mathematics Department, Tulane University*, New Orleans, LA, USA. gdidier@tulane.edu

***Abstract*—Multivariate selfsimilarity has recently been put forward as a relevant analysis concept for applications entailing the joint analysis of a set of correlated signals characterized by scale-free temporal dynamics. However, multivariate selfsimilarity analysis remained so far essentially restricted to the estimation of a collection of selfsimilarity parameters, or Hurst exponents, while a significant piece of information, relevant for applications, is conveyed in the potential correlations and mixing amongst components. This work aims to advance the state of the art by proposing the joint estimation of the selfsimilarity parameters and of the mixing and correlation matrices. This relies on embedding the statistical properties of multivariate wavelet coefficients of the multivariate fractional Brownian motion used as a model into an optimization framework. Estimation performance are quantified by means of Monte Carlo simulations under various multivariate settings. Application to real financial data is illustrated and discussed.**

***Index Terms*—Multivariate selfsimilarity, matrix logarithm, wavelet analysis, optimization**

I. Introduction

Context. Selfsimilarity [1] has been widely used to study theoretically or to model in a broad variety of applications scale-free dynamics, where temporal or spatial structure are governed by a broad continuum or scales, rather than by a few specific scales (cf. e.g., [2]–[6]). In univariate settings, selfsimilarity analysis consists of the estimation of the so-called Hurst exponent or selfsimilarity parameter H, which can be efficiently and robustly achieved by means of wavelet analysis [7]. However, most modern applications produce multivariate data and thus call for multivariate selfsimilarity modeling [8] and analysis [9]–[13]. Moreover, besides a vector of selfsimilarity parameters $\underline{H}$, multivariate selfsimilarity naturally entails correlations and mixing amongst components, which significantly both enriches and complicates its practical use for applications [8], [13], [14]. The present work is the first attempt for the joint estimation of these selfsimilarity, correlation and mixing parameters.

Related works. Fractional Brownian motion [15] has been widely used to model and analyze scale-free dynamics and selfsimilarity in univariate settings. It has also been massively documented that multiscale representations such as wavelet transforms are well suited to yield theoretically well-grounded and practically robust estimation of the selfsimilarity parameter H [7]. In the last two decades, models for multivariate selfsimilarity were proposed [8], [9], [13], [14], [16]. In the restricted setting with correlation but no mixing amongst components, several multivariate extensions of univariate analysis were proposed based on wavelet transforms [10] or increments [9], all exploiting component-wise power law scaling of multiscale statistics. But the general setting with mixing between components is far more involved because it blurs the correspondence between scale-free dynamics and scale covariant statistics so that univariate analysis based estimation strategies are ineffective [10]. Alternatively, it was shown and studied that the behaviors of the eigenvalues of the wavelet matrices across scales permits robust estimation of the vector of Hurst parameters [11]–[13]. However, except for one preliminary and computationally intensive attempt in bivariate settings [17], there exists no strategy for the joint estimation of the selfsimilarity, correlation and mixing parameters.

Goals, contributions and outline. The overarching goal of the present work is to devise a multivariate wavelet based joint estimation of the selfsimilarity, correlation and mixing parameters characterizing multivariate selfsimilarity, and to study and assess estimation performance. To that end, Section II briefly recalls the definition of multivariate fractional Brownian motion (mfBm) and the key equations associated with its wavelet decomposition. Section III describes the first key contributions of the present work: i) The principles of the wavelet-based multivariate wavelet based full identification of the multivariate fractional Brownian motion used as model for multivariate selfsimilarity, and ii) the matrix logarithm based corresponding optimization procedure. As an additional contribution, Section IV assesses estimation performance as a function of mfBm parameters, by means of Monte Carlo simulations conducted on a large number of synthetic realizations of mfBm, and quantifies and discusses benefits compared to state-of-the-art estimation procedures. Finally, the interest and benefits of the proposed demixing strategy are illustrated at work on real financial data in Section V.

II. Multivariate selfsimilarity

A. Model: Multivariate fractional Brownian motion

Elaborating on the general model for multivariate selfsimilarity, operator fractional Brownian motion [8], a more versatile and better suited to application modeling model

G.D. is partially supported by the award NSF DMS 2515732. G.D.'s long-term visits to ENS de Lyon were supported by the school, the CNRS and the Simons Foundation collaboration grant #714014.

referred to as multivariate fractional Brownian motion (mfBm) $B_{\mathbf{\Sigma},\mathbf{W},\underline{H}}(t)$ was proposed in [13] and defined as:

$$B_{\mathbf{W},\mathbf{\Sigma},\underline{H}}(t) := \mathbf{W} B_{\mathbf{\Sigma},\underline{H}}(t)^\top, \tag{1}$$

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

B. Analysis: Multivariate wavelet transform

1) Multivariate wavelet transform: Let ψ_0 denote a compact support reference pattern, referred to as the *mother wavelet*, labelled ψ_0, and characterized by its number of vanishing moments N_ψ. Let Y denote M-variate times series. The vector of coefficients of the multivariate Discrete Wavelet Transform is defined as $D_Y(2^j, k) = \int_{\mathbb{R}} \psi(j,k) Y(t) \mathrm{d}t$, with $\psi(j,k) = 2^{-j/2}\psi_0(2^{-j}t - k)$ a collection of dilated and translated templates of ψ_0.

2) Multivariate wavelet spectrum: Expanding on univariate selfsimilarity analysis [7], the multivariate wavelet spectrum is defined as the collection of $M \times M$ symmetric, positive semi-definite scale-dependent wavelet coefficient correlation matrices: [11]–[13]:

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

For mfBm $B_{\mathbf{W},\mathbf{\Sigma},\underline{H}}$, it has been shown that wavelet coefficients reproduce selfsimilarity as [11], [12]:

$$\left\{D_{B_{\mathbf{W},\mathbf{\Sigma},\underline{H}}}(2^j, k)\right\}_{k\in\mathbb{Z}} \overset{fdd}{=} \left\{W 2^{j(\underline{H}+\frac{1}{2}\mathbb{I})} D_{B_{\mathbf{\Sigma},\underline{H}}}(2^0, k)\right\}_{k\in\mathbb{Z}}, \tag{3}$$

where $2^{j(\underline{H}+\frac{1}{2})} = \mathrm{diag}(2^{j(H_1+1/2)}, \ldots, 2^{j(H_M+1/2)})$.
The wavelet spectrum of mfBm thus satisfies:

$$\mathbf{S}_{B_{\mathbf{W},\mathbf{\Sigma},\underline{H}}}(2^j) \overset{d}{=} \mathbf{W} 2^{j(\underline{H}+\frac{1}{2}\mathbb{I})}\, \mathbf{S}_{B_{\mathbf{\Sigma},\underline{H}}}(2^0)\, 2^{j(\underline{H}^\top+\frac{1}{2}\mathbb{I})} \mathbf{W}^\top, \tag{4}$$

where $\mathbf{S}_{B_{\mathbf{\Sigma},\underline{H}}}(2^0)$ corresponds to the wavelet spectrum of the unmixed mfBm $B_{\mathbf{\Sigma},\underline{H}}$ at scale $2^0 = 1$, and is thus directly related to $\mathbf{\Sigma}$, as well as to $\underline{H}$ [10], but not to $\mathbf{W}$.

3) Estimation of the vector of selfsimilarity parameters: When there is no mixing, i.e, $\mathbf{W} \equiv \mathbb{I}$, (4) straightforwardly shows that the diagonal entries satisfy, on average, $\mathbb{E}\left(\mathbf{S}_{B_{\mathbb{I},\mathbf{\Sigma},\underline{H}}}\right)_{m,m}(2^j) = \beta_m 2^{j(2H_m+1)}$, thus leading to the classical univariate linear regression-based estimators [7]:

$$\hat{H}_m^{\mathrm{U}} = \frac{1}{2}\left(\sum_{j=j_1}^{j_2} w_j \log_2 \mathbf{S}_{m,m}(2^j) - 1\right), \tag{5}$$

where $j \in (j_1, j_2)$ denote the range of octaves involved and where the regression weights satisfy: $\sum_{j=j_1}^{j_2} j w_j = 1$ and $\sum_{j=j_1}^{j_2} w_j = 0$.

In the presence of mixing, i.e., when $\mathbf{W} \neq \mathbb{I}$, (4) shows that (the expectation of) each entry of $\mathbf{S}_{B_{\mathbf{W},\mathbf{\Sigma},\underline{H}}}(2^j)$ generally consists of a mixture of power laws with respect to scales 2^j:

$$\mathbb{E}\left(\mathbf{S}_{B_{\mathbf{W},\mathbf{\Sigma},\underline{H}}}\right)_{m,m'}(2^j) = \sum_{k=1}^{M}\sum_{k'=1}^{M} \alpha_{k,k'}^{m,m'} 2^{j(H_k+H_{k'}+1)}, \tag{6}$$

with $\alpha_{k,k'}^{m,m'}$ a priori depending on both $\mathbf{W}$ and $\mathbf{\Sigma}$. In other words, the entry-wise identification of scale-free dynamics is lost, which precludes the use of (5).

However, it has been shown that (on average) each eigenvalue $\lambda_m(2^j)$ of $\mathbb{E}\mathbf{S}_{B_{\mathbf{W},\mathbf{\Sigma},\underline{H}}}(2^j)$ behaves asymptotically as a power law with respect to the scale 2^j, where the scaling exponent is controlled by H_m [11]–[13]:

$$\mathbb{E}\lambda_m(2^j) \sim \xi_m(2^0)\,(2^j)^{2H_m+1}, \quad \text{as } j \to +\infty. \tag{7}$$

(7) restores the relationship between scale-free dynamics and power-laws, and naturally suggests that the estimation of the vector of selfsimilarity parameters $\underline{H}$ can be performed by a collection of individual linear regressions on the logarithms of the eigenvalues $\hat{\lambda}_m(2^j)$, $m = 1, \ldots, M$, of $\mathbf{S}_{B_{\mathbf{W},\mathbf{\Sigma},\underline{H}}}(2^j)$ versus the logarithms of the scales 2^j:

$$\hat{H}_m^{\mathrm{M}} = \frac{1}{2}\left(\sum_{j=j_1}^{j_2} w_j \log_2 \hat{\lambda}_m(2^j) - 1\right). \tag{8}$$

The asymptotic estimation performance of $\hat{H}_m^{\mathrm{M}}$ were thoroughly studied theoretically and practically in [11]–[13].

III. Multivariate selfsimilarity: Demixing

A. Parametric estimation

By focusing on the behaviours of the eigenvalues of $\mathbf{S}_{B_{\mathbf{W},\mathbf{\Sigma},\underline{H}}}(2^j)$ across scales, relation (7) and the semi-parametric estimators defined in (8) permit to estimate the selfsimilarity parameters $\underline{H}$, but not $\mathbf{W}$ or $\mathbf{\Sigma}$.

Instead, (4) suggests that $\underline{H}$, $\mathbf{W}$ and $\mathbf{S}_{B_{\mathbf{\Sigma},\underline{H}}}(2^0)$ can be jointly estimated by adjustment to the behavior across scales of the collection of matrices $\mathbf{S}_{B_{\mathbf{W},\mathbf{\Sigma},\underline{H}}}(2^j)$. This motivates us to propose a parametric estimation procedure as:

$$\left(\hat{\underline{H}}^{(o)}, \hat{\mathbf{W}}^{(o)}, \widehat{\mathbf{S}^{(o)}_{B_{\mathbf{\Sigma},\underline{H}}(2^0)}}\right) := \mathrm{argmin}_{\underline{H},\mathbf{W},\mathbf{S}} \mathcal{L}(\underline{H}, \mathbf{W}, \mathbf{S}), \tag{9}$$

$$\mathcal{L}(\underline{H}, \mathbf{W}, \mathbf{S}) := \frac{1}{|\mathcal{J}|}\sum_{j\in\mathcal{J}} \Big\| \log \mathbf{S}_Y(2^j) - \log\left(\mathbf{W} 2^{j(\underline{H}+\frac{1}{2})} S 2^{j(\underline{H}+\frac{1}{2})} \mathbf{W}^\top\right)\Big\|_F^2, \tag{10}$$

with $\mathcal{J} = (j_1, \ldots, j_2)$ and where $\log$ denotes the matrix logarithm. For a symmetric semi-positive definite (SPD) matrix S, with eigenvalue decomposition $S = U\Lambda U^T$, the matrix logarithm takes the simple form $\log S = U \log(\Lambda) U^T$. An estimate for $\mathbf{\Sigma}$ can then be obtained by applying the inverse of the estimated mixing matrix to original data, and estimating the sample point covariance of the unmixed data, or by combining $\widehat{\mathbf{S}^{(o)}_{B_{\mathbf{\Sigma},\underline{H}}(2^0)}}$ and $\hat{\underline{H}}^{(o)}$.

Taking the Logarithm of the matrices $\mathbf{S}_{B_{\mathbf{W},\mathbf{\Sigma},\underline{H}}}(2^j)$ constitutes a key point of the parametric estimation proposed here. It turns algebraic into linear growths and thus ensures that all scales, from fine to coarse, contribute to estimation. The use of the matrix logarithm is thus reminiscent of the use of the classical logarithm in estimators in (5) or (8), and thus of the intimate relationship between power-laws and scale-free dynamics. Since all $\mathbf{S}_{B_{\mathbf{W},\mathbf{\Sigma},\underline{H}}}(2^j)$ involved in (9) are covariance matrices, the Frobenius norm applied to the matrix logarithm corresponds to the log Euclidean distance on the SPD cone, which preserves symmetry and positive-definiteness while remaining computationally simple. The use of advanced affine-invariant Riemannian metrics and optimization schemes on SPD manifolds [18] will be addressed in future works.

B. Optimization procedure

The minimization problem in (9) is solved using a joint, (unconstrained) optimization over the self-similarity parameters $\underline{H}$, the mixing matrix $\mathbf{W}$, and the scale-normalized covariance matrix ($S_{B_{\Sigma,\underline{H}}}(2^0)$). Specifically, we rely on a Quasi-Newton method with a cubic line search procedure using the BFGS formula [19], as implemented by the MATLAB function fminunc. No alternating or block-coordinate strategy is used.

IV. Estimation performance

A. Monte Carlo simulation set-up

Synthetic data. Estimation performance are quantified by means of Monte Carlo simulations conducted from $N = 100$ realizations of mfBm. Sample paths (of length $n = 2^{18}$) of mfBm are produced using the procedure and toolbox described in [20], for various configurations of $(\underline{H}, \mathbf{\Sigma}, \mathbf{W})$.

Multivariate wavelet transform. Wavelet analysis is performed using the procedures and toolbox described in [13], using the Daubechies-3 least asymmetric orthonormal wavelet. The range of scales involved in (10) is $\mathcal{J} = (3 : 15)$.

Optimisation. The vector $\underline{H}$ is initialized from the eigenvalue-based estimations using (8). Based on the thorough analysis of mfBm structure in [12], Matrix $\mathbf{W}$ is initialized as the eigenvectors obtained from the diagonalisation of $\mathbf{S}_Y(2^j)$ at the largest scale in $\mathcal{J}$. In the absence of prior knowledge, Matrix $\mathbf{S}_{B_{\mathbf{\Sigma},\underline{H}}}(2^0)$ is initialized as diagonal. The Quasi-Newton descent is stopped at 100.000 iterations or when the cost function is no longer decreasing.

Performance assessment. As classical in demixing contexts, the quality of the estimated mixing matrix $\widehat{\mathbf{W}}$ is quantified using the Mixing Error Ratio (MER), as implemented e.g., in the Blind Source Separation Evaluation (BSS Eval) toolbox [21]. MER consists of an SNR-like criterion that accounts for column permutation, sign, and scaling ambiguities. Performance for the estimation of $\underline{H}$ are quantified by means of bias, variance and root Mean-Squared Error (rMSE). To assess the benefits of correctly estimating $\mathbf{W}$, performance of $\hat{\underline{H}}^{(o)}$ are compared against those of the univariate and multivariate estimates obtained prior to demixing, $\hat{\underline{H}}^U_{pre}$, $\hat{\underline{H}}^M_{pre}$, from (5) and (8) applied to the original data Y, as well as against the univariate and multivariate estimates, $\hat{\underline{H}}^U_{post}$ and $\hat{\underline{H}}^M_{post}$ obtained by applying (5) and (8) to demixed data $(\hat{\mathbf{W}}^{(o)})^{-1}Y$. Estimation performance for Σ is essentially driven by the accuracy of estimation for W and are not reported here for space reasons.

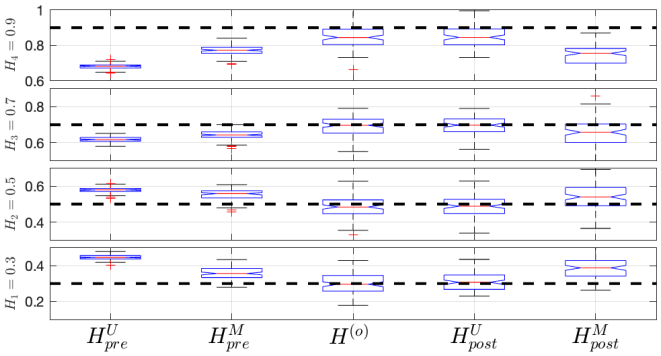

Fig. 1. Boxplots for the estimates of $\underline{H}$ across 100 realizations of mfBm.

B. Proof of concept: 4-variate setting

To illustrate the effectiveness of the proposed estimation strategy, a first Monte Carlo simulation is conducted in a 4-variate setting with $\underline{H} = (0.3, 0.5, 0.7, 0.9)$, a non diagonal covariance matrix $\mathbf{\Sigma}$ with significant off-diagonal entries, and a demanding mixing matrix $\mathbf{W} = [1, 0.5, 0, 0; 0.6, 1, 0.3, 0.4; 0, 0.4, 1, 0.5; 0, 0.3, 0.4, 1]$. Performance are reported in Table I and Fig. 1. They show that without demixing, univariate estimates, $\hat{\underline{H}}^U_{pre}$ are strongly biased, and differ significantly from the multivariate estimates $\hat{\underline{H}}^M_{pre}$, thus clearly indicating the presence of mixing. The demixing procedure proposed here yields significantly improved estimates, $\hat{\underline{H}}^{(o)}$, in terms of both biases and rMSE.

	$\hat{\underline{H}}^U_{pre}$	$\hat{\underline{H}}^M_{pre}$	$\hat{\underline{H}}^{(o)}$	$\hat{\underline{H}}^U_{post}$	$\hat{\underline{H}}^M_{post}$
MSE	0.13	0.08	0.07	0.05	0.05
Bias	0.11	0.08	0.04	0.02	0.02
std	0.02	0.03	0.05	0.04	0.04

TABLE I
Estimation performance for $\underline{H}$ averaged across 100 realizations of 4-variate mfBm.

The univariate and multivariate estimates, $\hat{\underline{H}}^U_{post}$ and $\hat{\underline{H}}^M_{post}$, obtained from the demixed data $(\hat{\mathbf{W}}^{(o)})^{-1}Y$, are observed to be consistent one with the other, and consistent with those obtained from the optimisation procedure $\hat{\underline{H}}^{(o)}$. Results also show that biases and rMSE for $\hat{\underline{H}}^U_{post}$ are essentially identical to those of $\hat{\underline{H}}^{(o)}$. This Monte Carlo simulation has been repeated for several different choices of matrices $\mathbf{W}$ and $\mathbf{\Sigma}$ or selfsimilarity parameters $\underline{H}$, yielding identical conclusions. This illustrates and validates the effectiveness of the proposed wavelet-based demixing strategy, in $M > 2$ settings, without the need to know the ground-truth mixing matrix $\mathbf{W}$.

C. Performance as a function of the difference in $\underline{H}$

For mfBm, the key difference amongst components lies in their temporal dynamics, driven by different selfsimilarity

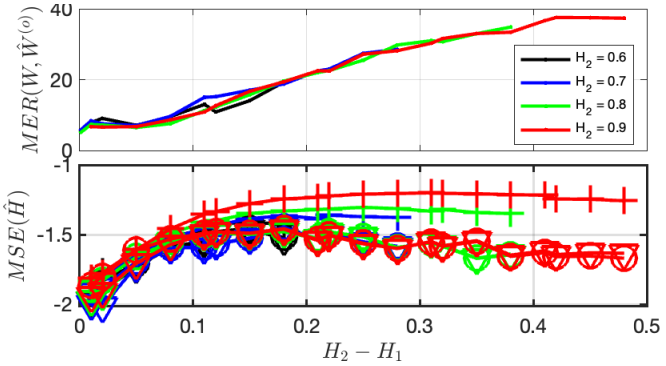

Fig. 2. Estimation performance as function of $\Delta_H = H_2 - H_1$, for four different values of H_2, averaged across 100 realizations of mfBm: Top row, Mixing Error Ratio (MER) for $\hat{\mathbf{W}}^{(o)}$; bottom row, rMSE for $\underline{H}$. Each color corresponds to a different H_2. For the bottom plot, the different markers correspond respectively to: '+',$\hat{\underline{H}}^{M}_{pre}$; 'o',$\hat{\underline{H}}^{(o)}$; 'v', $\hat{\underline{H}}^{U}_{post}$.

parameters H. To study the impact of differences amongst the values of $\underline{H}$, a second set of Monte Carlo simulations is conducted in bivariate settings, $(H_1 \leq H_2)$, for four values of H_2 and a large range of several different values for H_1. Mixing and covariance matrices are set to $\mathbf{W} = [1, 0.4; 0.5, 1]$ and $\mathbf{\Sigma} = [1, 0.7; 0.7, 1]$.

Estimation performance for $\hat{\mathbf{W}}^{(o)}$ are reported in Fig. 2 (top plot). Estimation performance for $\hat{\underline{H}}^{M}_{pre}$, $\hat{\underline{H}}^{(o)}$,and $\hat{\underline{H}}^{U}_{post}$, are shown in Fig. 2 (bottom plot). For clarity, performance for $\hat{\underline{H}}^{U}_{pre}$ / $\hat{\underline{H}}^{M}_{post}$ are no longer reported as they can only be poorer compared to those of $\hat{\underline{H}}^{M}_{pre}$ / $\hat{\underline{H}}^{U}_{post}$.

Fig. 2 shows, first, that for both $\mathbf{W}$ and $\underline{H}$ estimation performance do not depend on H_2 but only on $\Delta H = H_2 - H_1$. Fig. 2(top plot) shows, second, that the quality of the demixing decreases as $\Delta H \to 0$ (or $H_1 \to H_2$), thus confirming that the more similar the temporal dynamics of the two components are the more difficult the demixing operation becomes. Practically, a MER above 20 is regarded as relevant demixing, which shows that, even with long time series, as is the case here $(n = 2^{18})$, $\Delta H \geq 0.2$ is required to achieved correct demixing.

Fig. 2 (bottom plot) further shows two interesting outcomes. First, irrespective of ΔH, $\hat{\underline{H}}^{(o)}$ outperforms $\hat{\underline{H}}^{M}_{pre}$ in rMSE, indicating that the proposed demixing procedure more accurately account for the full multivariate joint statistical structure of mfBm (through (3)) than the sole eigenvalues (through (7)). This comes however with computational cost larger by two orders of magnitude (from 0.2s on average per realization to 42s). Second, once the mixing matrix is estimated, univariate estimation of H on demixed data, $\hat{\underline{H}}^{U}_{post}$, does not improve the estimation of $\underline{H}$, compared to that of $\hat{\underline{H}}^{(o)}$ obtained jointly with the estimation of the mixing matrix $\hat{\mathbf{W}}^{(o)}$.

Finally, let us note that the rMSE for $\underline{H}$ are almost not varying with ΔH as long as the demixing matrix is correctly estimated (MER $\leq$ 20 and $\Delta H \leq 0.2$). As an apparent paradox, rMSE tend to decrease for MER below 20, i.e., when the mixing matrix is less accurately estimated. However, this is only because H_1 and H_2 are close enough for the distibution of their estimates to overlap. Solving demixing ambiguities by sorting estimates of $\underline{H}$ then decreases rMSE.

D. Performance as function of the conditioning number of the mixing matrix and of the correlation level.

A third set of Monte Carlo simulations aims to investigate the impact of the conditioning of the mixing matrix and of the correlation levels. To that end, a bivariate setting is used with fixed $(H_1, H_2) = (0.6, 0.8)$. The correlation between pre-mixed components is varied as $\rho \in (0, 0.3, 0.5, 0.8)$. The mixing matrix is varied from diagonal (no mixing) to close to non invertible with a reciprocal of its condition number (as computed using the standard norm-1 LAPAC condition estimator) varying from 1 to $\simeq 2.1e^{-2}$.

Estimation performance are reported in Table II. First, Table II shows that estimation performance, both for $\mathbf{W}$ and $\underline{H}$, do not depend on the pre-mixing correlation level. Second, Table II suggests that the proposed demixing procedure performs equally satisfactorily as long as the reciprocal of the condition number of $\mathbf{W}$ remains above $\simeq 5.1e^{-2}$. Third, Table II confirms that $\hat{\underline{H}}^{(o)}$ outperforms $\hat{\underline{H}}^{M}_{pre}$ in rMSE, while $\hat{\underline{H}}^{U}_{post}$, does not further improve the estimation of H, compared to $\hat{\underline{H}}^{(o)}$.

	Reciprocal of the conditioning number ($\log_{10}$)							
ρ	0.0	-0.5	-1.0	-1.3	-1.4	-1.5	-1.6	-1.7
0.0	20.1	24.0	33.8	38.2	26.3	14.6	9.7	10.0
	(10.7)	(13.1)	(10.2)	(15.7)	(22.4)	(18.2)	(12.3)	(12.6)
0.3	20.8	26.1	31.5	34.1	25.3	16.6	9.2	11.6
	(9.1)	(10.7)	(11.0)	(18.4)	(22.0)	(19.5)	(11.7)	(15.6)
0.5	18.8	24.5	32.9	38.4	24.4	15.1	9.5	10.1
	(10.0)	(10.4)	(10.6)	(18.0)	(22.5)	(18.1)	(10.9)	(12.6)
0.8	22.4	23.7	33.9	39.7	27.3	16.3	9.8	10.8
	(10.2)	(10.7)	(11.3)	(14.3)	(22.5)	(19.4)	(12.5)	(14.0)

TABLE II
ESTIMATION PERFORMANCE FOR $\mathbf{W}$ QUANTIFIED BY MIXING ERROR RATIO (MER), AVERAGE (STD) ACROSS 100 REALIZATIONS OF MFBM, AS FUNCTION OF THE LOGARITHM OF THE RECIPROCAL CONDITION NUMBER OF $\mathbf{W}$, FOR FOUR DIFFERENT VALUES OF ρ.

V. FEDERAL RESERVE ECONOMIC DATA

As a pedagogical illustration, the estimation strategy proposed here is applied to a pair of time series extracted from the Federal Reserve Economic Data of monthly frequency macroeconomic data (FRED-MD) [22]. The two time series, chosen here because highly correlated and potentially cointegrated, are shown in Fig. 3 (top plot).

The unmixing procedure is applied using scales $\mathcal{J} = (2, \ldots, 7)$ corresponding to time scales ranging from 2 to 128 months. Table III shows that the univariate analysis applied to these yields two estimates $\hat{\underline{H}}^{U}_{pre}$ close to 0.5 (white noise). These estimates differ from the multivariate ones $\hat{\underline{H}}^{M}_{pre}$ that reveal much lower selfsimilarity parameter values embedded in the joint statistical structure of the pair. The unmixing procedure further confirms the estimation of distinct selfsimilarity parameters $\hat{\underline{H}}^{(o)}$ from the pair of time series.

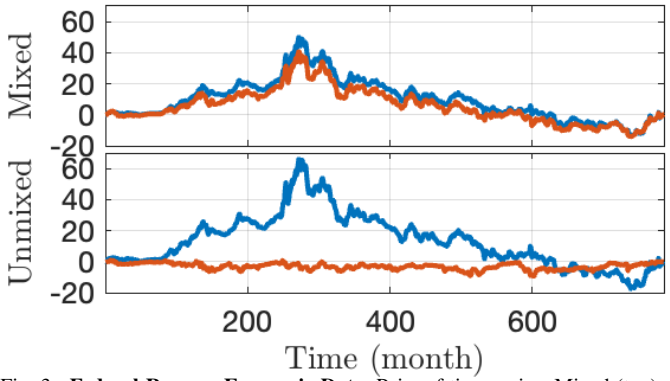

Fig. 3. **Federal Reserve Economic Data.** Pairs of time series: Mixed (top) ; Unmixed (bottom).

From the estimated mixing matrix,

$$\hat{W}_o = \begin{bmatrix} 0.76 & 0.20 \\ 0.65 & 0.98 \end{bmatrix}$$

one can construct demixed time series, $(\hat{\mathbf{W}}^{(o)})^{-1}\dot{Y}$, shown in Fig. 3 (bottom plot). The univariate and multivariate estimates of selfsimilarity parameters applied to these demixed time series $\hat{\underline{H}}^{U}_{post}$ and $\hat{\underline{H}}^{M}_{post}$ are very consistent and in agreement with the estimates obtained from the optimization procedure $\hat{\underline{H}}^{(o)}$, thus validating the relevance of the unmixing operation applied to these two empirical time series, without the knowledge of the ground truth mixing matrix $\mathbf{W}$. The estimated mixing matrix $\hat{W}_o$ further show that one of the observed time series essentially consists of one of the sources, while the other mixes the two sources. This illustrative pilot will be extended to larger sets of multivariate FRED time series in future work.

	Mixed		Demixed		
	$\hat{\underline{H}}^{U}_{pre}$	$\hat{\underline{H}}^{M}_{pre}$	$\hat{\underline{H}}^{(o)}$	$\hat{\underline{H}}^{U}_{post}$	$\hat{\underline{H}}^{M}_{post}$
X_1	0.50	0.29	0.27	0.30	0.27
X_2	0.55	0.54	0.56	0.55	0.55

TABLE III
ESTIMATED SELFSIMILARITY PARAMETERS BEFORE AND AFTER DEMIXING FOR A PAIR OF FRED DATA.

VI. CONCLUSIONS AND PERSPECTIVES

This work shows that multivariate selfsimilarity can be analysed beyond selfsimilarity parameters by estimating a potential mixing matrix between components, and concommitanly further improving the estimation performance for $\underline{H}$ compared to existing techniques. This procedure totally renews the potential of multivariate selfsimilarity in applications and is ready for use on real-world data. A toolbox will be shared at the time of publication.

The studies of estimation performance indicate that the main hypothesis for the current procedure to yield satisfactory performance is that the selfsimilarity parameters in $\underline{H}$ are sufficiently distinct, while performance only depend weakly on correlations and remain very stable with respect to the condition of the mixing matrix. Future works will address these limits by an in-depth use of the joint statistical structure of multivariate selfsimilarity, as studied in [12], [13]. Finally, large dimensional settings where M grows at the same rate as the sample size will be investigated.

REFERENCES

[1] G. Samorodnitsky and M. Taqqu, *Stable non-Gaussian Random Processes: Stochastic Models with Infinite Variance*, Chapman and Hall, New York, 1994.

[2] M. Palmer, "Fractal geometry: a tool for describing spatial patterns of plant communities," *Vegetatio*, vol. 75, pp. 91–102, 1988.

[3] R. Lopes and N. Betrouni, "Fractal and multifractal analysis: A review," *Medical image analysis*, vol. 13, no. 4, pp. 634–649, 2009.

[4] 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.

[5] B. J He, "Scale-free brain activity: past, present, and future," *Trends in cognitive sciences*, vol. 18, no. 9, pp. 480–487, 2014.

[6] 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.

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

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

[9] P. Amblard and J. Coeurjolly, "Identification of the multivariate fractional Brownian motion," *IEEE Transactions on Signal Processing*, vol. 59, no. 11, pp. 5152–5168, 2011.

[10] 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, 2017.

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

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

[13] C. Lucas, G. Didier, H. Wendt, and P. Abry, "Multivariate selfsimilarity: Multiscale eigen-structures for selfsimilarity parameter estimation," *IEEE Trans. Signal Process.*, vol. 72, pp. 1739–1749, 2024.

[14] G. Didier and V. Pipiras, "Exponents, symmetry groups and classification of operator fractional Brownian motions," *Journal of Theoretical Probability*, vol. 25, no. 2, pp. 353–395, 2012.

[15] B. Mandelbrot and J. Van Ness, "Fractional Brownian Motions, Fractional Noises and Applications," *SIAM Review*, vol. 10, no. 4, pp. 422–437, October 1968.

[16] 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.

[17] J. Frecon, G. Didier, N. Pustelnik, and P. Abry, "Non-linear wavelet regression and branch & bound optimization for the full identification of bivariate operator fractional brownian motion," *IEEE Transactions on Signal Processing*, vol. 64, no. 15, pp. 4040–4049, 2016.

[18] Z. Gao, Y. Wu, Y. Jia, and M. Harandi, "Learning to optimize on spd manifolds," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020, pp. 7700–7709.

[19] R. Fletcher, *Practical methods of optimization*, John Wiley & Sons, 2013.

[20] 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.

[21] E. Vincent, S. Araki, and P. Bofill, "The 2008 signal separation evaluation campaign: A community-based approach to large-scale evaluation," in *International Conference on Independent Component Analysis and Signal Separation*. Springer, 2009, pp. 734–741.

[22] M. W. McCracken and S. Ng, "Fred-md: A monthly database for macroeconomic research," *Journal of Business & Economic Statistics*, vol. 34, no. 4, pp. 574–589, 2016.