

Wavelet-based demixing for multivariate fractional Brownian motion

Kinan Abbas¹, Herwig Wendt², Gustavo Didier³ and Patrice Abry¹

¹CNRS, ENS de Lyon, Laboratoire de Physique, Lyon, France; {kinan.abbas, patrice.abry}@ens-lyon.fr

²CNRS, IRIT, University of Toulouse, Toulouse, France; herwig.wendt@irit.fr

³Mathematics Department, Tulane University, New Orleans, LA, USA; gdidier@tulane.edu

Multivariate fractional Brownian motion MFBM

- **Selfsimilarity**: scale-free dynamics governed by a broad continuum of scales.
- **Multivariate selfsimilarity**: M-components

$$\mathbf{X}(t) \stackrel{fdd}{=} a^H \mathbf{X}(t/a), \quad \underline{H}: \text{Matrix of selfsimilarity exponents} \quad (1)$$

- **Multivariate fractional Brownian motion** [1]:

$$B_{\underline{H}}(t) = [B_{\underline{H}, H_1}(t), \dots, B_{\underline{H}, H_M}(t)], \quad \underline{H} = (H_1, \dots, H_M), \quad 0 < H_m < 1.$$

The components are correlated through the point covariance matrix

$$\underline{\Sigma} = \text{diag}(\sigma_1^2, \dots, \sigma_M^2) \rho \text{diag}(\sigma_1^2, \dots, \sigma_M^2)^*, \quad (2)$$

where ρ is a semi-positive definite matrix of pairwise correlations.

$\mathbf{W}$ invertible mixing matrix, $\underline{H}$ vector of selfsimilarity exponent, $\underline{\Sigma}$ correlation matrix.

$$B_{\mathbf{W}, \underline{\Sigma}, \underline{H}}(t) := \mathbf{W} B_{\underline{\Sigma}, \underline{H}}(t)^\top. \quad (3)$$

$$\Rightarrow \underline{H} = \mathbf{W} \text{diag}(H_1, \dots, H_M) \mathbf{W}^\top. \quad (4)$$

- **Goal**: Jointly estimate $\underline{H}$, $\mathbf{W}$, $\underline{\Sigma}$,

Analysis: Multivariate wavelet transform

- Let ψ_0 be a compactly supported mother wavelet with N_ψ vanishing moments. For an M -variate time series Y , the multivariate wavelet coefficients are

$$D_Y(2^j, k) = \int_{\mathbb{R}} \psi(j, k) Y(t) dt, \quad \psi(j, k) = 2^{-j/2} \psi_0(2^{-j}t - k). \quad (5)$$

- The **multivariate wavelet spectrum** is the collection of scale-dependent covariance matrices [2, 1]:

$$\mathbf{S}_Y(2^j) = \frac{1}{n_j} \sum_{k=1}^{n_j} D_Y(2^j, k) D_Y(2^j, k)^\top. \quad (6)$$

- For mfBm, wavelet coefficients reproduce selfsimilarity across scales:

$$\left\{ D_{B_{\mathbf{W}, \underline{\Sigma}, \underline{H}}}(2^j, k) \right\}_{k \in \mathbb{Z}} \stackrel{fdd}{=} \left\{ \mathbf{W} 2^{j(\underline{H} + \frac{1}{2})} D_{B_{\underline{\Sigma}, \underline{H}}}(2^0, k) \right\}_{k \in \mathbb{Z}}. \quad (7)$$

Key wavelet scaling relation

$$\mathbf{S}_{B_{\mathbf{W}, \underline{\Sigma}, \underline{H}}}(2^j) \stackrel{d}{=} \mathbf{W} 2^{j(\underline{H} + \frac{1}{2})} \mathbf{S}_{B_{\underline{\Sigma}, \underline{H}}}(2^0) 2^{j(\underline{H} + \frac{1}{2})} \mathbf{W}^\top. \quad (8)$$

Without mixing: power law $\rightarrow$ univariate estimator

$$\mathbb{E}(S_{m,m}(2^j)) = \beta_m 2^{j(2H_m+1)} \Rightarrow \hat{H}_m^U = \frac{1}{2} \left(\sum_{j=j_1}^{j_2} w_j \log_2 S_{m,m}(2^j) - 1 \right). \quad (9)$$

With mixing: mixtures $\rightarrow$ eigenvalues $\rightarrow$ multivariate estimator

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

$$\mathbf{S}_Y(2^j) = \mathbf{U}_j \text{diag}(\lambda_1(2^j), \dots, \lambda_M(2^j)) \mathbf{U}_j^\top \Rightarrow \hat{H}_m^M = \frac{1}{2} \left(\sum_{j=j_1}^{j_2} w_j \log_2 \lambda_m(2^j) - 1 \right). \quad (11)$$

Demixing: optimization problem

- The proposed approach estimates the **selfsimilarity parameters**, the **mixing matrix**, and the scale-normalized covariance structure by fitting the multivariate wavelet spectra across scales.

$$\left(\hat{\underline{H}}^{(o)}, \hat{\mathbf{W}}^{(o)}, \hat{\underline{\Sigma}}^{(o)} \right) = \underset{\underline{H}, \mathbf{W}, \underline{\Sigma}}{\text{argmin}} \mathcal{L}(\underline{H}, \mathbf{W}, \underline{\Sigma}). \quad (12)$$

- The loss compares the observed wavelet spectrum to the mfBm model at each scale $j \in \mathcal{J}$:

Log-Euclidean multiscale loss

$$\mathcal{L}(\underline{H}, \mathbf{W}, \underline{\Sigma}) = \frac{1}{|\mathcal{J}|} \sum_{j \in \mathcal{J}} \left\| \log \mathbf{S}_Y(2^j) - \log \left(\mathbf{W} 2^{j(\underline{H} + \frac{1}{2})} \mathbf{S}_{B_{\underline{\Sigma}, \underline{H}}}(2^0) 2^{j(\underline{H} + \frac{1}{2})} \mathbf{W}^\top \right) \right\|_F^2. \quad (13)$$

- For $\mathbf{S} = \mathbf{U} \text{diag}(\lambda_1, \dots, \lambda_M) \mathbf{U}^\top$, **matrix logarithm** is $\log \mathbf{S} = \mathbf{U} \text{diag}(\log \lambda_1, \dots, \log \lambda_M) \mathbf{U}^\top$.
- The **matrix logarithm** turns algebraic growths into linear growths, so that all scales contribute to estimation.
- The minimization is performed jointly over $\underline{H}$, $\mathbf{W}$, $\underline{\Sigma} = \mathbf{S}_{B_{\underline{\Sigma}, \underline{H}}}(2^0)$, using a Quasi-Newton method with BFGS updates and cubic line search, as implemented by MATLAB `fminunc`.

Experiment setup

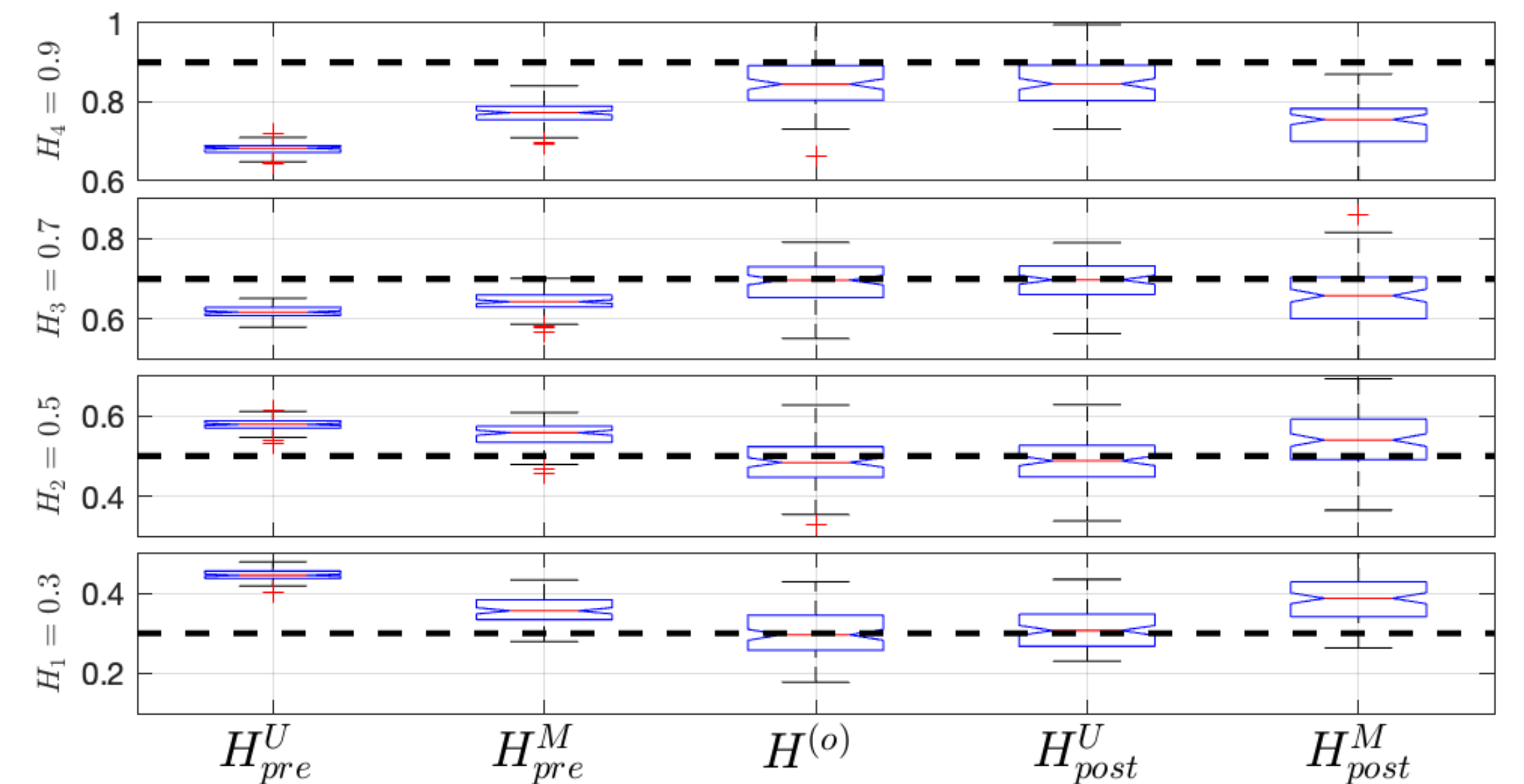
- Estimation performance is quantified by Monte Carlo simulations with $N = 100$ realizations of mfBm sample paths of length $n = 2^{18}$, for several configurations of $(\underline{H}, \underline{\Sigma}, \mathbf{W})$.
- Wavelet analysis is performed using a **Daubechies-3 least asymmetric orthonormal wavelet**, with scales $\mathcal{J} = (3 : 15)$.
- The optimization is initialized as follows: $\underline{H}$ from eigenvalue-based regression, $\mathbf{W}$ from eigenvectors of $\mathbf{S}_Y(2^j)$ at the largest scale, and $\mathbf{S}_{B_{\underline{\Sigma}, \underline{H}}}(2^0)$ as diagonal.
- The quality of $\hat{\mathbf{W}}$ is measured with **Mixing Error Ratio (MER)**; the quality of $\hat{\underline{H}}$ is measured by **bias**, **variance**, and **rMSE**.

Results: Monte Carlo simulations

- **Notation**. $\hat{H}_{pre}^U$ and $\hat{H}_{pre}^M$ are obtained from the original mixed data Y . $\hat{H}_{post}^U$ and $\hat{H}_{post}^M$ are obtained from the demixed data

$$\tilde{Y} = (\hat{\mathbf{W}}^{(o)})^{-1} Y.$$

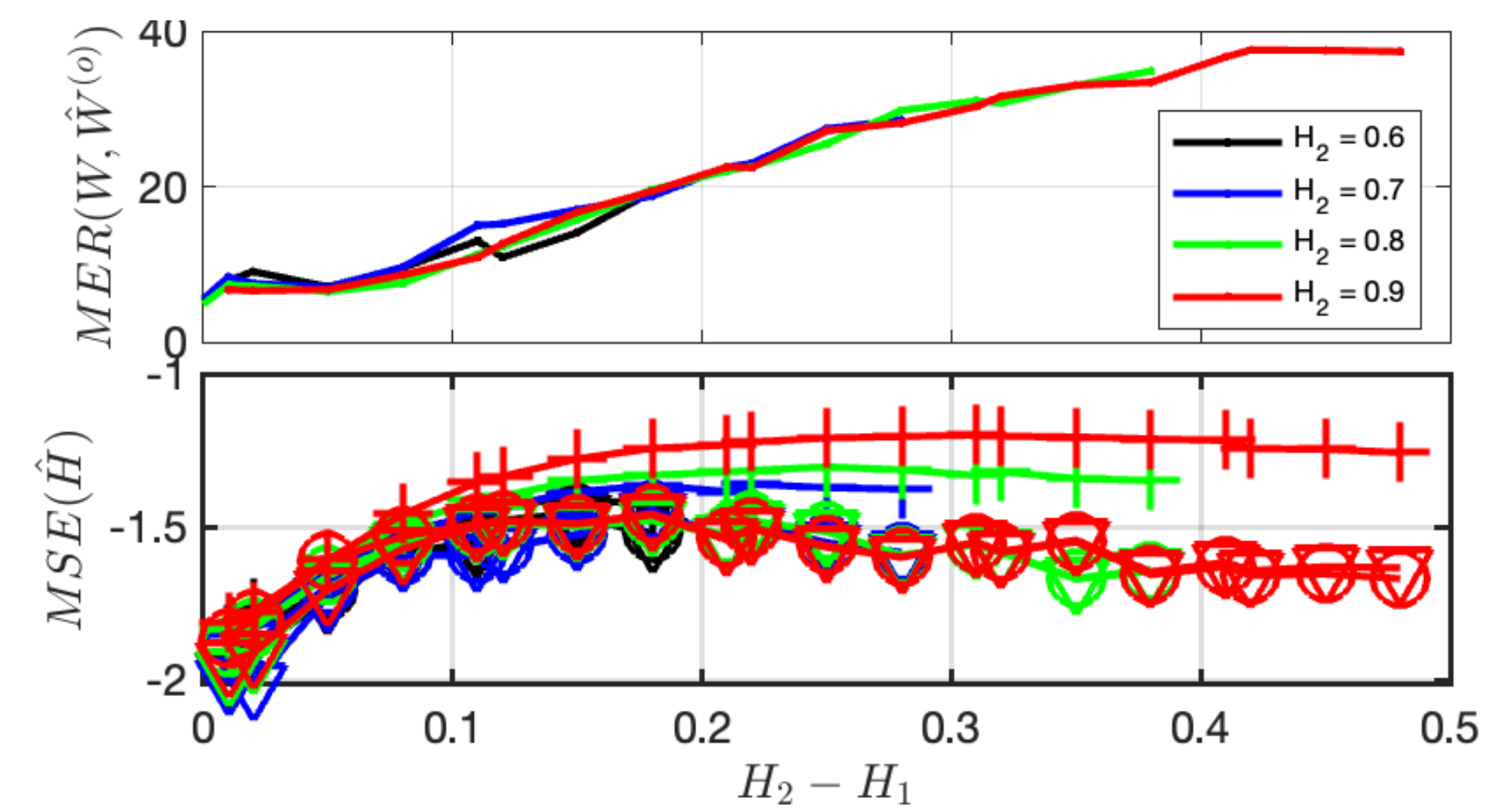
- **4-variate proof of concept**. $\underline{H} = (0.3, 0.5, 0.7, 0.9)$, a non-diagonal $\underline{\Sigma}$, and a mixing matrix $\mathbf{W}$ with high condition number, making the demixing task more demanding.



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

- **Effect of differences in selfsimilarity**. Performance is studied as a function of $\Delta_H = H_2 - H_1$, with

$$\mathbf{W} = \begin{bmatrix} 1 & 0.4 \\ 0.5 & 1 \end{bmatrix} \text{ and } \underline{\Sigma} = \begin{bmatrix} 1 & 0.7 \\ 0.7 & 1 \end{bmatrix}.$$

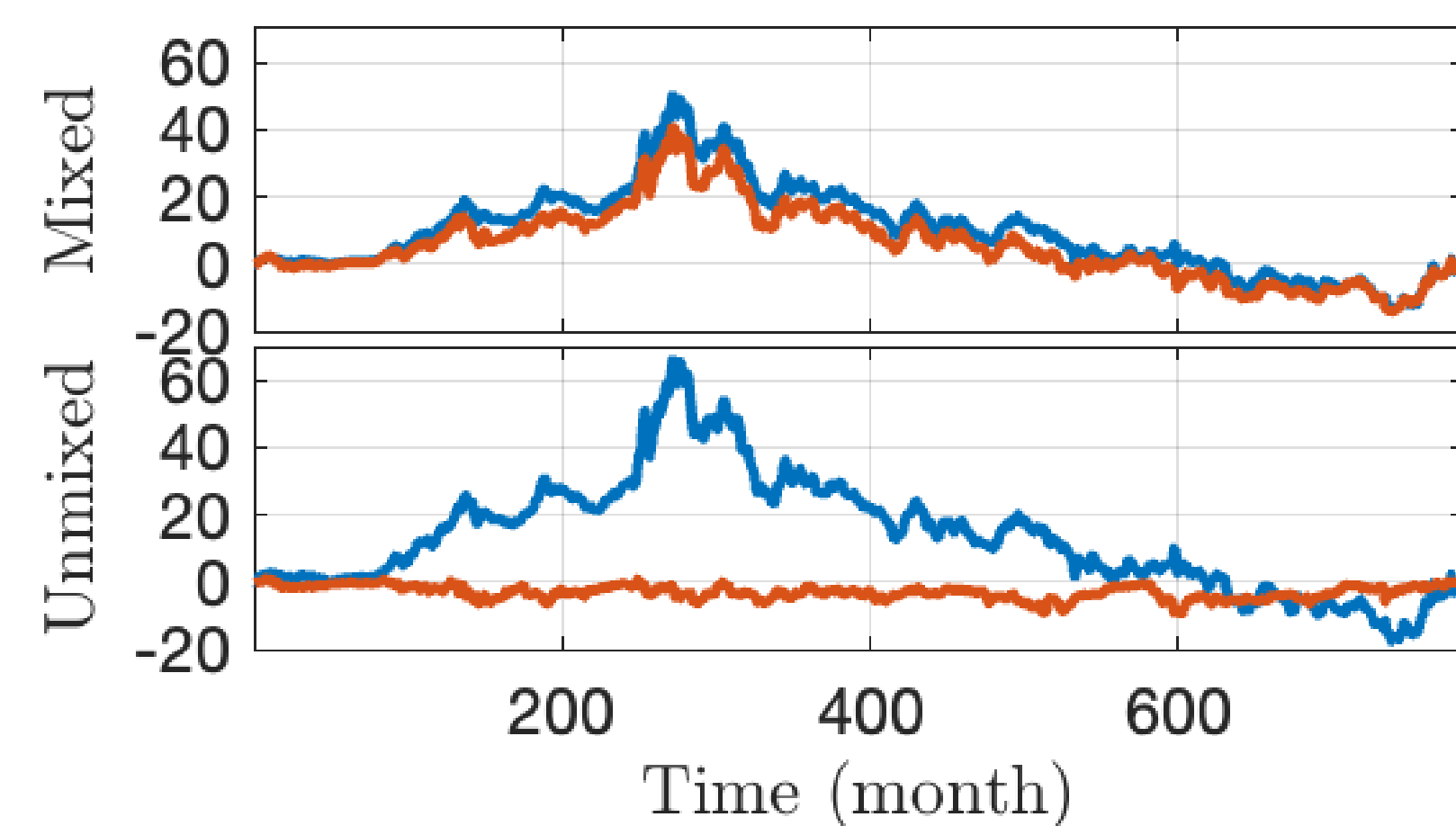


MER for $\hat{\mathbf{W}}^{(o)}$ and rMSE for $\hat{\underline{H}}$ as functions of Δ_H .

- **Conditioning of $\mathbf{W}$ and correlation level**. A bivariate setting uses fixed $(H_1, H_2) = (0.6, 0.8)$, with $\rho \in (0, 0.3, 0.5, 0.8)$ and $\mathbf{W}$ ranging from diagonal to close to non-invertible.

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

- **Federal Reserve Economic Data**. The estimation strategy is applied to a pair of monthly macroeconomic time series from FRED-MD [3]. The unmixing procedure is applied using $\mathcal{J} = (2, \dots, 7)$, corresponding to time scales from 2 to 128 months.



Pairs of time series: mixed (top) and unmixed (bottom).

From the estimated mixing matrix,

$$\hat{\mathbf{W}}^{(o)} = \begin{bmatrix} 0.76 & 0.20 \\ 0.65 & 0.98 \end{bmatrix},$$

one can construct demixed time series.

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

Estimated selfsimilarity parameters before and after demixing.

Conclusion and Future Work

- The proposed wavelet-based demixing procedure jointly estimates $\underline{H}$, $\mathbf{W}$, and $\underline{\Sigma}$.
- Monte Carlo simulations show improved estimation of $\underline{H}$ and effective recovery of the mixing matrix $\mathbf{W}$.
- The main limitation is the need for sufficiently distinct selfsimilarity parameters.
- Future work will address larger dimensional settings and further exploit the joint statistical structure of multivariate selfsimilarity.

References

- [1] C. Lucas, G. Didier, H. Wendt, and P. Abry. "Multivariate selfsimilarity: Multiscale eigen-structures for selfsimilarity parameter estimation". In: *IEEE Trans. Signal Process.* 72 (2024), pp. 1739–1749.
- [2] P. Abry and G. Didier. "Wavelet eigenvalue regression for n-variate operator fractional Brownian motion". In: *Journal of Multivariate Analysis* 168 (2018), pp. 75–104.
- [3] M. W. McCracken and S. Ng. "FRED-MD: A monthly database for macroeconomic research". In: *Journal of Business & Economic Statistics* 34.4 (2016), pp. 574–589.

