

# WAVELET-BASED DETECTION AND ESTIMATION OF FRACTIONAL LÉVY SIGNALS IN HIGH DIMENSIONS

B. Cooper Boniece<sup>1</sup>, Herwig Wendt<sup>2</sup>, Gustavo Didier<sup>3</sup>, Patrice Abry<sup>4</sup>,

<sup>1</sup> Dept. Math. and Stats., Washington University in St. Louis, St. Louis, USA.

<sup>2</sup> IRIT, CNRS (UMR 5505), Université de Toulouse, France.

<sup>3</sup> Math. Dept., Tulane University, New Orleans, USA.

<sup>4</sup> Univ Lyon, ENS de Lyon, Univ Claude Bernard, CNRS, Laboratoire de Physique, Lyon, France.

## ABSTRACT

In the modern world, systems are routinely monitored by multiple sensors, generating “Big Data” in the form of a large collection of time series. However, dynamic signals are often low-dimensional and characterized by joint scale-free dynamics (self-similarity) and non-Gaussianity. In this paper, we put forward a statistical methodology for identifying the number of multivariate self-similar, Lévy-driven components immersed in high-dimensional noise, as well as for estimating the underlying scaling exponents. It relies on the analysis of the evolution over scales of the eigenvalues of random wavelet matrices. Monte Carlo simulations show that the proposed methodology is accurate for realistic sample sizes. This holds even at low signal-to-noise ratios and for a large number of observed mixed and noisy time series. The mathematical framework further allows us to analyze the impact of the tails of the Lévy noise marginal distribution on the estimation performance.

**Index Terms**— multivariate self-similarity, operator fractional Lévy motion, wavelets, random matrices, high dimensions

## 1. INTRODUCTION

**Context: many sensors, few sources.** In the modern world of “Big Data,” a plethora of sensors monitor natural and artificial systems, generating large data sets in the form of several joint time series. This can be seen in many fields of application. In neuroscience, for example, the number of macroscopic brain activity time series ranges from hundreds (MEG data) to several tens of thousands (fMRI data) [1]. Likewise, in climate studies, dealing with large numbers of measured components has become standard [2]. Nevertheless, the number of physical mechanisms, or sources, driving the spatio-temporal dynamics of a system is usually far smaller than the number of recorded time series (sparsity in high dimensions). This has created a strong demand for data analytic methods that can statistically identify sources amidst noisy measurements and extract their informational content [3].

**(Covariance) self-similarity and non-Gaussian behavior.** A signal is called *scale invariant* when its temporal dynamics lack a characteristic scale. In the presence of scale invariance, the analyst’s main interest is in identifying mechanisms, often in the form of scaling exponents, that relate the continuum of scales [4]. Scale

invariance often manifests itself in physics and engineering in the form of *self-similarity*. A signal  $X$  is called self-similar when its finite-dimensional distributions (fdd) are invariant under suitable time scaling, i.e.,  $\{X(t)\}_{t \in \mathbb{R}} \stackrel{\text{fdd}}{=} \{a^H X(t/a)\}_{t \in \mathbb{R}}$ ,  $a > 0$ , where  $0 < H < 1$  is called the Hurst parameter. A cornerstone model of scale invariance, fractional Brownian motion (fBm) is the only Gaussian, self-similar process with stationary increments [5]. Estimation of the Hurst parameter  $H$  plays a key role in signal processing tasks such as characterization, diagnosis, classification and detection. The wavelet transform underpins efficient multiscale representations of self-similar or fractional systems, and it provides the analytical basis for well-established methodologies for the estimation of  $H$  [6].

On the other hand, non-Gaussian behavior is pervasive in a myriad of natural and artificial systems displaying scale invariance (e.g., turbulence [7] or Internet traffic [8]). Due to its breadth and flexibility, the family of fractional Lévy processes [9–11] has become popular in physical applications involving non-Gaussianity [12, 13]. Nevertheless, while of great importance in practice, the modeling of multivariate fractional Lévy signals is a research topic that has been relatively little explored [14, 15]. In [16], operator fractional Lévy motion (ofLm), a generalization of the univariate fBm and its multivariate counterpart, operator fBm (ofBm) [17, 18], is put forward as a model for non-Gaussian multivariate (covariance) self-similarity. In particular, the ofLm family comprises instances where there are multiple correlated fractional Lévy coordinate processes with possibly distinct self-similarity exponents  $H_m$ ,  $m = 1, \dots, M$ , occurring in a non-canonical set of coordinates (mixing).

**Related work.** In the statistical signal processing literature, the problem of statistically identifying the number and properties of sources in multivariate (fixed dimensions) or high-dimensional noisy signals has been studied for at least three decades now [19–22]. Examples of the proposed techniques include principal component analysis, factor analysis and sparse graphical Gaussian models [23]. Nevertheless, there has been a paucity of estimation methodologies for both *high-dimensional* and *scale invariant* signals [24]. To address this issue, the so-named wavelet eigenanalysis estimation methodology was recently constructed [25, 26]. It was shown to lead to efficient and robust estimation of self-similarity exponents in both Gaussian multivariate/high-dimensional [26, 27] and non-Gaussian multivariate [16] settings.

**Goal, contributions and outline.** In this paper, we tackle the problem of high-dimensional inference for non-Gaussian covariance self-similar signals. The observations are assumed to appear in the form of an ofLm plus high-dimensional noise. The concept of multi-

Work supported by Grant ANR-16-CE33-0020 MultiFracs. G.D. was partially supported by the ARO grant W911NF-14-1-0475. G.D.’s long term visits to France were supported by ENS de Lyon, CNRS and the Carol Lavin Bernick faculty grant.

variate (covariance) self-similarity is briefly laid out at the beginning of Section 2. The definition and properties of ofLm are recapped in Section 2.1. The recently proposed wavelet eigenanalysis methodology is detailed in Section 2.2. The high-dimensional model and model order selection procedure are described in Section 3. Performance with respect to non-Gaussianity and increasing dimension of both model order selection and Hurst exponent estimation are assessed by means of Monte Carlo studies and discussed in Section 4.

## 2. OPERATOR FRACTIONAL LÉVY MOTION AND WAVELET EIGENANALYSIS-BASED ESTIMATION

**Covariance operator self-similarity.** Let  $\underline{H} = P \text{diag}(\underline{H}) P^{-1}$  be the so-named Hurst matrix parameter, where the vector  $\underline{H}$  contains the Hurst exponents (eigenvalues). We say a stochastic process  $X = (X_1, \dots, X_M)$  is covariance (multivariate) operator self-similar (c.o.s.s.) with Hurst matrix parameter  $\underline{H}$  when

$$\mathbb{E}X(s)X(t)^* = a^{\underline{H}} \mathbb{E}X(s/a)X(t/a)^* a^{\underline{H}^*} \quad (1)$$

for all  $a > 0$ , where  $a^{\underline{H}} := \sum_{k=0}^{+\infty} \log^k(a) \underline{H}^k / k!$ . When the mixing matrix  $P$  is diagonal, namely, when we can set  $P \equiv I$ , the covariance self-similarity relation (1) takes the simple form of component-wise covariance self-similarity relations

$$\mathbb{E}X_\ell(s)X_{\ell'}(t) = a^{H_\ell + H_{\ell'}} \mathbb{E}X_\ell(s/a)X_{\ell'}(t/a) \quad (2)$$

for  $a > 0$  and  $\ell, \ell' = 1, 2, \dots, M$ .

### 2.1. Operator fractional Lévy motion

OfLm is a class of non-Gaussian multivariate fractional processes displaying the same covariance structure of ofBm [16]. In particular, ofLm is c.o.s.s. For clarity of exposition and without loss of generality, we construct ofLm in the bivariate case. Let  $\{L(t) = (L_1(t), L_2(t))\}_{t \in \mathbb{R}}$  be a two-sided symmetric Lévy process in  $\mathbb{R}^2$  with  $\mathbb{E}L(1)L(1)^* =: \Sigma_L, |\Sigma_L| < \infty$  (e.g., [28]). Let  $0 < H_1 \leq H_2 < 1$ , and define the (pre-mixed) process  $\mathcal{X}$  by the component-wise convolution  $\mathcal{X}(t) = (g_t * \dot{L})(t)$ , where  $g$  is the diagonal-matrix-valued fractional kernel  $g_t(u) := u_+^D - (u-t)_+^D$ ,  $D = \text{diag}(\underline{H}) - \frac{1}{2}I$ . Then, the entrywise processes  $\mathcal{X}_{H_\ell}$  ( $\neq X$ ),  $\ell = 1, 2$ , are generally *correlated* fractional Lévy processes with corresponding Hurst parameters  $H_\ell \in (0, 1)$  (e.g., [9, 10]). In particular, each has stationary increments and its covariance function is identical to that of fBm, i.e.,

$$\mathbb{E}\mathcal{X}_{H_\ell}(t)\mathcal{X}_{H_\ell}(s) = \{|t|^{2H_\ell} + |s|^{2H_\ell} - |t-s|^{2H_\ell}\} \sigma_\ell^2 / 2. \quad (3)$$

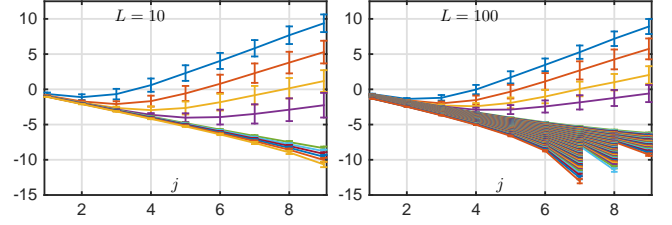
Note that, if non-Gaussian, the stochastic behavior of the process  $\mathcal{X}_{H_\ell}$  is *not* characterized by its covariance function. Let  $P$  be a  $2 \times 2$ , real-valued, invertible matrix. We define the (bivariate) ofLm  $X^{\underline{H}, L, P}$  by

$$\{X_1^{\underline{H}, L, P}(t), X_2^{\underline{H}, L, P}(t)\}_{t \in \mathbb{R}} = P\{\mathcal{X}_{H_1}(t), \mathcal{X}_{H_2}(t)\}_{t \in \mathbb{R}}$$

(in short,  $X^{\underline{H}, L, P} = P\mathcal{X}$ ). Note that the parameters  $\underline{H}$  and  $\rho := \text{Corr}(\mathcal{X}_{H_1}(1), \mathcal{X}_{H_2}(1))$  cannot be selected independently (see [16] for more details).

### 2.2. Wavelet-based scaling exponent estimation

**Multivariate wavelet transform.** Let  $\psi_0$  be a mother wavelet, i.e., an oscillating function that is well localized in both time and frequency [29]. A mother wavelet is parametrized by the so-named



**Fig. 1. Log-eigenvalues of  $S_Y(2^j)$ .** For  $\gamma = 10^{-5}$  and  $\text{SNR}=0\text{dB}$ , plots of  $\log_2 \Lambda_Y(2^j)$  for  $M = 4$  and  $L = 10, 100$  (left and right plots, respectively).

number of vanishing moments  $N_\psi$ , i.e., a positive integer such that, for  $\kappa = 0, \dots, N_\psi - 1$ ,  $\int_{\mathbb{R}} t^\kappa \psi_0(t) dt \equiv 0$  and  $\int_{\mathbb{R}} t^{N_\psi} \psi_0(t) dt \neq 0$ . Let  $\{\psi_{j,k}(t) = 2^{-j/2} \psi_0(2^{-j}t - k)\}_{(j,k) \in \mathbb{Z}^2}$  be the collection of dilated and translated templates of  $\psi_0$  that forms an orthonormal basis of  $\mathcal{L}^2(\mathbb{R})$ . For the  $M$ -variate process  $X = \{X(t)\}_{t \in \mathbb{R}}$  and its components, let  $D_{X_m}(2^j, k) := \langle 2^{-j/2} \psi(2^{-j}t - k) | X_m(t) \rangle$ ,  $m \in \{1, \dots, M\}$ . The multivariate DWT of  $X$  is defined as  $D_X(2^j, k) := (D_{X_1}(2^j, k), \dots, D_{X_M}(2^j, k))^*$ ,  $\forall k \in \mathbb{Z}, \forall j \in \mathbb{Z}$ .

**Covariance operator self-similarity in the wavelet domain.** It can be shown that the wavelet coefficients  $\{D_X(2^j, k)\}_{k \in \mathbb{Z}}$  satisfy the covariance (operator) self-similarity relation [30]

$$\begin{aligned} \mathbb{E}D_X(2^j, k)D_X(2^j, k)^* \\ = 2^{j(\underline{H} + \frac{1}{2}I)} \mathbb{E}D_X(1, k)D_X(1, k)^* 2^{j(\underline{H}^* + \frac{1}{2}I)}, \end{aligned} \quad (4)$$

for every fixed octave  $j$  and any shift  $k$ . When  $P = I$  (no mixing), similar to (2),  $M$ -variate covariance (operator) self-similarity in the wavelet domain boils down to the entrywise covariance self-similarity relations, for  $\ell, \ell' = 1, \dots, M$ ,

$$\begin{aligned} \mathbb{E}D_X(2^j, k)D_X(2^j, k)_{\ell\ell'}^* \\ = 2^{j(H_\ell + H_{\ell'} + 1)} \mathbb{E}D_X(1, k)D_X(1, k)_{\ell\ell'}^*. \end{aligned} \quad (5)$$

**Estimation of  $\underline{H}$ .** The sample wavelet spectrum (variance) is given by the  $M \times M$  random matrices

$$S_X(2^j) = \frac{1}{n_j} \sum_{k=1}^{n_j} D_X(2^j, k)D_X(2^j, k)^*, \quad j = j_1, \dots, j_2, \quad (6)$$

where  $n$  is the time series sample size and  $n_j = n/2^j$ . When there is no mixing ( $P = I$ ), i.e., when relation (5) holds, the scaling exponents may be estimated by means of entrywise multiscale log-regressions based on the system (6) (see [31]). Otherwise, i.e., starting from (4) (mixing, or non-diagonal  $P$ ), this procedure does not yield meaningful results (see [25, 26] for further discussions).

An estimation methodology for general (non-diagonal)  $P$  can be constructed as follows. Consider the eigenvalues  $\Lambda_X(2^j) = (\lambda_1(2^j), \dots, \lambda_M(2^j))^*$  of the random matrix  $S_X(2^j)$ . The estimators  $\hat{\underline{H}} = (\hat{H}_1, \dots, \hat{H}_M)$  for  $(H_1, \dots, H_M)$  are defined by means of the multiscale weighted linear regressions

$$\hat{H}_m = \left( \sum_{j=j_1}^{j_2} w_j \log_2 \lambda_m(2^j) \right) / 2 - \frac{1}{2}, \quad \forall m. \quad (7)$$

It was shown mathematically or based on Monte Carlo experiments that  $\hat{\underline{H}}$  has good performance over large and finite samples in multivariate/high-dimensional Gaussian frameworks [25–27], as well as in multivariate non-Gaussian (Lévy-driven) frameworks [16].

### 3. ESTIMATION IN HIGH DIMENSIONS

#### 3.1. High-dimensional model

Let  $W$  be a full rank  $L \times M$  matrix, where  $L \gg M$ . Also let  $X = \{X(t)\}_{t \in \mathbb{R}}$  be a  $M$ -variate ofLm with Hurst matrix  $\underline{H} = P \text{diag}(\underline{H}) P^{-1}$ , as described in Section 2. We assume the  $L$  observed time series follow the model

$$Y(t) := WX(t) + \sigma_{\mathcal{N}} \mathcal{N}(t). \quad (8)$$

In (8), the noise term  $\mathcal{N}(t) = (\mathcal{N}_1(t), \dots, \mathcal{N}_L(t))_{t \in \mathbb{R}}^*$  is made up of  $L$  independent vectors, each consisting of i.i.d. standardized Gaussian variables, and  $\sigma_{\mathcal{N}}^2$  is a control parameter for the noise variance. Assume, without loss of generality, that  $WP$  has rows with unit norm, and further suppose  $\sigma_{\mathcal{X}} := \sigma_1 = \dots = \sigma_M$  in (3). The Signal-to-Noise Ratio (SNR) is defined as

$$\text{SNR} := \frac{\sigma_{\mathcal{X}}}{\sigma_{\mathcal{N}}}. \quad (9)$$

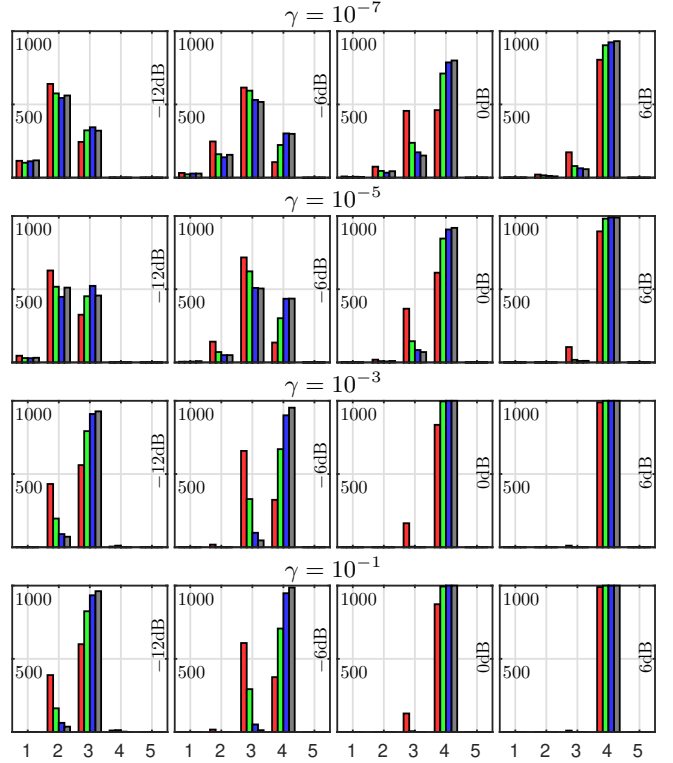
This definition is natural for scaling processes because, for a given value of the SNR, it sets the size of eigenvalues at fine scales  $j \rightarrow 0$ , regardless of the values of  $M$  and  $L$ ; see Fig. 1. Note that it is much more penalizing than the conventional definition of SNR because (9) implies that, for a fixed SNR value, the overall energy of noise in the observations increases linearly with  $L$ .

We assume the Lévy process driving the ofLm has a symmetric exponentially tempered  $\alpha$ -stable distribution (e.g. [32]). Such distributions have moments of all orders and are parametrized by a stability index  $\alpha \in (0, 2)$  and a tempering parameter  $\gamma > 0$ . A smaller tempering parameter  $\gamma > 0$  corresponds to heavier-tailed behavior. Furthermore, it can be shown [30] that at fine timescales the corresponding ofLm is approximately operator stable (heavy-tailed), whereas on coarse timescales it is close to an ofBm (i.e., Gaussian), with transition controlled by  $\gamma$ . Accordingly, for this class of ofLm, we treat  $\gamma$  as a proxy for measuring departures from Gaussianity.

#### 3.2. Model order selection procedure

The proposed procedure relies on the eigenstructure of the random matrices  $S_Y(2^j)$  over several octaves  $j$ , i.e., on the sample wavelet spectra of the  $L$ -variate observations  $Y$ . In the absence of noise (SNR =  $+\infty$ ),  $S_Y(2^j)$  has rank  $M$  for all scales  $2^j$  (as long as  $n_j > M$ ). Conversely, when noise is added (SNR <  $+\infty$ ),  $S_Y(2^j)$  has rank  $L$ . Hence, assuming the SNR is sufficiently large, heuristically speaking each random matrix  $S_Y(j)$  possesses  $L - M$  small eigenvalues corresponding to noise and  $M$  large eigenvalues corresponding to the hidden ofLm components. The former set of eigenvalues manifests itself in the form of log-eigenvalue curves with *negative* slopes, whereas the latter leads to log-eigenvalue curves with *positive* slopes driven by the exponents  $H_m > 0$ . This is illustrated in Fig. 1, which shows averages and error bars for log-eigenvalues generated based on 1000 independent realizations of ofLm (see Sec. 4 for details on numerical simulations). So, let  $\Lambda_Y(2^j) \in \mathbb{R}^L$  be the vector of eigenvalues of  $S_Y(2^j)$ . We propose to estimate the number of components  $M$  as the number of estimated slopes that are found to be positive, where each slope estimate is obtained as in (7) by log-linear regressions on the entries of  $\Lambda_Y(2^j)$ .

Once the number of ofLm components  $\widehat{M}$  has been estimated, a vector of estimated scaling exponents  $(\widehat{H}_1, \dots, \widehat{H}_{\widehat{M}})$  stems naturally from the procedure given by (7), applied to the  $\widehat{M}$  largest values of  $\Lambda_Y(2^j)$ .

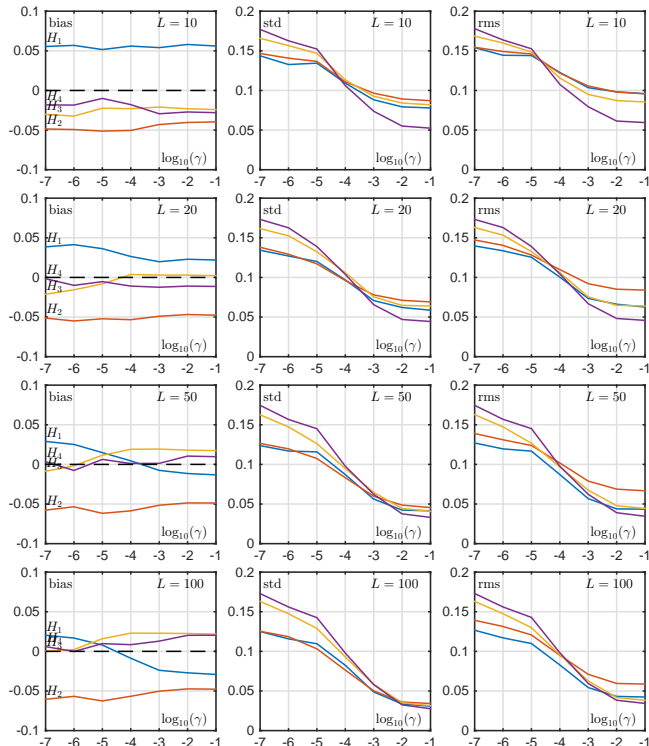


**Fig. 2. Model order selection performance.** Histograms of selected model orders  $\widehat{M}$  for different SNR (increasing from left to right columns) and  $\gamma$  (decreasing departure from Gaussian from top to bottom row), respectively, and  $M = 4$ . The red/green/blue/grey bars correspond to  $L = 10/20/50/100$ , respectively.

### 4. PERFORMANCE

**Monte Carlo simulation.** We evaluate the performance of the proposed model order detection and high-dimensional estimation procedure for 1000 independent realizations of size  $n = 2^{14}$  of ofLm with  $M = 4$  components and exponents  $\underline{H} = (0.4, 0.6, 0.75, 0.9)$ . We study different degrees of deviation from Gaussian behavior (tempering parameter  $\gamma = (10^{-7}, 10^{-6}, \dots, 10^{-1})$ ) with  $\alpha = 1.1$  fixed and  $L = (10, 20, 50, 100)$  observed time series, for various SNR values. The mixing matrix is set to  $P = ((111 \ -1)^T, (1 \ -111)^T, (1111)^T, (1 \ -1 \ -11)^T)$ . The matrix  $W$  is picked randomly and is normalized such that the rows of  $WP$  have unit norm. Estimation of the scaling exponents  $H_m$ ,  $m = 1, \dots, M$ , is performed as in (7) by means of weighted linear regressions for  $j \in (4, j_{\max}^{(m)})$ , where  $j_{\max}^{(m)}$  is the largest scale for which  $\lambda_m(j) > 0$ ,  $\forall m$ . The ofLm paths were generated via an adaptation of the FFT-based algorithm described in [33]. In the notation of [33], the mesh size was chosen  $m = 16$  with kernel cutoff parameter  $M = 2^{14}$ . The symmetric tempered stable processes driving the ofLm paths were simulated via an accept-reject procedure (see [34]).

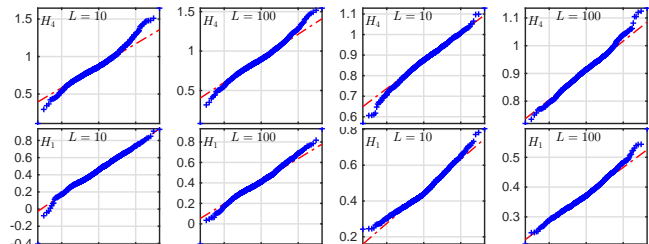
**Model order selection performance.** For different values of  $\gamma$  (namely, levels of departure from Gaussianity) and of the SNR, and for different dimensions  $L$ , Fig. 2 plots histograms of the estimated model orders  $\widehat{M}$  generated by means of the proposed model order selection procedure. The results demonstrate that for  $\gamma = 10^{-1}, 10^{-3}$



**Fig. 3. Estimation performance.** Bias, std and rms (left to right column) for  $L = 10/20/50/100$  (top to bottom row) and fixed SNR=30dB as a function of  $\gamma$ .

(weak departure from Gaussianity), the procedure selects the correct model order  $M = 4$  with high probability as long as SNR  $\geq 0$ dB. Note that the procedure's good performance is robust with respect to the value of  $L$ . Strikingly, we also observe a phenomenon that may be called the *blessing of dimensionality*: for lower SNR, larger dimensional observations turn out to be beneficial. For example, for SNR = -6dB, the selected model orders are correct with probability close to one for  $L \geq 50$ , yet the correct model order is only detected one out of two times for  $L = 10$ . This phenomenon has been observed in other frameworks (cf. [35]), in particular in Gaussian systems [27], and may be related to the fecund eigenvalue repulsion effect [36]. The fact that model order selection becomes easier in high dimensions can also be seen in Fig. 1, where the separation between the curves with negative and positive slopes becomes sharper as the dimension  $L$  increases. Turning back to Fig. 2, as expected, the stronger the departure from Gaussianity, the more difficult the model order selection problem. Nevertheless, the model order selection procedure remains quite reliable. For  $L \geq 20$  and SNR=0dB, the probability of correct detection is about 80% for  $\gamma = 10^{-5}, 10^{-7}$  (instead of  $\approx 100\%$  in lighter-tailed instances). For low SNR values, it is generally true that the majority of detection attempts are often too conservative ( $\widehat{M} < M$ ). Note that model order selection performance could be optimized by searching the best threshold value, which in our simulations is set a priori to 0.

**Estimation performance in high dimensions.** For instances involving low noise levels (30dB SNR), Fig. 3 plots bias, standard deviation (std) and root-mean-squared error (rms) for estimates  $\widehat{H}_m$  as a function of the dimension  $L = 10, 20, 50, 100$  and  $\gamma$ . The



**Fig. 4. Gaussianity of  $\widehat{H}_m$ .** For SNR=30dB,  $\gamma = \{10^{-7}, 10^{-3}\}$  (first two and last two columns, respectively) and  $L = 10, 100$ , qq-plots of the estimates  $\widehat{H}_4$  and  $\widehat{H}_1$  (largest and smallest exponents, top and bottom rows of subplots) against the standard Normal distribution.

plots reveal a number of properties. First, as expected, it can be seen that the estimation performance tends to deteriorate for stronger departures from Gaussianity (smaller  $\gamma$ ). This is essentially due to a reduction in the standard deviation as  $\gamma$  increases, whereas, remarkably, the bias remains unaffected. These results corroborate similar findings for the bivariate, bi-dimensional case  $L = M = 2$  without noise [16]. Second, and interestingly, the results indicate that there is an inversion in the relative estimation difficulty between the smallest exponent  $H_1$  than the largest exponent  $H_4$  as  $\gamma$  increases. In other words, for small values for  $\gamma$  (heavier tails), the smallest exponent  $H_1$  has lower std and rms than the largest exponent  $H_4$ ; by contrast, the opposite tends to be observed for large values for  $\gamma$ . Finally, estimation performance generally improves for larger  $L$ , both in terms of bias and std. This resonates with the blessing of dimensionality phenomenon (improved model order selection performance for large  $L$ ) observed in Figs. 1 and 2.

**Gaussianity of estimates  $\widehat{H}_m$ .** In Fig. 4, we investigate the influence of  $L$  and  $\gamma$  on the finite-sample distribution of  $\widehat{H}_m$ . We report qq-plots for estimates  $\widehat{H}_m$  against the standard Normal distribution for instances with very little noise (30dB SNR). We can observe that, regardless of  $L$ , the finite-sample distribution of  $\widehat{H}_m$  becomes more and more Gaussian as the regularization value  $\gamma$  increases, as expected and reported in [16] for  $M = L = 2$ . The results also show that there is no clear difference between the histograms for the smaller exponent  $H_1$  and for the largest exponent  $H_4$ . Moreover, the dimension  $L$  seems to bear little influence on the distribution of  $\widehat{H}_m$ .

## 5. CONCLUSIONS

In this paper, we construct statistical methodology for the identification of the number of multivariate self-similar, Lévy-driven components immersed in high-dimensional noise, as well as for the estimation of the associated scaling parameters. The technique draws upon the analysis of the evolution over scales of the eigenvalues of random wavelet matrices (wavelet eigenanalysis). Monte Carlo simulations reveal the effect of high dimensions on the estimation procedure and show that the proposed methodology is accurate for realistic sample sizes. This holds even at relatively low signal-to-noise ratios and for a large number of observed mixed and noisy time series. By tuning the tail parameter of the Lévy noise distribution, we further quantify the impact of non-Gaussianity on the estimation performance, where strongly non-Gaussian instances are shown to be typically more statistically challenging.

## 6. REFERENCES

- [1] P. Ciuciu, G. Varoquaux, P. Abry, S. Sadaghiani, and A. Kleinschmidt, "Scale-free and multifractal time dynamics of fMRI signals during rest and task," *Front. Physiol.*, vol. 3, 2012.
- [2] F. A. Isotta, C. Frei, V. Weilguni, M. Perčec Tadić, P. Lassegues, B. Rudolf, V. Pavan, C. Cacciamani, G. Antolini, S. M. Ratto, and M. Munari, "The climate of daily precipitation in the Alps: development and analysis of a high-resolution grid dataset from pan-Alpine rain-gauge data," *Int. J. Climatol.*, vol. 34, no. 5, pp. 1657–1675, 2014.
- [3] M. J. Wainwright, *High-Dimensional Statistics: a Non-Asymptotic Viewpoint*, vol. 48, Cambridge University Press, 2019.
- [4] P. Flandrin, "Wavelet analysis and synthesis of fractional Brownian motion," *IEEE Trans. Info. Theory*, vol. 38, pp. 910–917, March 1992.
- [5] M. S. Taqqu, "Fractional Brownian motion and long range dependence," in *Theory and Applications of Long-Range Dependence (P. Doukhan, G. Oppenheim and M. S. Taqqu, eds.)*, pp. 5–38. Birkhäuser, Boston, 2003.
- [6] 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.
- [7] B. C. Boniece, G. Didier, and F. Sabzikar, "On fractional Lévy processes: tempering, sample path properties and stochastic integration," *Preprint*, pp. 1–46, 2018.
- [8] W. Willinger, R. Govindan, S. Jamin, V. Paxson, and S. Shenker, "Scaling phenomena in the Internet: critically examining criticality," *Proc. Natl. Acad. Sci. U.S.A.*, vol. 99, no. suppl. 1, pp. 2573–2580, 2002.
- [9] A. Benassi, S. Cohen, and J. Istas, "Identification and properties of real harmonizable fractional Lévy motions," *Bernoulli*, vol. 8, no. 1, pp. 97–115, 2002.
- [10] T. Marquardt, "Fractional Lévy processes with an application to long memory moving average processes," *Bernoulli*, vol. 12, no. 6, pp. 1099–1126, 2006.
- [11] C. Lacaux and J.-M. Loubes, "Hurst exponent estimation of fractional Lévy motion," *ALEA: Lat. Am. J. Probab.*, vol. 3, pp. 143–164, 2007.
- [12] S. Zhang, Z. Lin, and X. Zhang, "A least squares estimator for Lévy-driven moving averages based on discrete time observations," *Commun. Stat. Theory Methods*, vol. 44, no. 6, pp. 1111–1129, 2015.
- [13] Y. Xu, Y. Li, H. Zhang, X. Li, and J. Kurths, "The switch in a genetic toggle system with Lévy noise," *Scientific Reports*, vol. 6, pp. 31505, 2016.
- [14] O. Barndorff-Nielsen and R. Stelzer, "Multivariate supOU processes," *Ann. Appl. Probab.*, vol. 21, no. 1, pp. 140–182, 2011.
- [15] M. Moser and R. Stelzer, "Functional regular variation of Lévy-driven multivariate mixed moving average processes," *Extremes*, vol. 16, no. 3, pp. 351–382, 2013.
- [16] B.C. Boniece, H. Wendt, G. Didier, and P. Abry, "On multivariate non-Gaussian scale invariance: fractional Lévy processes and wavelet estimation," in *Proc. Eur. Signal. Process. Conf. (EUSIPCO)*, 2019, pp. 1–5.
- [17] J. D. Mason and Y. Xiao, "Sample path properties of operator-self-similar Gaussian random fields," *Theory Probab. Appl.*, vol. 46, no. 1, pp. 58–78, 2002.
- [18] G. Didier and V. Pipiras, "Integral representations and properties of operator fractional Brownian motions," *Bernoulli*, vol. 17, no. 1, pp. 1–33, 2011.
- [19] P. Comon and C. Jutten, *Handbook of Blind Source Separation: Independent Component Analysis and Applications*, Academic Press, 2010.
- [20] P. Stoica and Y. Selen, "Model-order selection: a review of information criterion rules," *IEEE Signal Process. Mag.*, vol. 21, no. 4, pp. 36–47, 2004.
- [21] A. P. Liavas and P. A. Regalia, "On the behavior of information theoretic criteria for model order selection," *IEEE Trans. Signal Process.*, vol. 49, no. 8, pp. 1689–1695, 2001.
- [22] J. P. C. L. Da Costa, A. Thakre, F. Roemer, and M. Haardt, "Comparison of model order selection techniques for high-resolution parameter estimation algorithms," in *Proc. 54th Intern. Scient. Colloq. (IWK'09)*, 2009.
- [23] R. A. Johnson and D. W. Wichern, *Applied Multivariate Statistical Analysis*, vol. 4, Prentice-Hall, New Jersey, 2014.
- [24] R. Zhang, P. Robinson, and Q. Yao, "Identifying cointegration by eigenanalysis," *J. Am. Stat. Assoc.*, vol. 114, no. 526, pp. 916–927, 2018.
- [25] P. Abry and G. Didier, "Wavelet estimation for operator fractional Brownian motion," *Bernoulli*, vol. 24, no. 2, pp. 895–928, 2018.
- [26] P. Abry and G. Didier, "Wavelet eigenvalue regression for  $n$ -variate operator fractional Brownian motion," *J. Multivar. Anal.*, vol. 168, pp. 75–104, November 2018.
- [27] P. Abry, H. Wendt, and G. Didier, "Detecting and estimating multivariate self-similar sources in high-dimensional mixtures," in *IEEE Stat. Signal Process. Workshop*, 2019, pp. 1–5.
- [28] K. Sato, *Lévy Processes and Infinitely Divisible Distributions*, Cambridge University Press, U.K., 1999.
- [29] S. Mallat, *A Wavelet Tour of Signal Processing*, Academic Press, San Diego, CA, 1998.
- [30] B. C. Boniece, *On Scale Invariance and Wavelet Analysis: Transience, Operator Fractional Lévy Motion, and High-Dimensional Inference*, Ph.D. thesis, Tulane University, 2019.
- [31] P.-O. Amblard and J.-F. Coeurjolly, "Identification of the multivariate fractional Brownian motion," *IEEE Trans. Signal Process.*, vol. 59, no. 11, pp. 5152–5168, 2011.
- [32] Michael Grabchak, *Tempered Stable Distributions: Stochastic Models for Multiscale Processes*, Springer, 2016.
- [33] S. Stoev and M. S. Taqqu, "Simulation methods for linear fractional stable motion and FARIMA using the Fast Fourier Transform," *Fractals*, vol. 12, no. 1, pp. 95–121, 2004.
- [34] B. Baeumer and M. M. Meerschaert, "Tempered stable Lévy motion and transient super-diffusion," *Journal of Computational and Applied Mathematics*, vol. 233, no. 10, pp. 2438–2448, 2010.
- [35] C. Lam and Q. Yao, "Factor modeling for high-dimensional time series: inference for the number of factors," *Ann. Stat.*, vol. 40, no. 2, pp. 694–726, 2012.
- [36] T. Tao, *Topics in Random Matrix Theory*, vol. 132, American Mathematical Society, 2012.