# Bayesian Estimation of the Parameters of the Bivariate Multifractal Spectrum

Lorena Leon, Herwig Wendt, Jean-Yves Tourneret
Univ. Toulouse,
CNRS, IRIT, Toulouse, France.
`firstname.lastname@irit.fr`

Patrice Abry
Univ Lyon, Ens de Lyon, Univ Claude Bernard,
CNRS, Laboratoire de Physique, Lyon, France.
`patrice.abry@ens-lyon.fr`

*Abstract*—Multifractal analysis is a reference tool for the analysis of data based on local regularity and has proven useful in an increasing number of applications involving univariate data (scalar valued time series or single channel images). Recently the theoretical ground for a multivariate multifractal analysis has been explored, showing its potential for capturing and quantifying transient higher-order dependence beyond correlation among collections of data. Yet, the accurate estimation of the parameters associated with these multivariate multifractal models is challenging. Building on these first formulations of multivariate multifractal analysis, the present work proposes a Bayesian model and studies an estimation framework for the parameters of a quadratic model for the joint multifractal spectrum of bivariate time series. The approach relies on a novel joint Gaussian model for the logarithm of wavelet leaders and leverages on a Whittle approximation and data augmentation for the matrix-valued parameters of interest. Monte Carlo simulations demonstrate the benefits of the method with respect to previous formulations. In particular, we obtain significant performance improvements at only moderately larger computational cost, for large ranges of sample size and multifractal parameter values.

## I. INTRODUCTION

**Multifractal analysis.** Over the last decades, multifractal analysis has grown into a standard signal processing tool, characterizing data in terms of its pointwise regularity [1], [2]. This is theoretically achieved via the so-named multifractal spectrum, which describes geometrically the fluctuations of pointwise regularity in time or space. On average, these regularity fluctuations lead to scale-free statistics. These can be efficiently modeled and analyzed through the paradigm of scale-invariance and are in practice used to produce an upper-bound estimate $\mathcal{L}$ for the multifractal spectrum. In the past, multifractal analysis has led to significant successes in many real-world applications in very different contexts [3]–[8]. However, because the methodology remained fundamentally univariate, multifractal analysis has essentially been tied to the independent analysis of individual univariate time series/images. This severely impairs its use for data consisting of several jointly registered components, for which crucial information is potentially hidden in the coupling and dependencies between different components. While this limitation had been recognized early on and partially addressed in specific applicative contexts [9], [10], the theoretical foundation for the multifractal analysis of multivariate data was laid only recently [11], [12] (see Sec. II-A for a definition of the bivariate multifractal spectrum). Its first practical uses showed that the multivariate multifractal spectrum can effectively capture and quantify transient, local dependencies in data that cannot be accounted for by second order statistics such as the correlation function [13]–[15].

**Multifractal parameter estimation.** The state-of-the-art for multifractal analysis relies on the use of wavelet leaders, defined through a nonlinear transformation of wavelet coefficients, see Section II-B and [16] for definitions and pointers to alternative formalisms. In essence, these estimation frameworks rely on simple log-log regressions for computing power law exponents of sample moments or cumulants. However, such estimators can lead to significant estimation variance, in particular for small sample size. To overcome this issue, in [17], [18] a Bayesian model has been proposed for univariate multifractal analysis. It relies on a Gaussian model for the joint statistics of log-leaders, which yields significant performance improvements for univariate multifractal parameter estimates. The estimation of multivariate multifractal parameters has first been explored in [13]–[15] in bivariate settings using sample moments and cumulants of wavelet leaders. However, as in the univariate case, log-log regression-based multifractal parameter estimates yield modest statistical performance that limit the actual use of such features in applications.

**Goals, contributions and outline.** The present contribution aims to empower multivariate multifractal analysis by complementing it with a statistical framework that leads to improved estimation performance. To this end, we make use of the quadratic (i.e., second order) model for the bivariate multifractal spectrum that is proposed and studied in [13], [14] (briefly recalled in Sec. II-B). Our contributions are as follows. First, the Bayesian model for log-leaders defined in [17] is generalized to the bivariate setting by including an appropriate modeling of the cross-statistics, cf. Sec. III, and an adequate modification of the prior distributions for the parameters. Second, we design an efficient Gibbs sampling algorithm for estimating the multifractal parameters under this model. This is achieved by building on the Whittle approximation [19] and by proposing a data augmentation scheme for the matrix-valued parameters of interest, see Sec. IV. Finally, we report results for extensive Monte Carlo simulations using the synthetic bivariate multifractal process defined in [13], [14], validating the proposed estimation framework and demonstrating its practical

benefits in terms of estimation performance, at moderate extra computational cost.

## II. BIVARIATE MULTIFRACTAL ANALYSIS

### A. Multifractal spectrum

The goal of multifractal analysis is the quantification of the fluctuations along time of the regularity of a signal or function $X(t) \in \mathbb{R}$ at position $t \in \mathbb{R}$, where pointwise regularity is usually measured by the Hölder exponent, $h(t) \geq 0$. The closer $h(t)$ is to 0, the more irregular $X$ is around $t$, see, e.g., [1] for details. Let $\mathbf{h}(t) \triangleq (h_1(t), h_2(t))$ denote the Hölder exponents of the components of the bivariate signal $\mathbf{X}(t) = (X_1(t), X_2(t)) \in \mathbb{R}^2$. The *bivariate multifractal spectrum* $\mathcal{D}(h_1, h_2)$ of $\mathbf{X}$ is defined as the collection of Hausdorff dimensions $\dim_H$ of the sets of points $t \in \mathbb{R}$ at which $(h_1(t), h_2(t))$ take on the values $\mathbf{h} = (h_1, h_2)$

$$\mathcal{D}(h_1, h_2) \triangleq \dim_H \{t : (h_1(t), h_2(t)) = (h_1, h_2)\}, \quad (1)$$

(see [9], [11], [12] for details). It provides a global, geometrical description of the pointwise regularity of $\mathbf{X}$. Specifically, its precise shape, width, and orientation with respect to the $h_1, h_2$ axes quantify the degree and coupling of the local fluctuations of the regularity of the components of $\mathbf{X}$. The state-of-the-art procedure for its estimation is constructed from the multiscale statistics of wavelet leaders [1], [2], [11], [12].

### B. Bivariate multifractal formalism using wavelet leaders

**Wavelet leaders.** Let $\psi$ denote the mother wavelet, an oscillating reference pattern that is characterized by its number of vanishing moments $N_\psi$, a positive integer defined as $\psi \in \mathcal{C}^{N_\psi - 1}$ and $\forall n = 0, \dots, N_\psi - 1$, $\int_{\mathbb{R}} t^n \psi(t) dt \equiv 0$ and $\int_{\mathbb{R}} t^{N_\psi} \psi(t) dt \neq 0$. The mother wavelet is designed such that the collection $\{\psi_{j,k}(t) = 2^{-j/2} \psi(2^{-j} t - k)\}_{(j,k) \in \mathbb{Z}^2}$ of its dilated and translated templates forms an orthonormal basis of $L^2(\mathbb{R})$ [20]. The $L^1$ normalized discrete wavelet transform coefficients $d_X(j, k)$ of a time series $X \in \mathbb{R}$ are defined as $d_X(j, k) = 2^{-j/2} \langle \psi_{j,k} | X \rangle$. Then, the *wavelet leaders* of $X$ are defined as $L_X(j, k) \triangleq \sup_{\lambda' \subset 3\lambda_{j,k}} |d_X(\lambda')|$, where $\lambda_{j,k} = [k2^j, (k+1)2^j]$ denotes the dyadic interval of size $2^j$ and $3\lambda_{j,k}$ stands for the union of $\lambda_{j,k}$ with its 2 neighbors; they reproduce Hölder exponents in the limit of fine scales, $L_X(j, k) \sim C 2^{-jh(t)}$ for $2^j \to 0$, $k2^{-j} \sim t$ ($t$ fixed) [1], [2]. **Bivariate multifractal formalism.** The bivariate cumulants $C_{p_1 p_2}(j)$ of the vector of log-leaders $\boldsymbol{\ell}_X(j, k) = (\ln L_{X_1}(j, k), \ln L_{X_2}(j, k)) \in \mathbb{R}^2$ at scale $2^j$ can be shown to take the form [13], [21]

$$C_{p_1 p_2}(j) = c_{p_1 p_2}^0 + j\, c_{p_1 p_2} \ln 2, \quad p_1 + p_2 \geq 1. \quad (2)$$

The coefficients $c_{p_1 p_2}$ yield an approximation of $\mathcal{D}(\mathbf{h})$

$$\mathcal{L}(\boldsymbol{h}) \approx 1 + \frac{c_{02} b}{2}\left(\frac{h_1 - c_{10}}{b}\right)^2 + \frac{c_{20} b}{2}\left(\frac{h_2 - c_{01}}{b}\right)^2$$
$$- c_{11} b \left(\frac{h_1 - c_{10}}{b}\right)\left(\frac{h_2 - c_{01}}{b}\right), \quad (3)$$

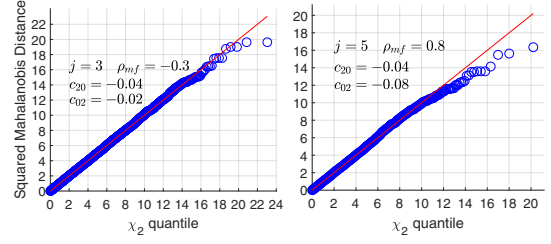where $b \triangleq c_{20} c_{02} - c_{11}^2 \geq 0$ [14] and



Fig. 1. Gamma plots for the empirical log-leaders joint distribution, at scales $j = 3, 5$ (left column and right column), associated with 100 independent copies of a bMRW for $N = 2^{10}$. The closer to the red line, the better the approximation of the distribution by a bivariate normal distribution.

- $(c_{10}, c_{01})$ indicate the position of the maximum of $\mathcal{D}(\mathbf{h})$
- $c_{20}, c_{02}$ quantify the amount of regularity fluctuations for each component independently
- $c_{11}$ characterizes the coupling between the components' regularity fluctuations.

## III. STATISTICAL MODEL FOR BIVARIATE LOG-LEADERS

The estimation of the coefficients of $\mathcal{L}(\boldsymbol{h})$ relies on linear regressions for sample cumulants as in (2). This procedure can lead to large estimation variance, in particular for the second order cumulants $c_{02}, c_{20}, c_{11}$ [13]–[15]. Instead, we propose here a Bayesian model and estimators for these parameters.

### A. Time domain model

Our model is inspired by and generalizes a model for univariate data studied in [17], [18]. In particular, it was shown that the distribution of the collection of scalar log-leaders $\ln L_X(j, \cdot)$ at scale $2^j$ can be reasonably well approximated by a Gaussian distribution.
**Bivariate marginal distributions.** Let $l_X(j, \cdot) \in \mathbb{R}^{n_j \times 2}$ denote the vector of all centered log-leaders $\boldsymbol{\ell}_X(j, \cdot)$ at scale $2^j$. We empirically show, using numerical simulations for synthetic bivariate multifractal processes as defined in Sec. V-A, for large ranges of sample size and multifractal parameter values, that $l_X(j, \cdot)$ can be well modeled as a Gaussian random vector. Illustrative examples are given in Fig. 1, which shows bivariate gamma plots [22], [23] for scales $j = 3, 5$ and for several different parameter settings, providing numerical evidence of bivariate normality.
**Covariance.** In our previous work, we have shown that the covariance of scalar log-leaders of multifractal processes has an asymptotic log-linear decay, see, e.g., [17], [18]. Here we propose a generalization to the bivariate case by amending the covariance model in [17], [18] with a cross-term describing the covariance between log-leaders of different components. Let $\boldsymbol{S}_j(r) \triangleq \mathbb{E}\{l_X(j, k)^T l_X(j, k + r)\}$, with $r \triangleq |\Delta k|$, denote the covariance matrix of $l_X(j, k)$ at a fixed scale $2^j$. Extensive numerical simulations indicate that

$$\boldsymbol{S}_j(r) \approx \boldsymbol{\Sigma}_1 f_1(j, r) + \boldsymbol{\Sigma}_2 f_2(j, r), \quad (4)$$

with $\boldsymbol{\Sigma}_1 = \begin{pmatrix} v_{20} & v_{11} \\ v_{11} & v_{02} \end{pmatrix}$, $\boldsymbol{\Sigma}_2 = \begin{pmatrix} v_{20}^0 & v_{11}^0 \\ v_{11}^0 & v_{02}^0 \end{pmatrix}$, $v_{20} = -c_{20}, v_{02} = -c_{02}, v_{11} = -c_{11}, f_1(j, r) = \max(0, -\ln((r +$
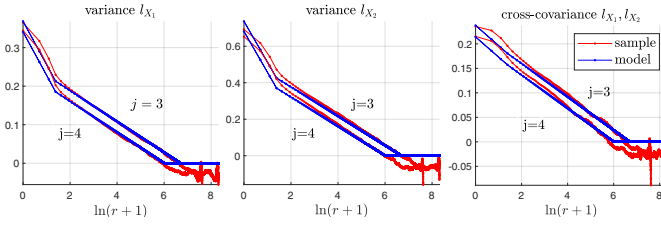
Fig. 2. Comparison between the proposed model (blue line) and the sample covariance (red line) at two scales $j = 3, 4$, averaged over 100 independent realizations of a bMRW for $N = 2^{15}$, $(\lambda_1, \lambda_2) = (\sqrt{0.04}, \sqrt{0.08})$ and $\rho_{\mathrm{mf}} = 0.6$. Plots (a), (b) and (c) correspond to each covariance matrix element.

$1)/(r_j + 1))$ and $f_2(j, r) = \max(0, 1 - \ln(r + 1)/\ln 4)$ with $r_j = \lfloor n_j/5 \rfloor$, where $\lfloor \cdot \rfloor$ truncates to integer values. Illustrative examples of this covariance model are provided in Fig. 2. Theoretically, $c_{20}, c_{02} < 0$ and $c_{20}c_{02} - c_{11}^2 > 0$, hence $v_{20}, v_{02} > 0$, $v_{20}v_{02} - v_{11}^2 > 0$ and thus $\mathbf{\Sigma}_1$ is positive definite (p.d.). Moreover, for all $i \in \{1, 2\}$ and scales $j$, $f_i(j, \cdot)$ is a non-negative function. Therefore, assuming that $\mathbf{\Sigma}_2$ is p.d. ensures that $\mathbf{S}(r)$ is positive semi-definite (p.s.d.) $\forall r$, as it is the sum of two p.s.d. matrices [24]. Finally, $\mathbf{\Sigma}_1$ and $\mathbf{\Sigma}_2$ are the two matrix-valued parameters to be estimated.

**Likelihood.** Denote as $\boldsymbol{y}_j = \left[\boldsymbol{l}_{X_1}(j, \cdot)^T, \boldsymbol{l}_{X_2}(j, \cdot)^T\right]^T$ the $2n_j \times 1$ vector resulting from the concatenation of the component-wise column vectors of centered log-leaders $\boldsymbol{l}_X(j, \cdot)$ at scale $2^j$. We assume independence of leaders at different scales $j' \neq j$. Together with the above model, the likelihood of the vector $\boldsymbol{y} = \left[\boldsymbol{y}_{j_1}^T, \ldots, \boldsymbol{y}_{j_2}^T\right]^T$ of the log-leaders at scales $j = j_1, \ldots, j_2$ given $\mathbf{\Sigma}_1$ and $\mathbf{\Sigma}_2$ reads

$$p(\boldsymbol{y} \mid \mathbf{\Sigma}_1, \mathbf{\Sigma}_2) = \prod_{j=j_1}^{j_2} p(\boldsymbol{y}_j \mid \mathbf{\Sigma}_1, \mathbf{\Sigma}_2)$$
$$\propto \prod_{j=j_1}^{j_2} (\det \mathbf{\Xi}_j)^{-1/2} \exp\left(-\frac{1}{2}\boldsymbol{y}_j^T \mathbf{\Xi}_j^{-1} \boldsymbol{y}_j\right), \quad (5)$$

where $\mathbf{\Xi}_j = \mathbf{\Sigma}_1 \otimes \boldsymbol{F}_{1j} + \mathbf{\Sigma}_2 \otimes \boldsymbol{F}_{2j}$ is a $2n_j \times 2n_j$ covariance matrix, $\otimes$ is the Kronecker product, the matrices $\boldsymbol{F}_{ij}$ are defined element-wise as $[\boldsymbol{F}_{ij}]_{u,v} = f_i(j, |k_u - k_v|)$, $i = 1, 2$. For all $i \in \{1, 2\}$ and scales $j$, the positive definiteness of $\boldsymbol{F}_{ij}$ can only be assessed numerically to check that $\mathbf{\Sigma}_i \otimes \boldsymbol{F}_{ij}$ and $\mathbf{\Xi}_{ij}$ are p.d.

### B. Whittle approximation

The numerical evaluation of the likelihood (5) being problematic for large sample size (in particular requiring computation of the matrix inverses $\mathbf{\Xi}_j^{-1}$), we resort to a Whittle approximation [19], [25]. This standard technique approximates the time-domain likelihood (5) in the frequency domain as

$$p_W(\boldsymbol{y}_j \mid \mathbf{\Sigma}_1, \mathbf{\Sigma}_2) \propto \prod_{m \in I_j} \left(\det \tilde{\boldsymbol{S}}_j(\omega_m)\right)^{-1}$$
$$\exp\left(-\boldsymbol{z}_j^H(\omega_m)\tilde{\boldsymbol{S}}_j^{-1}(\omega_m)\boldsymbol{z}_j(\omega_m)\right), \quad (6)$$

where $\cdot^H$ denotes the conjugate transpose operator and $\boldsymbol{z}_j(\omega_m)$ is the bivariate discrete Fourier transform (DFT) coefficient $\boldsymbol{z}_j(\omega_m) = \frac{1}{\sqrt{n_j}} \sum_{k=1}^{n_j} \begin{bmatrix} \boldsymbol{l}_{X_1}(j, k) \\ \boldsymbol{l}_{X_2}(j, k) \end{bmatrix} \exp(-ik\omega_m)$, with

$\omega_m = 2\pi m/\sqrt{n_j}$ and $I_j = [\![-\lceil\sqrt{n_j}/2\rceil + 1, \lfloor\sqrt{n_j}/2\rfloor]\!] \setminus \{0\}$ ($[\![a_1, a_2]\!]$ denotes the set of integers ranging from $a_1$ to $a_2$). Moreover, the power spectral matrix $\tilde{\boldsymbol{S}}_j(\omega_m)$ forms a Fourier pair with the covariance matrix $\boldsymbol{S}_j(r)$, which can be approximated using a discrete Fourier transform (DFT) of $f_i(j, \cdot)$ denoted as $f_i^{\mathrm{DFT}}(j, \cdot)$, $i = 1, 2$, as

$$\tilde{\boldsymbol{S}}_j(\omega_m) = \mathbf{\Sigma}_1 f_1^{\mathrm{DFT}}(j, \omega_m) + \mathbf{\Sigma}_2 f_2^{\mathrm{DFT}}(j, \omega_m). \quad (7)$$

For all $i \in \{1, 2\}$ and scales $j$, the imaginary part of $f_i^{\mathrm{DFT}}(j, \cdot)$ is equal to 0. Thus, only the real part is considered, which is also strictly positive. Since $\mathbf{\Sigma}_1$ and $\mathbf{\Sigma}_2$ are assumed to be p.d., $\tilde{\boldsymbol{S}}_j(\cdot)$ is a real-valued p.d. matrix for all scales $j$.

### C. Data augmentation

Equation (6) suggests that $p_W(\boldsymbol{y}_j \mid \mathbf{\Sigma}_1, \mathbf{\Sigma}_2)$ is equivalent to modeling $\boldsymbol{z}_j = [\boldsymbol{z}_j(\omega_1), \ldots, \boldsymbol{z}_j(\omega_{N_I})]^T$ with $N_I = \mathrm{card}(I_j)$ by a random vector with a non-degenerate centered circular-symmetric complex Gaussian distribution ($\mathcal{CN}$). Thus, we consider the collection of Fourier coefficients $\boldsymbol{z} = \left[\boldsymbol{z}_{j_1}^T, \ldots, \boldsymbol{z}_{j_2}^T\right]^T$ as the observed data. The likelihood of $\boldsymbol{z}$ is given by

$$p(\boldsymbol{z} \mid \mathbf{\Sigma}_1, \mathbf{\Sigma}_2) = \prod_{j=j_1}^{j_2} p(\boldsymbol{z}_j \mid \mathbf{\Sigma}_1, \mathbf{\Sigma}_2)$$
$$\propto (\det \mathbf{\Omega})^{-1} \exp\left(-\boldsymbol{z}^H \mathbf{\Omega}^{-1} \boldsymbol{z}\right), \quad (8)$$

where $\mathbf{\Omega} \triangleq \mathbf{\Sigma}_1 \otimes \boldsymbol{G}_1 + \mathbf{\Sigma}_2 \otimes \boldsymbol{G}_2$ is a real-valued block diagonal p.d. covariance matrix, with $\boldsymbol{G}_i \triangleq \mathrm{diag}(\boldsymbol{g}_i)$, $\boldsymbol{g}_i \triangleq [\boldsymbol{g}_{ij_1}, \ldots, \boldsymbol{g}_{ij_2}]^T$, $\boldsymbol{g}_{ij} \triangleq (f_i^{\mathrm{DFT}}(j, \omega_m))_{m \in I_j}$, for $i = 1, 2$.

Model (8) is simple and cheap to evaluate numerically. Its main inconvenience regarding the estimation of $\mathbf{\Sigma}_1$ and $\mathbf{\Sigma}_2$ is that these matrices are additively tied together in $\mathbf{\Omega}$ so that it is not possible to design conjugate priors leading to simple conditional distributions. To bypass this difficulty we use data augmentation and introduce a complex-valued vector of latent variables $\boldsymbol{u}$ defining the conditional distributions

$$\boldsymbol{z} \mid \boldsymbol{u}, \mathbf{\Sigma}_1 \sim \mathcal{CN}(\boldsymbol{u}, \mathbf{\Sigma}_1 \otimes \boldsymbol{G}_1), \ \boldsymbol{u} \mid \mathbf{\Sigma}_2 \sim \mathcal{CN}(\boldsymbol{0}, \mathbf{\Sigma}_2 \otimes \boldsymbol{G}_2),$$

leading to the augmented likelihood

$$p(\boldsymbol{z}, \boldsymbol{u} \mid \mathbf{\Sigma}_1, \mathbf{\Sigma}_2) \propto p(\boldsymbol{z} \mid \boldsymbol{u}, \mathbf{\Sigma}_1)p(\boldsymbol{u} \mid \mathbf{\Sigma}_2). \quad (9)$$

By construction, the likelihood (8) is obtained by marginalizing (9) with respect to $\boldsymbol{u}$.

## IV. BAYESIAN ESTIMATION

In this section, a Bayesian model and a procedure for the estimation of $\mathbf{\Sigma}_1$ and $\mathbf{\Sigma}_2$, are described.

### A. Likelihood, priors and posterior

The proposed Bayesian model is based on the augmented likelihood (9), which is the product of complex Gaussian distributions having $\mathbf{\Sigma}_1$ and $\mathbf{\Sigma}_2$ as covariance matrices. The natural conjugate prior for these matrices is the inverse Wishart (IW) prior [26], $\mathbf{\Sigma}_i \sim \mathcal{IW}(\nu_i, \mathbf{\Lambda}_i)$, $i = 1, 2$, with $\nu_i \in \mathbb{R}$ ($\nu_i > 3$) degrees of freedom and a $2-$dimensional p.d. scale matrix $\mathbf{\Lambda}_i$. The posterior distribution built on the augmented likelihood and the IW priors is

$$p(\mathbf{\Sigma}_1, \mathbf{\Sigma}_2, \boldsymbol{u} \mid \boldsymbol{z}) \propto p(\boldsymbol{z}, \boldsymbol{u} \mid \mathbf{\Sigma}_1)p(\boldsymbol{u} \mid \mathbf{\Sigma}_2)p(\mathbf{\Sigma}_1)p(\mathbf{\Sigma}_2). \quad (10)$$

## B. Bayesian estimators

The posterior distribution (10) can be used to define the marginal minimum mean square error (MMSE) estimator $\Sigma_i^{\mathrm{MMSE}} = \mathbb{E}\left[\Sigma_i \mid z, u\right]$, $i = 1, 2$, approximated here via a Markov chain Monte Carlo (MCMC) algorithm (Gibbs sampler). For details about MCMC, see, e.g., [27].

## C. Gibbs sampler

The Gibbs sampler successively generates samples $\left\{u^\lambda, \Sigma_1^\lambda, \Sigma_2^\lambda\right\}_{\lambda=1}^{N_{\mathrm{mc}}}$ according to the conditional distributions of (10), i.e., a complex Gaussian distribution $p(u \mid \Sigma_1, \Sigma_2, z)$, and two inverse Wishart distributions $p(\Sigma_1 \mid z, u)$, $p(\Sigma_2 \mid u)$. After a burn-in period, where the first $N_{\mathrm{bi}}$ samples are discarded, the MMSE estimator is approximated by $\Sigma_i^{\mathrm{MMSE}} \approx (N_{\mathrm{mc}} - N_{\mathrm{bi}})^{-1} \sum_{\lambda=N_{\mathrm{bi}}+1}^{N_{\mathrm{mc}}} \Sigma_i^\lambda$, $i = 1, 2$.

## V. ESTIMATION PERFORMANCE ASSESSMENT

In this section, we compare the proposed Bayesian approach using IW priors, denoted as MMSE, to the weighted linear regression-based method (see, e.g., [2]) denoted as WLR. The comparison is performed applying both methods on a large number of independent realizations of a synthetic bivariate multifractal process described next.

## A. Bivariate multifractal random walk (bMRW)

The construction of bMRW [13] requires two pairs of stochastic processes: a pair of increments of fBm, $(G_1(t), G_2(t))$, which is determined by two self-similarity parameters $H_1$ and $H_2$ and a point covariance $\Sigma_{ss}$, and a pair of Gaussian processes $(\omega_1(t), \omega_2(t))$ with prescribed co-variance function $\Sigma_{mf}$, with entries given by $[\Sigma_{mf}]_{iv}(k,l) = [\rho_{mf}]_{iv} \lambda_i \lambda_v \log\left(\frac{T}{|k-l|+1}\right)$, $i, v = 1, 2$, for $|k - l| \le T - 1$ and 0 otherwise, where $T$ is an arbitrary integral scale. To simplify notations, we consider $[\rho_{mf}]_{11} = [\rho_{mf}]_{22} = 1$, and $[\rho_{mf}]_{12} = \rho_{mf}$. Both pairs of processes are numerically synthesized as described in [28]. Each component $i = 1, 2$ of bMRW is then defined as $X_i(t) = \sum_{k=1}^{t} G_i(k)e^{\omega_i(k)}$. Following [13], [29], the bivariate multifractality parameters of bMRW are given by $c_{10} = H_1 + \lambda_1^2/2$, $c_{01} = H_2 + \lambda_2^2/2$, $c_{20} = -\lambda_1^2$, $c_{02} = -\lambda_2^2$, and $c_{11} = -\rho_{mf}\lambda_1\lambda_2$. Moreover, $c_{p_1 p_2} \equiv 0$, $\forall (p_1, p_2)$ such that $p_1 + p_2 \ge 3$.

## B. Monte Carlo simulations and parameter setting

The parameters of the process are set to $(H_1, H_2) = (0.6, 0.8)$, $(\lambda_1, \lambda_2) = (\sqrt{0.02}, \sqrt{0.02})$, $\rho_{ss} = 0$ and $\rho_{mf} = 0.5$. The considered sample sizes are $N = \{2^8, 2^{10}, 2^{12}, 2^{14}, 2^{16}\}$ and the integral scale is set to $T = N$. A wavelet analysis is conducted with a Daubechies least asymmetric wavelet, with $N_\psi = 3$ and scale ranges $j_1 = 3$ and $j_2 = \{5, 7, 9, 10, 10\}$. The Gibbs sampler uses $N_{\mathrm{mc}} = 2000$ and $N_{\mathrm{bi}} = 1000$. For $i \in \{1, 2\}$, the hyperparameters of the IW prior for $\Sigma_i$ are set to $\nu_i = 4$ and $\Lambda_i = I_2$ (identity matrix), as suggested in [30]. The estimation performance is analyzed through Monte Carlo simulations over 100 independent copies of bMRW and evaluated via the sample mean or bias, the

TABLE I
BMRW ESTIMATION PERFORMANCE FOR $N = \{2^8, 2^{10}, 2^{12}, 2^{14}, 2^{16}\}$, $j_1 = 3$ AND $j_2 = \{5, 7, 9, 10, 10\}$. BEST RESULTS ARE MARKED IN BOLD.

| | | N | $2^8$ | $2^{10}$ | $2^{12}$ | $2^{14}$ | $2^{16}$ |
|---|---|---|---|---|---|---|---|
| $-c_{20}=0.02$ | $b$ | WLR | **0.0190** | **0.0025** | **0.0042** | **0.0075** | 0.0070 |
| | | MMSE | 0.0620 | 0.0211 | 0.0127 | 0.0078 | **0.0056** |
| | $s$ | WLR | 0.0662 | 0.0288 | 0.0155 | 0.0083 | 0.0048 |
| | | MMSE | **0.0142** | **0.0058** | **0.0043** | **0.0029** | **0.0020** |
| | $r$ | WLR | 0.0689 | 0.0289 | 0.0161 | 0.0112 | 0.0085 |
| | | MMSE | **0.0636** | **0.0219** | **0.0134** | **0.0083** | **0.0060** |
| $-c_{02}=0.02$ | $b$ | WLR | **0.0199** | **0.0064** | **0.0026** | **0.0047** | **0.0051** |
| | | MMSE | 0.0704 | 0.0308 | 0.0173 | 0.0113 | 0.0082 |
| | $s$ | WLR | 0.0998 | 0.0371 | 0.0176 | 0.0083 | 0.0054 |
| | | MMSE | **0.0142** | **0.0066** | **0.0047** | **0.0033** | **0.0020** |
| | $r$ | WLR | 0.1017 | 0.0377 | **0.0178** | **0.0096** | **0.0075** |
| | | MMSE | **0.0718** | **0.0315** | 0.0180 | 0.0118 | 0.0084 |
| $-c_{11}=0.01$ | $b$ | WLR | **0.0004** | **0.0005** | **0.0002** | **0.0001** | 0.0004 |
| | | MMSE | 0.0065 | 0.0056 | 0.0035 | 0.0011 | **0** |
| | $s$ | WLR | 0.0485 | 0.0198 | 0.0121 | 0.0061 | 0.0031 |
| | | MMSE | **0.0080** | **0.0040** | **0.0029** | **0.0021** | **0.0012** |
| | $r$ | WLR | 0.0485 | 0.0198 | 0.0121 | 0.0061 | 0.0031 |
| | | MMSE | **0.0103** | **0.0069** | **0.0046** | **0.0024** | **0.0012** |

TABLE II
BMRW ESTIMATION PERFORMANCE FOR $N = 2^{14}$, $j_1 = 3$, $j_2 = 10$ AND $-c_{02} = \{0.02, 0.04, 0.06, 0.08, 0.10\}$. BEST RESULTS ARE MARKED IN BOLD.

| | | $-c_{02}$ | 0.02 | 0.04 | 0.06 | 0.08 | 0.10 |
|---|---|---|---|---|---|---|---|
| $-c_{20}=0.02$ | $b$ | WLR | **0.0075** | **0.0066** | **0.0067** | **0.0062** | **0.0064** |
| | | MMSE | 0.0078 | 0.0072 | 0.0070 | 0.0072 | 0.0066 |
| | $s$ | WLR | 0.0083 | 0.0070 | 0.0080 | 0.0081 | 0.0080 |
| | | MMSE | **0.0029** | **0.0030** | **0.0029** | **0.0029** | **0.0030** |
| | $r$ | WLR | 0.0112 | 0.0096 | 0.0104 | 0.0102 | 0.0102 |
| | | MMSE | **0.0083** | **0.0078** | **0.0075** | **0.0078** | **0.0072** |
| $-c_{02}$ | $b$ | WLR | **0.0047** | **0.0067** | 0.0069 | 0.0082 | 0.0097 |
| | | MMSE | 0.0113 | 0.0114 | **0.0065** | **0.0042** | **0.0010** |
| | $s$ | WLR | 0.0083 | 0.0132 | 0.0145 | 0.0178 | 0.0207 |
| | | MMSE | **0.0033** | **0.0051** | **0.0066** | **0.0078** | **0.0089** |
| | $r$ | WLR | **0.0096** | 0.0148 | 0.0161 | 0.0195 | 0.0228 |
| | | MMSE | 0.0118 | **0.0125** | **0.0092** | **0.0089** | **0.0090** |
| $-c_{11}$ | | | 0.0100 | 0.0141 | 0.0173 | 0.0200 | 0.0224 |
| | $b$ | WLR | **0.0001** | **0.0007** | **0.0002** | **0.0010** | **0.0010** |
| | | MMSE | 0.0011 | 0.0021 | 0.0032 | 0.0048 | 0.0054 |
| | $s$ | WLR | 0.0061 | 0.0070 | 0.0076 | 0.0081 | 0.0095 |
| | | MMSE | **0.0021** | **0.0027** | **0.0031** | **0.0034** | **0.0036** |
| | $r$ | WLR | 0.0061 | 0.0071 | 0.0076 | 0.0081 | 0.0095 |
| | | MMSE | **0.0024** | **0.0034** | **0.0045** | **0.0058** | **0.0065** |

sample standard deviation (STD) and the root mean square error (RMSE) across realizations of the estimates for $\theta \in \{-c_{20}, -c_{02}, -c_{11}\}$, defined as $m(\hat{\theta}) = \widehat{\mathbb{E}}[\hat{\theta}]$, $b(\hat{\theta}) = m(\hat{\theta}) - \theta$, $s(\hat{\theta}) = \sqrt{\widehat{\mathrm{Var}}[\hat{\theta}]}$ and $r(\hat{\theta}) = \sqrt{b(\hat{\theta})^2 + s(\hat{\theta})^2}$.

## C. Results

Table I summarizes the estimation performance of WLR and MMSE estimators (using IW priors) for bMRW for a range of sample sizes $N$. Overall, it is observed that the MMSE estimator tends to have larger bias than WLR, which decreases when the sample size $N$ increases, as expected. On the other hand, the Bayesian estimator yields significantly reduced standard deviations, with those of WLR being more than 2.5 times larger. These estimation variance reductions, which are more pronounced for small sample size, influence

directly the overall RMSE values, which are in general smaller for MMSE than for WLR. These performance gains are particularly important for the parameter $c_{11}$ measuring joint multifractality, for which RMSE values are twice as large for WLR. The price to pay with this performance improvement is an increased, yet reasonable, computational cost of $16s$ for the Bayesian estimator for $N = 2^{14}$ (versus $10^{-4}s$ for WLR).

Table II contains the estimation performance of the WLR and MMSE estimators for bMRW process with sample size $N = 2^{14}$ for a range of weak to strong multifractality for the second component $X_2$, i.e., $-c_{02} = \{0.02, 0.04, 0.06, 0.08, 0.10\}$. One can observe that the bias of the MMSE estimator of $-c_{02}$ decreases if $-c_{02}$ increases and drops below 10% of the value of $-c_{02}$ for $-c_{02} > 0.06$. Conversely, the bias of the MMSE estimator of $-c_{11}$ increases if $-c_{02}$ increases, which reflects a dependence between the estimation of these parameters. In terms of STD, the MMSE estimator shows again reduced values as compare to WLR, which has overall twice as large STD. Likewise, the RMSE values for the MMSE estimator remain below those of WLR.

## VI. Conclusion

This paper proposed a Bayesian estimation procedure for the multifractal parameters $c_{02}, c_{20}, c_{11}$ given a bivariate time series. It relies on a novel joint Gaussian model for a bivariate vector of centered logarithms of wavelet leaders, based on a Whittle approximation and data augmentation for the matrix-valued parameters of interest. Monte Carlo simulations demonstrated the benefits of the method with respect to alternative estimators (weighted linear regression). In particular, the proposed Bayesian estimator showed a significant reduction of variance and mean square error when compared to the linear regression method, especially for the parameter of joint multifractality $c_{11}$. An increase in bias was observed for $c_{02}, c_{20}$ close to zero, as well as dependence between the estimator of $c_{11}$ and the diagonal elements of $\mathbf{\Sigma_1}$, both potentially induced by the inverse Wishart priors used here. Future work will include the study of alternative prior distributions, generalization of our model to texture (images), and applications to real-world data.

## References

[1] S. Jaffard, "Wavelet techniques in multifractal analysis," in *Fractal Geometry and Applications: A Jubilee of Benoît Mandelbrot, Proceedings of Symposia in Pure Mathematics*, M. Lapidus and M. van Frankenhuijsen, Eds., vol. 72(2). AMS, 2004, pp. 91–152.

[2] H. Wendt, P. Abry, and S. Jaffard, "Bootstrap for empirical multifractal analysis," *IEEE Signal Proces. Mag.*, vol. 24, no. 4, pp. 38–48, 2007.

[3] K. Kiyono, Z. R. Struzik, N. Aoyagi, and Y. Yamamoto, "Multiscale probability density function analysis: non-Gaussian and scale-invariant fluctuations of healthy human heart rate," *IEEE Trans. Biomed. Eng.*, vol. 53, no. 1, pp. 95–102, Jan. 2006.

[4] B. B. Mandelbrot, "Intermittent turbulence in self-similar cascades: divergence of high moments and dimension of the carrier," *J. Fluid Mech.*, vol. 62, pp. 331–358, 1974.

[5] 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, Jun. 2012.

[6] M. Doret, H. Helgason, P. Abry, P. Gonçalvès, C. Gharib, and P. Gaucherand, "Multifractal analysis of fetal heart rate variability in fetuses with and without severe acidosis during labor," *Am. J. Perinatol.*, vol. 28, no. 4, pp. 259–266, 2011.

[7] 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 T. Networking*, vol. 25, no. 4, 2017.

[8] A. Johansen and D. Sornette, "Finite-time singularity in the dynamics of the world population, economic and financial indices," *Physica A*, vol. 294, pp. 465–502, 2001.

[9] C. Meneveau, K. Sreenivasan, P. Kailasnath, and M. Fan, "Joint multifractal measures - theory and applications to turbulence," *Phys. Rev. A*, vol. 41, no. 2, pp. 894–913, 1990.

[10] T. Lux, "Higher dimensional multifractal processes: A GMM approach," *Journal of Business and Economic Statistics*, vol. 26, pp. 194–210, 2007.

[11] S. Jaffard, S. Seuret, H. Wendt, R. Leonarduzzi, S. Roux, and P. Abry, "Multivariate multifractal analysis," *Applied and Computational Harmonic Analysis*, vol. 46, no. 3, pp. 653–663, 2019.

[12] S. Jaffard, S. Seuret, H. Wendt, R. Leonarduzzi, and P. Abry, "Multifractal formalisms for multivariate analysis," *Proc. Royal Society A*, vol. 475, no. 2229, 2019.

[13] H. Wendt, R. Leonarduzzi, P. Abry, S. Roux, S. Jaffard, and S. Seuret, "Assessing cross-dependencies using bivariate multifractal analysis," in *IEEE Int. Conf. Acoust., Speech, and Signal Proces. (ICASSP)*, Calgary, Canada, April 2018.

[14] R. Leonarduzzi, P. Abry, S. G. Roux, H. Wendt, S. Jaffard, and S. Seuret, "Multifractal characterization for bivariate data," in *Proc. European Signal Process. Conf. (EUSIPCO)*, Rome, Italy, September 2018.

[15] P. Abry, R. Leonarduzzi, H. Wendt, S. Jaffard, and S. Seuret, "A multivariate multifractal analysis of lacunary wavelet series," in *Proc. IEEE Int. Workshop CAMSAP*, Guadeloupe, France, Dec. 2019.

[16] S. Jaffard, P. Abry, and H. Wendt, *Irregularities and Scaling in Signal and Image Processing: Multifractal Analysis*. Singapore: World scientific publishing, 2015, pp. 31–116.

[17] S. Combrexelle, H. Wendt, N. Dobigeon, J.-Y. Tourneret, S. McLaughlin, and P. Abry, "Bayesian estimation of the multifractality parameter for image texture using a whittle approximation," *IEEE T. Image Proces.*, vol. 24, no. 8, pp. 2540–2551, 2015.

[18] H. Wendt, S. Combrexelle, Y. Altmann, J.-Y. Tourneret, S. McLaughlin, and P. Abry, "Multifractal analysis of multivariate images using gamma markov random field priors," *SIAM J. on Imaging Sciences (SIIMS)*, vol. 11, no. 2, pp. 1294–1316, 2018.

[19] P. Whittle, "Estimation and information in stationary time series," *Ark. Mat.*, vol. 2, no. 5, pp. 423–434, 1953.

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

[21] B. Castaing, Y. Gagne, and M. Marchand, "Log-similarity for turbulent flows," *Physica D*, vol. 68, no. 3-4, pp. 387–400, 1993.

[22] R. A. Johnson and D. W. Wichern, *Applied multivariate statistical analysis*. Prentice hall Upper Saddle River, NJ, 2002, vol. 5, no. 8.

[23] F. B. Oppong and S. Y. Agbedra, "Assessing Univariate and Multivariate Normality , A Guide For Non - Statisticians," *Mathematical Theory and Modeling*, vol. 6, no. 2, pp. 26–33, 2016.

[24] R. A. Horn and C. R. Johnson, *Matrix Analysis*, 2nd ed. Cambridge University Press, 2012.

[25] A. M. Sykulski, S. C. Olhede, J. M. Lilly, and J. J. Early, "Frequency-domain stochastic modeling of stationary bivariate or complex-valued signals," *IEEE T. Signal Process.*, vol. 65, no. 12, pp. 3136–3151, 2017.

[26] J. Barnard, R. McCulloch, and X.-L. Meng, "Modelling covariance matrices in terms of standard deviations and correlations, with application to shrinkage," *Statistica Sinica*, vol. 10, pp. 1281–1311, 10 2000.

[27] C. P. Robert and G. Casella, *Monte Carlo Statistical Methods*. New York, USA: Springer, 2005.

[28] H. Helgason, V. Pipiras, and P. Abry, "Fast and exact synthesis of stationary multivariate Gaussian time series using circulant embedding," *Signal Proc.*, vol. 91, no. 5, pp. 1123–1133, 2011.

[29] E. Bacry, J. Delour, and J.-F. Muzy, "Multifractal random walk," *Phys. Rev. E*, vol. 64: 026103, 2001.

[30] I. Alvarez, J. Niemi, and M. Simpson, "Bayesian inference for a covariance matrix," in *Conf. Applied Statistics in Agriculture*, 2014.