

# A Bayesian Approach for the Multifractal Analysis of Spatio-Temporal Data

S. Combrexelle, H. Wendt, J.-Y. Tourneret  
IRIT - ENSEEIHT, CNRS,  
Univ. of Toulouse, France

firstname.lastname@irit.fr

Y. Altmann, S. McLaughlin  
School of Engineering, Heriot-Watt Univ.,  
Edinburgh, Scotland

initial.lastname@hw.ac.uk

P. Abry  
CNRS, Physics Dept.,  
ENS Lyon, France

patrice.abry@ens-lyon.fr

**Abstract**—Multifractal (MF) analysis enables the theoretical study of scale invariance models and their practical assessment via wavelet leaders. Yet, the accurate estimation of MF parameters remains a challenging task. For a range of applications, notably biomedical, the performance can potentially be improved by taking advantage of the multivariate nature of data. However, this has barely been considered in the context of MF analysis. This paper proposes a Bayesian model that enables the joint estimation of MF parameters for multivariate time series. It builds on a recently introduced statistical model for leaders and is formulated using a 3D gamma Markov random field joint prior for the MF parameters of the voxels of spatio-temporal data, represented as a multivariate time series, that counteracts the statistical variability induced by small sample size. Numerical simulations indicate that the proposed Bayesian estimator significantly outperforms current state-of-the-art algorithms.

## I. CONTEXT, RELATED WORK AND CONTRIBUTIONS

**Context.** Multifractal analysis is a widely used signal processing tool and enables the study of the scale invariance properties of data. It has been successfully used in a large variety of applications, ranging from biomedical [1], [2], physics [3] to finance [4] and Internet [5], cf., e.g., [6] for a review. Scale invariance implies that the dynamics of a time series  $X(t)$  are driven by a large continuum of time scales instead of only a few characteristic scales. This translates into power law behaviors of the time averages of well chosen multiresolution quantities  $T_X(j, k)$  of  $X$  (i.e., quantities that depend jointly on scale  $2^j$  and time instance  $k$ ) over a large range of scales  $2^j$

$$S(q, j) \triangleq \frac{1}{n_j} \sum_k |T_X(j, k)|^q \simeq (2^j)^{\zeta(q)}, \quad 2^{j_1} \leq 2^j \leq 2^{j_2} \quad (1)$$

where  $n_j$  is the number of  $T_X(j, k)$  at scale  $j$ . In this work, wavelet leaders  $l(j, k)$  are used as multiresolution quantities, which can be shown to be well suited for this purpose [6], [7] (and are defined in Section II). Given a time series  $X$ , the goal is the estimation of the so-called *scaling exponents*  $\zeta(q)$  of the power law in (1), which fully characterize the scale invariance properties of  $X$ . Notably, the scaling exponents permit discrimination between the two fundamental classes of scale invariance models: self-similar processes, which are obtained by *additive* construction mechanisms and characterized by a linear function  $\zeta(q) = qH$  [8]; multifractal multiplicative cascades (MMC), which have a *multiplicative*

structure and yield a strictly concave function  $\zeta(q)$  [4]. In order to understand the construction mechanisms underlying data, it is crucial to decide which model better fits the data in applications. This decision can be reached by considering the polynomial development  $\zeta(q) = \sum_{m \geq 1} c_m q^m / m!$  at  $q = 0$ . It can be shown that the coefficient  $c_2$ , called the *multifractality parameter*, is strictly negative for MMC but equals zero for self-similar processes, cf., e.g., [7], [9]. Therefore, the estimation of  $c_2$  is central in multifractal analysis.

**Estimation of  $c_2$ .** The multifractality parameter  $c_2$  can be directly linked to the variance of the logarithm of  $l(j, k)$  [9]

$$C_2(j) \triangleq \text{Var} [\ln l(j, \cdot)] = c_2^0 + c_2 \ln 2^j. \quad (2)$$

This motivates estimation of  $c_2$  as a linear regression of the sample variance  $\widehat{\text{Var}}[\cdot]$  of log-leaders with respect to scale  $j$

$$\hat{c}_2 = (\ln 2)^{-1} \sum_{j=j_1}^{j_2} w_j \widehat{\text{Var}} [\ln l(j, \cdot)] \quad (3)$$

where  $w_j$  are appropriate regression weights [7]. The estimator (3) is widely used but is known to yield poor performance for small sample size [10], [11]. Alternative estimators have been described in, e.g., [12], [13], but they make assumptions, (e.g. fully parametric model, specific multifractal process), that are often too restrictive in real-world applications. More recently, Bayesian estimators for  $c_2$  have been proposed in [10], [11]. Their advantage lies in the use of a semi-parametric model for the statistics of the log-leaders that is generically valid for MMC processes and induces considerable performance gains when compared to (3). These gains were obtained at the price of increased computational cost since the Bayesian inference was achieved by a Markov chain Monte Carlo (MCMC) algorithm with a Metropolis-Hastings within Gibbs (MHG) sampler. A significantly more efficient algorithm was obtained very recently in [14] by considering a data augmented formulation of the Bayesian model. None of these developments addressed the estimation of  $c_2$  for multivariate data.

**Contributions.** This paper devises a Bayesian procedure for the joint estimation of  $c_2$  associated with multivariate time series registered on a volume (voxels) that makes use of the dependence of neighboring voxels in order to improve estimation accuracy. The algorithm combines the statistical model introduced in [10], [11] with the data augmentation strategy proposed for images in [14] (summarized in Section II). The

key contribution (described in Section III) resides in the design of an appropriate joint prior for the multifractality parameters for voxels. It consists of a hidden 3D gamma Markov random field (GMRF) [15] with eight-fold spatial neighborhood that models the dependence between the parameters of neighboring voxels. The Bayesian model is designed in such a way that the conditional distributions of the resulting joint posterior can be sampled without MHG steps. Consequently, the approximation of the associated Bayesian estimator by means of an MCMC algorithm is very efficient (inducing approximately 10 times only the overall cost of estimation based on (3)). Numerical simulations with synthetic multifractal data demonstrate that the proposed method reduces standard deviations as compared to the linear regression (3) by more than one order of magnitude and permits, for the first time, the accurate assessment of small differences of the values of  $c_2$  associated with voxels of multivariate time series (cf., Section IV).

## II. STATISTICAL MODEL FOR LOG-LEADERS

### A. Direct statistical model in the time-domain

**Wavelet leaders.** A mother wavelet  $\psi_0(t)$  is a reference pattern that has narrow supports in the time and frequency domains. It is chosen such that the collection  $\{\psi_{j,k}(t) \equiv 2^{-j/2}\psi_0(2^{-j}t - k), j \in \mathbb{N}, k \in \mathbb{N}\}$  forms a basis of  $L^2(\mathbb{R})$  and is characterized by its number of vanishing moments  $N_\psi \geq 1$  ( $\forall k = 0, 1, \dots, N_\psi - 1, \int_{\mathbb{R}} t^k \psi_0(t) dt \equiv 0$  and  $\int_{\mathbb{R}} t^{N_\psi} \psi_0(t) dt \neq 0$ ). The ( $L^1$ -normalized) discrete wavelet transform coefficients of  $X$  are defined as  $d_X(j, k) = \langle X, 2^{-j/2}\psi_{j,k} \rangle$ , cf., e.g., [16] for details.

Let  $\lambda_{j,k} = [k2^j, (k+1)2^j)$  denote the dyadic interval of size  $2^j$  and  $3\lambda_{j,k}$  the union of  $\lambda_{j,k}$  with its 2 neighbors. The wavelet leaders are defined as the largest wavelet coefficient within  $3\lambda_{j,k}$  over all finer scales [6], [7]

$$l(j, k) \triangleq \sup_{\lambda' \subset 3\lambda_{j,k}} |d_X(\lambda')|. \quad (4)$$

**Statistical model.** Denote as  $\ell_j$  the vector of the log-leaders  $\ell(j, \cdot) \triangleq \ln l(j, \cdot)$  at scale  $j$  after mean subtraction (since it conveys no information on  $c_2$ ) and  $\boldsymbol{\ell} \triangleq [\ell_{j_1}^T, \dots, \ell_{j_2}^T]^T$ . The statistics of  $\ell_j$  of MMC based processes can be well approximated by a multivariate Gaussian distribution whose covariance  $C_j(k, \Delta k) \triangleq \text{Cov}[\ell(j, k), \ell(j, k + \Delta k)]$  is [10]

$$C_j(k, \Delta k) \approx \varrho_j^0(\Delta k; \boldsymbol{\theta}) + \varrho_j^1(\Delta k; \boldsymbol{\theta}) \quad (5)$$

where  $\boldsymbol{\theta} = (c_2, c_2^0)$ ,  $\varrho_j^1(r; \boldsymbol{\theta}) \triangleq c_2 \ln(4|r|/n_j) \mathbb{I}_{(3, n_j/4)}(r)$ ,  $\varrho_j^0(r; \boldsymbol{\theta}) \triangleq \left(\frac{\ln(1+|r|)}{\ln 4}\right)(\varrho_j^1(3; \boldsymbol{\theta}) - c_2^0 - c_2 \ln 2^j) + c_2^0 + c_2 \ln 2^j \mathbb{I}_{(0, 3]}(r)$  and where  $\mathbb{I}_A$  is the indicator function of the set  $A$ . Assuming independence between  $\ell_j$  at different scales  $j$  leads to the following likelihood for  $\boldsymbol{\ell}$

$$p(\boldsymbol{\ell}|\boldsymbol{\theta}) = \prod_{j=j_1}^{j_2} |\Sigma_{j,\boldsymbol{\theta}}|^{-\frac{1}{2}} \exp\left(-\frac{1}{2}\boldsymbol{\ell}_j^T \Sigma_{j,\boldsymbol{\theta}}^{-1} \boldsymbol{\ell}_j\right) \quad (6)$$

where the matrices  $\Sigma_{j,\boldsymbol{\theta}}$  are defined element-wise by (5),  $|\cdot|$  is denoting the determinant and  $^T$  the transpose operator.

**Whittle approximation.** The likelihood (6) is problematic to evaluate numerically since it requires the computation of the matrix inverses  $\Sigma_{j,\boldsymbol{\theta}}^{-1}$ . Thus, it has been proposed in [11]

to approximate (6) with the asymptotic Whittle likelihood [17]

$$p_W(\boldsymbol{\ell}|\boldsymbol{\theta}) = |\Gamma_{\boldsymbol{\theta}}|^{-1} \exp(-\mathbf{y}^H \Gamma_{\boldsymbol{\theta}}^{-1} \mathbf{y}), \quad (7)$$

$$\mathbf{y} \triangleq [\mathbf{y}_{j_1}^T, \dots, \mathbf{y}_{j_2}^T]^T, \quad \mathbf{y}_j = \mathcal{F}(\ell_j)$$

where  $\Gamma_{\boldsymbol{\theta}} \triangleq c_2 \mathbf{F} + c_2^0 \mathbf{G}$  is an  $N_Y \times N_Y$  diagonal covariance matrix, with  $N_Y \triangleq \text{card}(\mathbf{y})$ ,  $\mathbf{F} \triangleq \text{diag}(\mathbf{f})$ ,  $\mathbf{G} \triangleq \text{diag}(\mathbf{g})$ ,  $\mathbf{f} \triangleq [\mathbf{f}_{j_1}^T, \dots, \mathbf{f}_{j_2}^T]^T$  and  $\mathbf{g} \triangleq [\mathbf{g}_{j_1}^T, \dots, \mathbf{g}_{j_2}^T]^T$ . The diagonal elements of  $\Gamma_{\boldsymbol{\theta}}$  correspond to the discretized spectral densities  $c_2 \mathbf{f}_j(m) + c_2^0 \mathbf{g}_j(m)$  associated with the model (5), for the positive frequencies  $\omega_m = 2\pi m / \sqrt{n_j}$ ,  $m \in \mathbb{N}^+$ . Here,  $\mathbf{y}_j \triangleq \mathcal{F}(\ell_j)$  is the periodogram of  $\ell_j$ , where the operator  $\mathcal{F}(\cdot)$  computes and vectorizes the discrete Fourier transform coefficients for  $\omega_m$ ,  $m \in \mathbb{N}^+$  and  $^H$  is the conjugate transpose operator. Note that  $\mathbf{f}_j$  and  $\mathbf{g}_j$  do not depend on  $\boldsymbol{\theta}$  and can be precomputed (and stored) using the fast Fourier transform.

### B. Data augmented statistical model in the Fourier domain

The parameters  $\boldsymbol{\theta}$  are encoded in  $\Sigma_{j,\boldsymbol{\theta}}^{-1}$ , and their conditional distributions are not standard. Sampling the posterior distribution with an MCMC method would hence require accept/reject procedures [10], [11]. A more efficient algorithm can be obtained by interpreting (7) as a statistical model for the Fourier coefficients  $\mathbf{y}$  [14]. Assuming that  $\Gamma_{\boldsymbol{\theta}}$  is positive definite, (7) amounts to modeling  $\mathbf{y}$  by a random vector with a centered circular-symmetric complex Gaussian distribution  $\mathcal{CN}(\mathbf{0}, \Gamma_{\boldsymbol{\theta}})$ , hence to the use of the likelihood

$$p(\mathbf{y}|\boldsymbol{\theta}) = |\Gamma_{\boldsymbol{\theta}}|^{-1} \exp(-\mathbf{y}^H \Gamma_{\boldsymbol{\theta}}^{-1} \mathbf{y}). \quad (8)$$

**Reparametrization.** The matrix  $\Gamma_{\boldsymbol{\theta}}$  is positive definite as long as the parameters  $\boldsymbol{\theta} = (c_2, c_2^0)$  belong to the admissible set

$$\mathcal{A} = \{\boldsymbol{\theta} \in \mathbb{R}_*^- \times \mathbb{R}_*^+ \mid c_2 \mathbf{f}(m) + c_2^0 \mathbf{g}(m) > 0, m = 1, \dots, N_Y\}. \quad (9)$$

Since  $\forall m, c_2^0 \mathbf{g}(m) > 0$  (while  $c_2 \mathbf{f}(m) < 0$ ), (9) can be mapped onto independent positivity constraints by a one-to-one transformation from  $\boldsymbol{\theta} \in \mathcal{A}$  to  $\mathbf{v} \in \mathbb{R}_*^{+2}$  defined as  $\boldsymbol{\theta} \mapsto \mathbf{v} = (v_1, v_2) \triangleq (-c_2, c_2^0/\gamma + c_2)$ , where  $\gamma = \sup_m \mathbf{f}(m)/\mathbf{g}(m)$  [14]. Consequently, (8) can be expressed using  $\mathbf{v} \in \mathbb{R}_*^{+2}$  as

$$p(\mathbf{y}|\mathbf{v}) \propto |\Gamma_{\mathbf{v}}|^{-1} \exp(-\mathbf{y}^H \Gamma_{\mathbf{v}}^{-1} \mathbf{y}) \quad (10)$$

$$\Gamma_{\mathbf{v}} = v_1 \tilde{\mathbf{F}} + v_2 \tilde{\mathbf{G}}, \quad \tilde{\mathbf{F}} = -\mathbf{F} + \mathbf{G}\gamma, \quad \tilde{\mathbf{G}} = \mathbf{G}\gamma$$

where  $\tilde{\mathbf{F}}$ ,  $\tilde{\mathbf{G}}$  and  $\Gamma_{\mathbf{v}}$  are positive definite diagonal matrices.

**Data augmentation.** One can now introduce an  $N_Y \times 1$  vector of latent variables  $\boldsymbol{\mu}$  that enables us to augment (10) using the model  $\mathbf{y}|\boldsymbol{\mu}, v_2 \sim \mathcal{CN}(\boldsymbol{\mu}, v_2 \tilde{\mathbf{G}})$ ,  $\boldsymbol{\mu}|v_1 \sim \mathcal{CN}(0, v_1 \tilde{\mathbf{F}})$ , which is associated with the extended likelihood [14]

$$p(\mathbf{y}, \boldsymbol{\mu}|\mathbf{v}) \propto v_2^{-N_Y} \exp\left(-v_2^{-1}(\mathbf{y} - \boldsymbol{\mu})^H \tilde{\mathbf{G}}^{-1}(\mathbf{y} - \boldsymbol{\mu})\right)$$

$$\times v_1^{-N_Y} \exp\left(-v_1^{-1} \boldsymbol{\mu}^H \tilde{\mathbf{F}}^{-1} \boldsymbol{\mu}\right). \quad (11)$$

Simple calculations show that (11) leads to standard conditional distributions when inverse-gamma priors are used for  $v_i$ ,  $i = 1, 2$ , and that (10) is recovered by marginalization of (11) with respect to  $\boldsymbol{\mu}$ .

### III. BAYESIAN MODEL FOR MULTIVARIATE TIME SERIES

#### A. Likelihood

Based on the likelihood (11) for one single time series  $X(t)$ , we now design a joint Bayesian model for the analysis of multivariate time series. Let  $X_m$ ,  $\mathbf{m} \triangleq (m_1, m_2, m_3)$ ,  $m_d = 1, \dots, M_d$ , denote  $M_1 \times M_2 \times M_3$  discrete time series (voxels) of length  $N$  (as illustrated in Fig. 1). Denote as  $\mathbf{y}_m$ ,  $\boldsymbol{\mu}_m$  and  $\mathbf{v}_m$  the Fourier coefficients, latent variables and parameter vector associated with  $X_m$  and as  $\mathbf{Y} \triangleq \{\mathbf{y}_m\}$ ,  $\mathbf{M} \triangleq \{\boldsymbol{\mu}_m\}$ , and  $\mathbf{V} \triangleq \{\mathbf{V}_1, \mathbf{V}_2\}$  (where  $\mathbf{V}_i \triangleq \{v_{i,m}\}$ ,  $i = 1, 2$ ) the corresponding collections for all voxels  $\{X_m\}$ . Assuming independence between the vectors  $\mathbf{y}_m$ , the joint likelihood of  $\mathbf{Y}$  can be written as

$$p(\mathbf{Y}, \mathbf{M} | \mathbf{V}) \propto \prod_m p(\mathbf{y}_m, \boldsymbol{\mu}_m | \mathbf{v}_m). \quad (12)$$

#### B. Gamma Markov random field prior

Inverse-gamma distributions  $\mathcal{IG}(\alpha_{i,m}, \beta_{i,m})$  are conjugate priors for the parameters  $v_{i,m}$  in (12), and we propose to specify  $(\alpha_{i,m}, \beta_{i,m})$  such that the resulting prior for  $\mathbf{V}_i$  is a hidden GMRF [15]. A GMRF makes use of a set of positive auxiliary variables  $\mathbf{Z} = \{\mathbf{Z}_1, \mathbf{Z}_2\}$ ,  $\mathbf{Z}_i = \{z_{i,m}\}$ , to induce positive dependence between the neighboring elements of  $\mathbf{V}_i$  (and thus spatial regularization) [15]. Specifically, each  $v_{i,m}$  is connected to the eight auxiliary variables  $z_{i,m'} > 0$ ,  $\mathbf{m}' \in \mathcal{V}_v(\mathbf{m}) \triangleq \{\mathbf{m} + (d_1, d_2, d_3)\}_{d_1, d_2, d_3=0,1}$  (and therefore, each  $z_{i,m}$  to  $v_{i,m'}$ ,  $\mathbf{m}' \in \mathcal{V}_z(\mathbf{m}) \triangleq \{\mathbf{m} + (d_1, d_2, d_3)\}_{d_1, d_2, d_3=-1,0}$ , via edges with weights  $\rho_i$ ,  $i = 1, 2$ , that are hyperparameters and control the amount of smoothness. It can be shown that this prior for  $(\mathbf{V}_i, \mathbf{Z}_i)$  is associated with the density [15]

$$p(\mathbf{V}_i, \mathbf{Z}_i | \rho_i) \propto \prod_k e^{(8\rho_i - 1) \log z_{i,m}} e^{-(8\rho_i + 1) \log v_{i,m}} \times e^{-\frac{\rho_i}{v_{i,m}} \sum_{\mathbf{m}' \in \mathcal{V}_v(\mathbf{m})} z_{i,m'}}. \quad (13)$$

#### C. Posterior distribution and Bayesian estimators

Under the assumption of prior independence between  $(\mathbf{V}_1, \mathbf{Z}_1)$  and  $(\mathbf{V}_2, \mathbf{Z}_2)$ , the joint posterior distribution associated with the proposed model is obtained as

$$p(\mathbf{V}, \mathbf{Z}, \mathbf{M} | \mathbf{Y}, \rho_1, \rho_2) \propto p(\mathbf{Y} | \mathbf{V}_2, \mathbf{M}) p(\mathbf{M} | \mathbf{V}_1) \times p(\mathbf{V}_1, \mathbf{Z}_1 | \rho_1) p(\mathbf{V}_2, \mathbf{Z}_2 | \rho_2) \quad (14)$$

using Bayes' theorem. To infer the parameters of interest  $\mathbf{V}_i$ , we consider the marginal posterior mean (minimum mean square error) estimator, denoted MMSE, which is defined as  $\mathbf{V}_i^{\text{MMSE}} \triangleq \mathbb{E}[\mathbf{V}_i | \mathbf{Y}, \rho_i]$ , where the expectation is taken with respect to the marginal posterior density  $p(\mathbf{V}_i | \mathbf{Y}, \rho_i)$ . The direct computation of  $\mathbf{V}_i^{\text{MMSE}}$  is not tractable since it requires integrating the posterior (14) over the variables  $\mathbf{Z}$  and  $\mathbf{M}$  and computation of the expectation. Instead, by considering a Gibbs sampler (GS) generating samples  $(\{\mathbf{V}_i^{(q)}\}, \mathbf{M}^{(q)}, \{\mathbf{Z}_i^{(q)}\})_{q=0}^{N_{mc}}$  that are asymptotically distributed according to (14), it can be approximated as [18]

$$\mathbf{V}_i^{\text{MMSE}} \approx (N_{mc} - N_{bi})^{-1} \sum_{q=N_{bi}}^{N_{mc}} \mathbf{V}_i^{(q)} \quad (15)$$

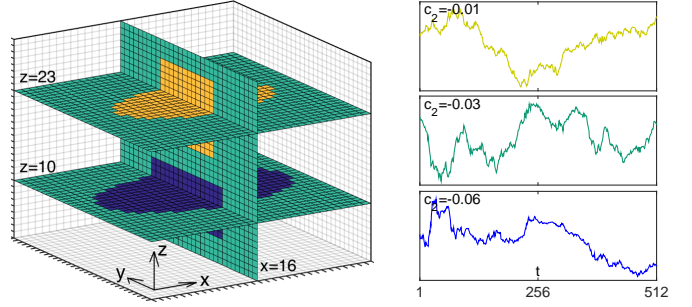


Fig. 1. Illustration of the cube of  $32 \times 32 \times 32$  voxels of time series (left panel) with prescribed multifractal properties  $c_2 \in \{-0.01, -0.03, -0.06\}$  (indicated as green, yellow, dark blue, respectively); the slices correspond to those analyzed in Fig. 2. Single realizations of time series corresponding to 3 voxels with different value of  $c_2$  (right panel).

	LF	IG	GMRF
b	0.0158	0.0051	0.0092
std	0.0800	0.0255	0.0020
rmse	0.0819	0.0262	0.0094

TABLE I

ESTIMATION PERFORMANCE FOR 100 INDEPENDENT REALIZATIONS.

where  $N_{bi}$  is the number of samples of the burn-in period.

#### D. Gibbs sampler

Here, the GS consists of successively drawing samples from the conditional distributions that are associated with the posterior (14) [18]. Simple calculations lead to

$$p(\boldsymbol{\mu} | \mathbf{Y}, \mathbf{V}) \sim \mathcal{CN}(\mathbf{v}_1 \tilde{\mathbf{F}} \mathbf{T}_v^{-1} \mathbf{y}, ((\mathbf{v}_1 \tilde{\mathbf{F}})^{-1} + (\mathbf{v}_2 \tilde{\mathbf{G}})^{-1})^{-1}) \quad (16a)$$

$$p(v_i | \mathbf{Y}, \mathbf{M}, \mathbf{Z}_i) \sim \mathcal{IG}(N_Y + \alpha_i, \Xi_i + \beta_i) \quad (16b)$$

$$p(z_i | \mathbf{V}_i) \sim \mathcal{G}(\alpha_i, \gamma_i) \quad (16c)$$

where the subscript  $\mathbf{m}$  has been omitted for notational convenience and where  $\Xi_1 = \|\boldsymbol{\mu}\|_{\tilde{\mathbf{F}}^{-1}}$ ,  $\Xi_2 = \|\mathbf{y} - \boldsymbol{\mu}\|_{\tilde{\mathbf{G}}^{-1}}$  with  $\|\mathbf{x}\|_{\Pi} \triangleq \mathbf{x}^H \Pi \mathbf{x}$ ,  $\alpha_{i,m} = 8\rho_i$ ,  $\beta_{i,m} = \rho_i \sum_{\mathbf{m}' \in \mathcal{V}_v(\mathbf{m})} z_{i,m'}$  and  $\gamma_{i,m} = (\rho_i \sum_{\mathbf{m}' \in \mathcal{V}_z(\mathbf{m})} v_{i,m'}^{-1})^{-1}$ . All conditionals (16a–16c) are *standard* laws that can be sampled efficiently, without MHG steps. Finally, note that when the parameters  $v_{i,m}$  are assumed to be independent and have  $\mathcal{IG}(c_i, d_i)$  priors instead of (13) (i.e., no smooth spatial evolution is assumed), a Bayesian model is obtained that can also be sampled using the GS steps (16a–16b), with  $\alpha_{i,m} = c_i$  and  $\beta_{i,m} = d_i$ .

### IV. NUMERICAL EXPERIMENTS

We compare the performance of the proposed estimator (denoted as GMRF) with its counterpart with an  $\mathcal{IG}$  prior (denoted as IG) and with the linear regression estimator (3) (denoted as LF, with weights as in [7]) by applying it to 100 independent realizations of a cube of  $32^3$  voxels of length  $N = 512$ . Each voxel is an independent realization of MRW, with prescribed values  $c_2 \in \{-0.01, -0.03, -0.06\}$  as illustrated in Fig. 1. MRW belongs to the class of MMC processes and possesses multifractal properties similar to those of Mandelbrot's multiplicative log-normal cascades, with scaling exponents  $\zeta(q) = (H - c_2)q + c_2 q^2 / 2$ , cf., [19] for details

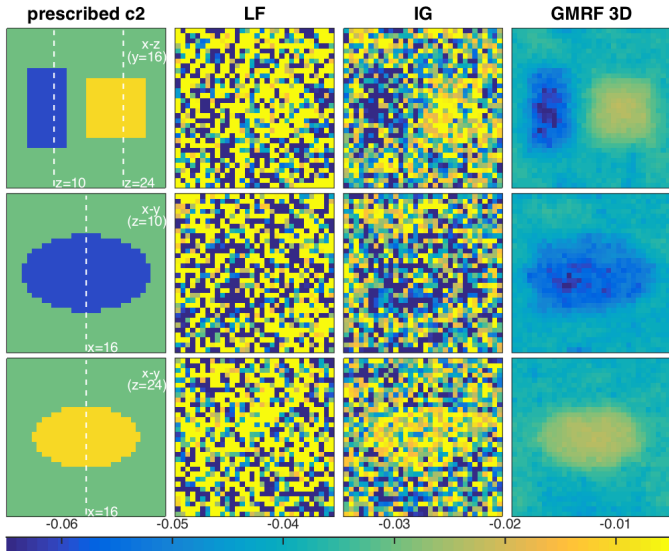


Fig. 2. Estimation of  $c_2$  for one single realization: ground truth (1st column) and estimates obtained using LF, IG and GMRF (2nd to 4th column, respectively) for the 3 slices shown in Fig. 1 (from top to bottom).

( $H = 0.72$  for the results presented below). The regularization parameters have been fixed a priori using cross-validation.

**Illustration for a single realization.** Fig. 2 displays estimates obtained for one single realization using LF, IG and GMRF (2nd to 4th row, respectively) for the slices  $x = 16$ ,  $z = 10$  and  $z = 23$  shown in Fig. 1 and yields the following conclusions. First, the LF estimator completely fails to reveal the existence of two zones of voxels with constant  $c_2 \in \{-0.03, -0.06\}$  in the background of voxels with  $c_2 = -0.01$ . The IG estimator improves estimation accuracy with respect to the LF such that the three groups of voxels can be evidenced visually, but its variability is too large to permit accurate identification of the voxels sharing the same value for  $c_2$ . In contrast, the GMRF estimator yields excellent estimates that accurately capture the geometry of the three zones of voxels and the corresponding values for  $c_2$ .

**Performance.** The estimation performance for  $c_2$  is quantified via the bias, standard deviation and root mean squared error defined as  $b = \mathbb{E}[\hat{c}_2] - c_2$ ,  $\text{std} = (\widehat{\text{Var}}[\hat{c}_2])^{\frac{1}{2}}$  and  $\text{rmse} = (b^2 + \text{std}^2)^{\frac{1}{2}}$ , respectively. The results are given in Table I and confirm the above conclusions. While IG reduces std values to  $\frac{1}{3}$  of those of LF, GMRF further and dramatically reduces std values to  $\frac{1}{40}$  of those of LF, which clearly demonstrates the effectiveness of the proposed joint Bayesian model. The bias is found to be small but non-negligible for all three methods (largest for LF and smallest for IG). As a result, gains in rmse values for GMRF are smaller but still significant (rmse values of GMRF are one order of magnitude below those of LF).

## V. CONCLUSIONS

This paper has proposed a Bayesian procedure for the joint estimation of  $c_2$  for spatio-temporal data (voxels). The Bayesian model is composed of a data-augmented Whittle

likelihood for log-leaders and a GMRF joint prior for the multifractality parameters for voxels and yields a significant improvement in estimation performance for each voxel (by more than one order of magnitude when compared to linear regression). Moreover, it is designed in such a way that the associated estimators can be approximated efficiently by an MCMC algorithm. The proposed joint estimator enables, for the first time, the reliable assessment of small differences of  $c_2$  for voxels with as little as  $N = 512$  samples. Future work will include incorporation of the regularization parameters  $\rho_i$  in the model and allowing them to differ for each voxel in order to permit abrupt changes in the data while maintaining strong smoothing within zones of voxels with constant  $c_2$ .

## REFERENCES

- [1] A. L. Goldberger, L. A. Amaral, J. M. Hausdorff, P. C. Ivanov, C. K. Peng, and H. E. Stanley, "Fractal dynamics in physiology: alterations with disease and aging," *Proc. Natl. Acad. Sci. USA*, vol. 99, no. Suppl 1, pp. 2466–2472, 2002.
- [2] P. Ciuciu, P. Abry, and B. J. He, "Interplay between functional connectivity and scale-free dynamics in intrinsic fMRI networks," *Neuroimage*, vol. 95, pp. 248–263, 2014.
- [3] 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.
- [4] B. B. Mandelbrot, "A multifractal walk down Wall Street," *Sci. Am.*, vol. 280, no. 2, pp. 70–73, 1999.
- [5] P. Abry, R. Baraniuk, P. Flandrin, R. Riedi, and D. Veitch, "Multiscale nature of network traffic," *IEEE Signal Proces. Mag.*, vol. 19, no. 3, pp. 28–46, 2002.
- [6] S. Jaffard, P. Abry, and H. Wendt, "Irregularities and scaling in signal and image processing: Multifractal analysis," in *Benoit Mandelbrot: A Life in Many Dimensions*, M. Frame, Ed. World Scientific Pub., 2015.
- [7] H. Wendt, P. Abry, and S. Jaffard, "Bootstrap for empirical multifractal analysis," *IEEE Signal Proces. Mag.*, vol. 24, no. 4, pp. 38–48, 2007.
- [8] B. B. Mandelbrot and J. W. van Ness, "Fractional Brownian motion, fractional noises and applications," *SIAM Review*, vol. 10, pp. 422–437, 1968.
- [9] B. Castaing, Y. Gagne, and M. Marchand, "Log-similarity for turbulent flows," *Physica D*, vol. 68, no. 3–4, pp. 387–400, 1993.
- [10] H. Wendt, N. Dobigeon, J.-Y. Tourneret, and P. Abry, "Bayesian estimation for the multifractality parameter," in *Proc. ICASSP*, Vancouver, Canada, May 2013.
- [11] S. Combexelle, H. Wendt, P. Abry, N. Dobigeon, S. McLaughlin, and J.-Y. Tourneret, "A Bayesian approach for the joint estimation of the multifractality parameter and integral scale based on the Whittle approximation," in *Proc. ICASSP*, Brisbane, Australia, April 2015.
- [12] T. Lux, "The Markov-switching multifractal model of asset returns," *J. Business & Economic Stat.*, vol. 26, no. 2, pp. 194–210, 2008.
- [13] O. Løvstetten and M. Rypdal, "Approximated maximum likelihood estimation in multifractal random walks," *Phys. Rev. E*, vol. 85, pp. 046705, 2012.
- [14] S. Combexelle, H. Wendt, Y. Altmann, J.-Y. Tourneret, S. McLaughlin, and P. Abry, "A Bayesian framework for the multifractal analysis of images using data augmentation and a Whittle approximation," in *Proc. ICASSP*, Shanghai, China, March 2016.
- [15] O. Dikmen and A.T. Cemgil, "Gamma Markov random fields for audio source modeling," *IEEE T. Audio, Speech, and Language Proces.*, vol. 18, no. 3, pp. 589–601, 2010.
- [16] S. Mallat, *A Wavelet Tour of Signal Processing*, Academic Press, San Diego, CA, 1998.
- [17] P. Whittle, "On stationary processes in the plane," *Biometrika*, vol. 41, pp. 434–449, 1954.
- [18] C. P. Robert and G. Casella, *Monte Carlo Statistical Methods*, Springer, New York, USA, 2005.
- [19] E. Bacry, J. Delour, and J.-F. Muzy, "Multifractal random walk," *Phys. Rev. E*, vol. 64: 026103, 2001.