BAYESIAN ESTIMATION FOR THE LOCAL ASSESSMENT OF THE MULTIFRACTALITY PARAMETER OF MULTIVARIATE TIME SERIES

S. Combrexelle[1], *H. Wendt*[1], *Y. Altmann*[2], *J.-Y. Tourneret*[1], *S. McLaughlin*[2], *P. Abry*[3]

[1] IRIT - ENSEEIHT, CNRS, Univ. of Toulouse, F-31062 Toulouse, France, firstname.lastname@enseeiht.fr

[2] School of Engineering and Physical Sciences, Heriot-Watt Univ., Edinburgh, UK, initial.lastname@hw.ac.uk

[3] CNRS, Physics Dept., Ecole Normale Supérieure de Lyon, F-69364 Lyon, France, patrice.abry@ens-lyon.fr

ABSTRACT

Multifractal analysis (MF) is a widely used signal processing tool that enables the study of scale invariance models. Classical MF assumes *homogeneous* MF properties, which cannot always be guaranteed in practice. Yet, the *local* estimation of MF parameters has barely been considered due to the challenging statistical nature of MF processes (non-Gaussian, intricate dependence), requiring large sample sizes. This present work addresses this limitation and proposes a Bayesian estimator for local MF parameters of multivariate time series. The proposed Bayesian model builds on a recently introduced statistical model for leaders (i.e., specific multiresolution quantities designed for MF analysis purposes) that enabled the Bayesian estimation of MF parameters and extends it to multivariate non-overlapping time windows. It is formulated using spatially smoothing gamma Markov random field priors that counteract the large statistical variability of estimates for short time windows. Numerical simulations demonstrate that the proposed algorithm significantly outperforms current state-of-the-art estimators.

Index Terms— Multifractal analysis, Bayesian estimation, Multivariate time series, Whittle likelihood, GMRF

1. INTRODUCTION

Context. The paradigm of scale invariance has proven relevant in a large variety of applications involving real-world data of very different natures, cf., e.g., [1] and references therein. Scale invariance models assume that the temporal dynamics of data are generated at a large continuum of time scales (instead of a few particular scales that could thus play a privileged role in the analysis). This is revealed via the power law behaviors, over a range of scales 2^j, of the sample moments of suitable multiresolution quantities $T_X(j,k)$ of a time series X

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

This work was supported by ANR BLANC 2011 AMATIS BS0101102.
S. Combrexelle was supported by the Direction Générale de l'Armement (DGA).
SML acknowledges the support of EPSRC via grant EP/J015180/1.

i.e., quantities that depend jointly on scale 2^j and time instance k (where $n_j = \text{card}(T_X(k,\cdot))$ is the number of $T_X(j,k)$ at scale j). The relations between different scales are thus characterized by the *scaling exponents* $\zeta(q)$, and the goal is to infer them from a given time series X. In this work, *wavelet leaders* $l(j,k)$, which can be shown to be well adapted in this context [1, 2], are used as multiresolution quantities (and defined in Section 2). Multifractal analysis consists of a specific instance of scale invariance analysis (cf. e.g., [1]). It notably permits discrimination between two fundamental classes of scale invariance models: self-similar models that are characterized by a linear function $\zeta(q) = qH$ and are tied to *additive* construction mechanisms [3]; multifractal multiplicative cascades (MMC), translating into a strictly concave $\zeta(q)$ and characterized by a *multiplicative* structure [4]. The decision on which model is better adapted for data is therefore fundamental for understanding the underlying data production mechanisms. A Taylor expansion of $\zeta(q)$ at the origin, i.e., $\zeta(q) = \sum_{m\ge1} c_m q^m/m!$, is useful for this discrimination, since it can be shown that the second coefficient c_2, called the intermittency or *multifractality parameter*, equals zero for self-similar processes but is strictly negative for MMC, cf., e.g., [2, 5].

Local estimation of c_2. Eq. (1) implicitly assumes that the scaling properties are *homogeneous* (since the sample moments correspond to the entire time series X). Thus, when c_2 varies along time, the analysis must be localized to small time intervals (time series segments). It can be shown that c_2 is directly related to the variance of the logarithm of $l(j,k)$ [5]

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

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

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

where w_j are appropriate regression weights [2]. This simple estimator is commonly used in applications but is known to yield prohibitively large variance for small sample size [6, 7].

978-0-9928-6265-7/16/$31.00 ©2016 IEEE

Estimation procedures with better performance were described in, e.g., [8, 9]. However, these estimators make use of fully parametric models or are designed for specific multifractal processes, which is often too restrictive for real-world data analysis. Alternatively, Bayesian estimators of c_2 for time series have been studied in [6, 7]. They rely on a semi-parametric model for the statistics of the log-leaders that is generically valid for MMC processes. Estimating the parameters of this model can be achieved by a Markov chain Monte Carlo (MCMC) algorithm with a Metropolis-Hastings within Gibbs (MHG) sampler [6, 7, 10]. However, very recently, a novel data-augmented Whittle likelihood based formulation was proposed for the estimation of c_2 for images [11], which yields a significantly more efficient algorithm.

Contributions. This paper studies a Bayesian procedure that makes use of the dependence of neighboring components of multivariate time series enabling the estimation of the time evolution of c_2. The method is developed here using the example of time series organized on a planar grid (corresponding, e.g., to measurements from a planar sensor array). The procedure makes use of an adaptation for time series of the statistical model for images introduced in [11] (described in Section 2) and relies on the following original key ingredients (detailed in Section 3). First, the local analysis of each component of the multivariate time series is formulated via non-overlapping time windows and the joint likelihood of all windows is expressed as the product of the augmented likelihood of each individual time window. Then, a joint prior for the multifractality parameters associated with the different windows is assigned, using a hidden gamma Markov random field (GaMRF) [12] with eight-fold temporal-spatial elementary cells and modeling the dependence between the parameters of neighboring (in space and time) windows. The design of the proposed Bayesian model is such that the conditional distributions of the resulting joint posterior can be sampled without MHG steps, and the associated Bayesian estimator can be approximated efficiently by means of an MCMC algorithm. The performance of the proposed estimator of parameters c_2 associated with short time intervals of multivariate time series is assessed by numerical simulations conducted with synthetic multifractal data (cf., Section 4). The proposed method significantly outperforms the linear regression (3), reducing standard deviation by up to one order of magnitude, and permits, for the first time, to accurately assess the time-evolution of c_2 of multivariate time series.

2. STATISTICAL MODEL FOR LOG-LEADERS

2.1. Time-domain statistical model

Wavelet leaders. A mother wavelet $\psi_0(t)$ is a reference pattern with narrow supports in both time and frequency domains that is characterized by its number of vanishing moments $N_\psi \geq 1$ ($\forall k = 0, 1, \ldots, N_\psi - 1$, $\int_{\mathbb{R}} t^k \psi_0(t) dt = 0$ and $\int_{\mathbb{R}} t^{N_\psi} \psi_0(t) dt \neq 0$). It is chosen such that the collection $\{\psi_{j,k}(t) \equiv 2^{-j/2}\, \psi_0(2^{-j}t - k), j \in \mathcal{N}, k \in \mathcal{N}\}$ forms a basis of $L^2(\mathbb{R})$. The discrete wavelet transform coefficients of X are defined as $d_X(j,k) = \langle X, \psi_{j,k} \rangle$, cf., e.g., [13] for further 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 [1, 2]

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

Statistical model. We denote by $\boldsymbol{\ell}_j$ the vector of all log-leaders $\ell(j,\cdot) \triangleq \ln l(j,\cdot)$ at scale j after subtracting the mean (which does not convey any information on c_2). It can be shown that the statistics of $\boldsymbol{\ell}_j$ of MMC-based processes can be well approximated by a multivariate Gaussian distribution whose covariance $\mathcal{C}_j(k,\Delta k) \triangleq \mathrm{Cov}[\ell(j,k), \ell(j,k+\Delta k)]$ is [6]

$$\mathcal{C}_j(k,\Delta k) \approx \varrho_j(\Delta k;\boldsymbol{\theta}) \triangleq \begin{cases} \varrho_j^0(|\Delta k|;\boldsymbol{\theta}) & |\Delta k| \leq 3 \\ \varrho_j^1(|\Delta k|;\boldsymbol{\theta}) & 3 < |\Delta k| \end{cases} \tag{5}$$

where $\boldsymbol{\theta} = (c_2, c_2^0)$, $\varrho_j^1(r;\boldsymbol{\theta}) \triangleq c_2 \ln(4r/n_j) \mathbb{I}_{[0,n_j/4]}(r)$, $\mathbb{I}_A$ is the indicator function of the set A and $\varrho_j^0(r;\boldsymbol{\theta}) \triangleq \frac{\ln(1+r)}{\ln 4}(\varrho_j^1(3;\boldsymbol{\theta}) - c_2^0 - c_2 \ln 2^j) + c_2^0 + c_2 \ln 2^j$. Furthermore, independence is assumed between different scales j, which leads to the following likelihood for $\boldsymbol{\ell} \triangleq [\boldsymbol{\ell}_{j_1}^T, \ldots, \boldsymbol{\ell}_{j_2}^T]^T$

$$p(\boldsymbol{\ell}|\boldsymbol{\theta}) \propto \prod_{j=j_1}^{j_2} |\boldsymbol{\Sigma}_{j,\boldsymbol{\theta}}|^{-\frac{1}{2}} \exp\Big(-\frac{1}{2} \boldsymbol{\ell}_j^T \boldsymbol{\Sigma}_{j,\boldsymbol{\theta}}^{-1} \boldsymbol{\ell}_j \Big) \tag{6}$$

where $\propto$ means "proportional to" and where the elements of $\boldsymbol{\Sigma}_{j,\boldsymbol{\theta}}$ are defined by $\varrho_j(|\Delta k|;\boldsymbol{\theta})$, $|\cdot|$ is denoting the determinant and ^{T} the transpose operator.

Whittle approximation. The evaluation of the above likelihood is problematic even for moderate sample size since it requires computing the matrix inverses $\boldsymbol{\Sigma}_{j,\boldsymbol{\theta}}^{-1}$. Therefore, it has been proposed in [7] to approximate (6) with an asymptotic Whittle likelihood [14], which can be written as

$$p_W(\boldsymbol{\ell}|\boldsymbol{\theta}) \propto |\boldsymbol{\Gamma}_{\boldsymbol{\theta}}|^{-1} \exp\left(-\boldsymbol{y}^H \boldsymbol{\Gamma}_{\boldsymbol{\theta}}^{-1} \boldsymbol{y}\right), \tag{7}$$
$$\boldsymbol{y} \triangleq [\boldsymbol{y}_{j_1}^T, ..., \boldsymbol{y}_{j_2}^T]^T, \quad \boldsymbol{y}_j = \mathcal{F}(\boldsymbol{\ell}_j).$$

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

2.2. Data augmented Fourier domain statistical model

The parameter vector $\boldsymbol{\theta}$ is encoded in $\boldsymbol{\Sigma}_{j,\boldsymbol{\theta}}^{-1}$, and its conditional distribution is not standard. Sampling the posterior distribution with an MCMC method would hence require accept/reject procedures [6, 7]. A more efficient algorithm can be obtained by interpreting (7) as a statistical model for the Fourier coefficients $\boldsymbol{y}$ [11]. Assuming that $\boldsymbol{\Gamma}_{\boldsymbol{\theta}}$ is positive definite, (7) amounts to modeling $\boldsymbol{y}$ by a random vector with a centered circular-symmetric complex Gaussian distribution $\mathcal{CN}(\boldsymbol{0}, \boldsymbol{\Gamma}_{\boldsymbol{\theta}})$, hence the use of the likelihood

$$p(\boldsymbol{y}|\boldsymbol{\theta}) \propto |\boldsymbol{\Gamma}_{\boldsymbol{\theta}}|^{-1} \exp\left(-\boldsymbol{y}^H \boldsymbol{\Gamma}_{\boldsymbol{\theta}}^{-1} \boldsymbol{y}\right). \tag{8}$$

Reparametrization. The matrix $\boldsymbol{\Gamma}_{\boldsymbol{\theta}}$ is positive definite when the parameters $\boldsymbol{\theta} = (c_2, c_2^0)$ belong to the admissible set $\mathcal{A} = \{\boldsymbol{\theta} \in \mathbb{R}_\star^- \times \mathbb{R}_\star^+ | c_2 \mathbf{f}(\mathrm{m}) + c_2^0 \mathbf{g}(\mathrm{m}) > 0, \mathrm{m} = 1,..., N_Y\}$. Since $\forall \mathrm{m}, c_2^0 \mathbf{g}(\mathrm{m}) > 0$ (while $c_2 \mathbf{f}(\mathrm{m}) < 0$), the set $\mathcal{A}$ can be mapped onto independent positivity constraints by a one-to-one transformation from $\boldsymbol{\theta} \in \mathcal{A}$ to $\boldsymbol{v} \in \mathbb{R}_\star^{+2}$ defined as

$$\boldsymbol{\theta} \mapsto \boldsymbol{v} = (v_1, v_2) \triangleq (-c_2, c_2^0/\gamma + c_2)$$

where $\gamma = \sup_{\mathrm{m}} \mathbf{f}(\mathrm{m})/\mathbf{g}(\mathrm{m})$ [11]. Consequently, (8) can be expressed using $\boldsymbol{v} \in \mathbb{R}_\star^{+2}$ as

$$\begin{aligned} p(\boldsymbol{y}|\boldsymbol{v}) \propto & |\boldsymbol{\Gamma}_{\boldsymbol{v}}|^{-1} \exp\left(-\boldsymbol{y}^H \boldsymbol{\Gamma}_{\boldsymbol{v}}^{-1} \boldsymbol{y}\right) \\ & \boldsymbol{\Gamma}_{\boldsymbol{v}} = v_1 \tilde{\boldsymbol{F}} + v_2 \tilde{\boldsymbol{G}}, \quad \tilde{\boldsymbol{F}} = -\boldsymbol{F} + \tilde{\boldsymbol{G}}, \quad \tilde{\boldsymbol{G}} = \boldsymbol{G}\gamma \end{aligned} \tag{9}$$

with positive definite diagonal matrices $\tilde{\boldsymbol{F}}$, $\tilde{\boldsymbol{G}}$ and $\boldsymbol{\Gamma}_{\boldsymbol{v}}$.

Data augmentation. The introduction of an $N_Y \times 1$ vector of latent variables $\boldsymbol{\mu}$ allows the following augmented model to be defined, $\boldsymbol{y}|\boldsymbol{\mu}, v_2 \sim \mathcal{CN}(\boldsymbol{\mu}, v_2 \tilde{\boldsymbol{G}})$, $\boldsymbol{\mu}|v_1 \sim \mathcal{CN}(0, v_1 \tilde{\boldsymbol{F}})$, which is associated with the extended likelihood [11]

$$\begin{aligned} p(\boldsymbol{y}, \boldsymbol{\mu}|\boldsymbol{v}) \propto & {v_2}^{-N_Y} \exp\Big(-v_2^{-1}(\boldsymbol{y}-\boldsymbol{\mu})^H \tilde{\boldsymbol{G}}^{-1}(\boldsymbol{y}-\boldsymbol{\mu})\Big) \\ & \times {v_1}^{-N_Y} \exp\Big(-v_1^{-1} \boldsymbol{\mu}^H \tilde{\boldsymbol{F}}^{-1} \boldsymbol{\mu}\Big). \end{aligned} \tag{10}$$

One easily verifies that marginalization of (10) with respect to $\boldsymbol{\mu}$ yields (9) and that (10) leads to standard conditional distributions when inverse-gamma priors are used for $\boldsymbol{v}$.

3. LOCAL BAYESIAN ESTIMATION FOR MULTIVARIATE TIME SERIES

3.1. Likelihood

We formulate a joint Bayesian model for the time-localized analysis of multivariate time series based on the likelihood (10) for one single time series $X(t)$. Let $X_{\boldsymbol{m}}$, $\boldsymbol{m} \triangleq (m_1, m_2)$, $m_i = 1, \ldots, M_i$, denote $M_1 \times M_2$ discrete time series of length N that have been simultaneously recorded at positions $\boldsymbol{m}$ on an equally-spaced planar grid as illustrated in Fig. 1. Each time series $X_{\boldsymbol{m}}$ is sub-divided, using n_S non-overlapping windows of length $L = N/n_S$, into segments $X_{(\boldsymbol{m},n)} = (X_{\boldsymbol{m}}(k))_{k=(n-1)L+1}^{nL}$, $n = 1, \ldots, n_S$. Denote as $\boldsymbol{y}_{(\boldsymbol{m},n)}$, $\boldsymbol{\mu}_{(\boldsymbol{m},n)}$ and $\boldsymbol{v}_{(\boldsymbol{m},n)}$ the Fourier coefficients, latent variables and parameter vector associated with $X_{(\boldsymbol{m},n)}$ and as $\boldsymbol{Y} \triangleq \{\boldsymbol{y}_{(\boldsymbol{m},n)}\}$, $\boldsymbol{M} \triangleq \{\boldsymbol{\mu}_{(\boldsymbol{m},n)}\}$, and $\boldsymbol{V} \triangleq \{\boldsymbol{V}_1, \boldsymbol{V}_2\}$ (where $\boldsymbol{V}_i \triangleq \{v_{i,(\boldsymbol{m},n)}\}$, $i = 1, 2$) the corresponding collections of parameters for all segments $\{X_{(\boldsymbol{m},n)}\}$. Under the assumptions of Section 2, the joint likelihood of $\boldsymbol{Y}$ is given by

$$p(\boldsymbol{Y}, \boldsymbol{M}|\boldsymbol{V}) \propto \prod\nolimits_{\boldsymbol{m},n} p(\boldsymbol{y}_{(\boldsymbol{m},n)}, \boldsymbol{\mu}_{(\boldsymbol{m},n)}|\boldsymbol{v}_{(\boldsymbol{m},n)}). \tag{11}$$

3.2. Gamma Markov random field prior

Inverse-gamma distributions $\mathcal{IG}(\alpha_{i,(\boldsymbol{m},n)}, \beta_{i,(\boldsymbol{m},n)})$ are conjugate priors for the parameters $v_{i,(\boldsymbol{m},n)}$ in (11). We propose to specify $(\alpha_{i,(\boldsymbol{m},n)}, \beta_{i,(\boldsymbol{m},n)})$ such that the resulting prior for $\boldsymbol{V}_i$ is a hidden GaMRF [12]. A GaMRF relies on the use of a set of positive auxiliary variables $\boldsymbol{Z} = \{\boldsymbol{Z}_1, \boldsymbol{Z}_2\}$, $\boldsymbol{Z}_i = \{z_{i,(\boldsymbol{m},n)}\}$, to induce positive dependence between the neighbooring elements of $\boldsymbol{V}_i$ [12] and hence spatial/temporal regularization. More precisely, each $v_{i,(\boldsymbol{m},n)}$ is connected to its eight neighboring auxiliary variables $z_{i,(\boldsymbol{m}',n')} > 0$, $(\boldsymbol{m}', n') \in \mathcal{V}_v((\boldsymbol{m}, n)) \triangleq \{((m_1 + i_1, m_2 + i_2), n + j)\}_{i_1,i_2,j=0,1}$ (and thus, each $z_{i,(\boldsymbol{m},n)}$ to $v_{i,(\boldsymbol{m}',n')}$, $\boldsymbol{m}' \in \mathcal{V}_z((\boldsymbol{m}, n)) \triangleq \{((m_1 + i_1, m_2 + i_2), n + j)\}_{i_1,i_2,j=-1,0}$), via edges weighted by r_i. Here, the weights r_i, $i = 1, 2$, are hyperparameters that control the amount of smoothness. This GaMRF can be shown to be associated with the density [12]

$$\begin{aligned} p(\boldsymbol{V}_i, \boldsymbol{Z}_i|r_i) = & C(r_i)^{-1} \prod\nolimits_{\boldsymbol{m},n} \mathrm{e}^{(8r_i - 1)\log z_{i,(\boldsymbol{m},n)}} \\ & \times \mathrm{e}^{-(8r_i+1)\log v_{i,(\boldsymbol{m},n)}} \mathrm{e}^{-\frac{r_i}{v_{i,(\boldsymbol{m},n)}} \sum_{(\boldsymbol{m}',n') \in \mathcal{V}_v(\boldsymbol{m})} z_{i,(\boldsymbol{m}',n')}} \end{aligned} \tag{12}$$

which is used here as a joint prior for $(\boldsymbol{V}_i, \boldsymbol{Z}_i)$, where $C(r_i)$ is a normalization constant.

Posterior. Assuming prior independence between $(\boldsymbol{V}_1, \boldsymbol{Z}_1)$ and $(\boldsymbol{V}_2, \boldsymbol{M}, \boldsymbol{Z}_2)$, the joint posterior distribution associated with the proposed model is obtained using Bayes' theorem

$$\begin{aligned} p(\boldsymbol{V}, \boldsymbol{Z}, \boldsymbol{M}|\boldsymbol{Y}, r_1, r_2) \propto & \; p(\boldsymbol{Y}|\boldsymbol{V}_2, \boldsymbol{M})\, p(\boldsymbol{M}|\boldsymbol{V}_1) \\ & \times p(\boldsymbol{V}_1, \boldsymbol{Z}_1|r_1)\, p(\boldsymbol{V}_2, \boldsymbol{Z}_2|r_2). \end{aligned} \tag{13}$$

3.3. Bayesian estimators

We consider the marginal posterior mean estimator for the parameters of interest $\boldsymbol{V}_i$, denoted MMSE (minimum mean square error) estimator and defined as $\boldsymbol{V}_i^{\mathrm{MMSE}} \triangleq \mathbb{E}[\boldsymbol{V}_i|\boldsymbol{Y}, r_i]$, where the expectation is taken with respect to the marginal posterior density $p(\boldsymbol{V}_i|\boldsymbol{Y}, r_i)$. The direct computation of $\boldsymbol{V}_i^{\mathrm{MMSE}}$ is not tractable since it requires integrating (13) over the variables $\boldsymbol{Z}$ and $\boldsymbol{M}$. Here, we consider a Gibbs sampler (GS) drawing samples $(\{\boldsymbol{V}_i^{(q)}\}, \boldsymbol{M}^{(q)}, \{\boldsymbol{Z}_i{}^{(k)}\})_{q=0}^{N_{mc}}$ that are asymptotically distributed according to (13). An approximation of $\boldsymbol{V}_i^{\mathrm{MMSE}}$ can then be obtained as [10]

$$\widehat{\boldsymbol{V}}_i^{\mathrm{MMSE}} \approx (N_{mc} - N_{bi})^{-1} \sum\nolimits_{q=N_{bi}}^{N_{mc}} \boldsymbol{V}_i^{(q)} \tag{14}$$

where N_{bi} is the length of the burn-in period.

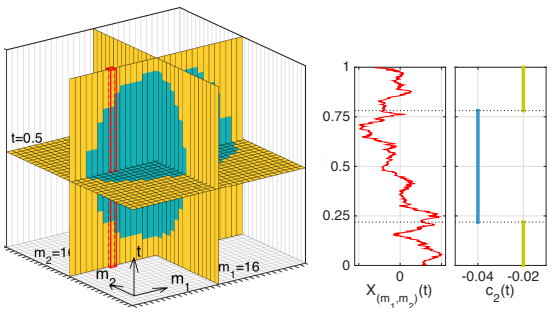

Fig. 1. Prescribed time evolution of c_2 for 32×32 time series $X_{\boldsymbol{m}}(t)$ (left panel). A single realization for the time series at position $(m_1, m_2) = (16, 16)$ highlighted in red on the left panel and the corresponding c_2 profile (right panel).

3.4. Gibbs sampler

The GS considered in this paper successively generates samples from the conditional distributions associated with the posterior (13) [10] which are easily calculated and given by

$$\boldsymbol{\mu} \,|\, \boldsymbol{Y}, \boldsymbol{V} \sim \mathcal{CN}\Big(v_1 \tilde{\boldsymbol{F}} \boldsymbol{\Gamma}_{\boldsymbol{v}}^{-1} \boldsymbol{y}, \Big((v_1 \tilde{\boldsymbol{F}})^{-1} + (v_2 \tilde{\boldsymbol{G}})^{-1}\Big)^{-1}\Big) \quad (15a)$$

$$v_1 \,|\, \boldsymbol{M}, \boldsymbol{Z}_1 \sim \mathcal{IG}\big(N_Y + \alpha_1, ||\boldsymbol{\mu}||_{\tilde{\boldsymbol{F}}^{-1}} + \beta_1\big) \quad (15b)$$

$$v_2 \,|\, \boldsymbol{Y}, \boldsymbol{M}, \boldsymbol{Z}_2 \sim \mathcal{IG}\big(N_Y + \alpha_2, ||\boldsymbol{y} - \boldsymbol{\mu}||_{\tilde{\boldsymbol{G}}^{-1}} + \beta_2\big) \quad (15c)$$

$$z_i \,|\, \boldsymbol{V}_i \sim \mathcal{G}(\tilde{\alpha}_i, \tilde{\beta}_i) \quad (15d)$$

where we have omitted the subscript $_{(\boldsymbol{m},n)}$ for notational convenience and where $||\mathbf{x}||_\Pi \triangleq \mathbf{x}^H \Pi \mathbf{x}$, $\alpha_{i,(\boldsymbol{m},n)} = \tilde{\alpha}_{i,(\boldsymbol{m},n)} = 8 r_i$, $\beta_{i,(\boldsymbol{m},n)} = r_i \sum_{(\boldsymbol{m}',n') \in \mathcal{V}_v((\boldsymbol{m},n))} z_{i,(\boldsymbol{m}',n')}$, $\tilde{\beta}_{i,(\boldsymbol{m},n)} = (r_i \sum_{(\boldsymbol{m}',n') \in \mathcal{V}_z((\boldsymbol{m},n))} v_{i,(\boldsymbol{m}',n')}^{-1})^{-1}$ and $\mathcal{G}$ is the gamma distribution. Note that all conditionals (15a–15d) are standard distributions that can be sampled efficiently, without Metropolis-Hastings steps. Moreover, when independence is assumed between the parameters $v_{i,(\boldsymbol{m},n)}$, that have $\mathcal{IG}(c_i, d_i)$ priors instead of (12), a Bayesian model without spatial/temporal regularization as in [11] is obtained that can also be sampled using the GS steps defined in (15a–15c) with parameters given by $\alpha_{i,(\boldsymbol{m},n)} = c_i$ and $\beta_{i,(\boldsymbol{m},n)} = d_i$.

4. SIMULATION RESULTS

The proposed procedure, denoted GaMRF, was applied to 50 independent realizations of 32×32 multifractal random walks (MRW) of length $N = 2^{14}$ with prescribed time-evolution of c_2. It was compared to its counterpart with $\mathcal{IG}$ priors for v_i (denoted IG) and to linear regression (3) (denoted LF). MRW is an MMC process whose multifractal properties mimic those of Mandelbrot's multiplicative log-normal cascades and has scaling exponents $\zeta(q) = (H - c_2)q + c_2 q^2/2$, see [15] for more details. We set $H = 0.72$ and prescribe shifts from $c_2 = -0.02$ to $c_2 = -0.04$ around $t = 0.5$ for the time series located in an ellipsoid at the center of the (m_1, m_2) plane ($c_2 = -0.02$ is constant outside this ellipsoid) as shown in Fig. 1. This piece-wise constant evolution for c_2 was chosen as a limit test case for GaMRF (which is designed for smooth evolutions). The cases of $n_S = 2^W$, $W = 2, \ldots, 6$, windows (resulting in window lengths $L = \{2^{12}, 2^{11}, 2^{10}, 2^9, 2^8\}$, respectively) were investigated, using $j_1 = 2$ and j_2 the largest available scale. The hyperparameters r_i were fixed a priori for each window size independently by preliminary visual inspection of estimates obtained for a single realization.

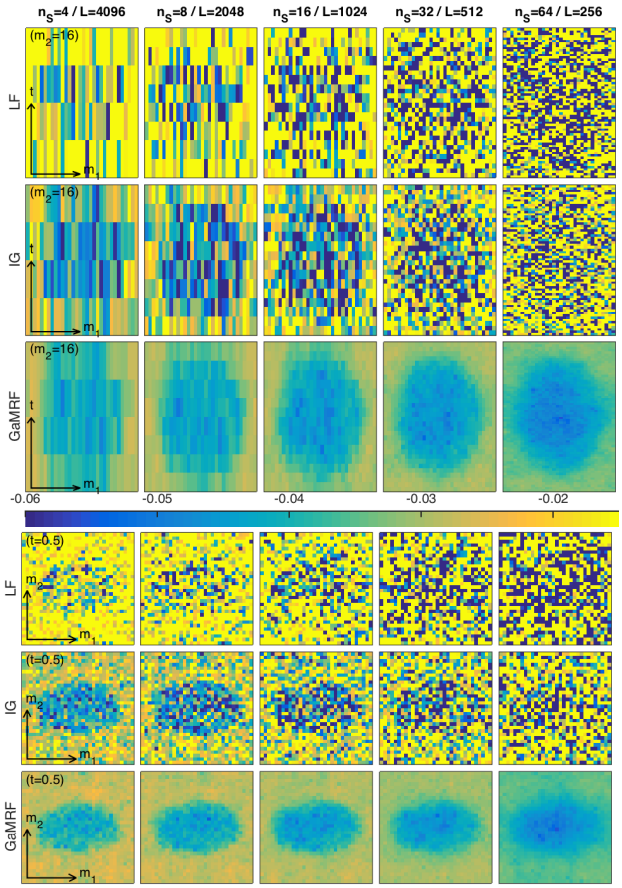

Fig. 2. Time-localized estimation of c_2 for a single realization of 32×32 time series of length $N = 2^{14}$ with time-varying c_2 as plotted in Fig. 1. From left to right: decreasing window size L. Top panel: estimates in the (m_1, t) plane for $m_2 = 16$ fixed. Bottom panel: Estimates in the (m_1, m_2) plane with $t = 0.5$ fixed. The estimates in rows 1 to 3 of the panels are obtained with LF, IG and GaMRF, respectively.

Illustration for a single realization. Fig. 2 plots estimates obtained with $n_S = \{4, 8, 16, 32, 64\}$ (left to right column, respectively) in the $(m_1, m_2 = 16, t)$ plane (hence illustrating the time resolution capability; top panel) and in the $(m_1, m_2, t{=}1/2)$ plane (i.e., illustrating the spatial coherence across time series; bottom panel) for LF, IG and GaMRF (rows 1 to 3, respectively). Clearly, LF fails to provide relevant local estimates of c_2 for any window size L (due to large variance). The Bayesian estimator IG improves the estimation accuracy with respect to LF and allows us to identify,

Table 1. RMSE values for different window sizes L (the lower, the better; best results are marked in bold).

n_S / L	4 / 2^{12}	8 / 2^{11}	16 / 2^{10}	32 / 2^9	64 / 2^8
LF	0.020	0.026	0.037	0.058	0.102
IG	0.011	0.013	0.018	0.024	0.036
GaMRF	**0.008**	**0.008**	**0.009**	**0.009**	**0.013**

at least visually, the spatial/temporal c_2 profile for intermediate window sizes $L = \{2048, 1024\}$. Yet, for smaller windows ($L < 1024$), the variability of IG is too large, and the time evolutions of c_2 cannot be finely resolved. In contrast, GaMRF provides excellent estimates of the evolution of c_2 for any of the considered window sizes. Even for the smallest window ($L = 256$), local estimates obtained with GaMRF accurately reproduce the time evolution of c_2 and the spatial coherence across time series.

Estimation performance. The estimation performance for c_2 is quantified via the root mean squared error defined as $\text{RMSE} = \left((\widehat{\mathbb{E}}[\hat{c}_2] - c_2)^2 + \widehat{\text{Var}}[\hat{c}_2]\right)^{\frac{1}{2}}$, which is given in Tab. 1 as a function of the number of used time windows (the bias is found to be comparable for all three estimators). The RMSE values of LF and IG are dominated by the variance for $L \leq 2048$ and accordingly scale as $L^{-\frac{1}{2}}$. A comparison of the three estimators yields the following conclusions. First, IG decreases RMSE values to $\frac{1}{2}$ to $\frac{1}{3}$ times those of LF, with larger performance gains for small window size L. Second, GaMRF further and dramatically decreases RMSE values to as little as $\frac{1}{8}$ times those of LF (for $L = 256$). Finally, the RMSE values of GaMRF increase only slowly with decreasing window size L, thus enabling the accurate local estimation of c_2 even for small time windows.

5. CONCLUSIONS AND PERSPECTIVES

This work proposed a Bayesian model for the local estimation of c_2 (in non-overlapping time windows) for multivariate time series in a plane. It relies on the use of an extended Whittle likelihood for log-leaders and a GaMRF joint prior for the parameters for each windowed time series that counteracts the variance increase induced by short time windows. The procedure significantly improves the estimation performance as compared to linear regression and enables, for the first time, reliable and accurate assessment of small changes of c_2 along time, at the price of moderately increased computational cost (by a factor 10 as compared to linear regression). Future work will include incorporation of the regularization parameters r_i in the model and extensions of the model to other types of multivariate data, e.g., volumetric time series (voxels in 3D).

6. REFERENCES

[1] 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.

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

[3] B. B. Mandelbrot and J. W. van Ness, "Fractional Brownian motion, fractional noises and applications," *SIAM Review*, vol. 10, pp. 422–437, 1968.

[4] B. B. Mandelbrot, "A multifractal walk down Wall Street," *Sci. Am.*, vol. 280, no. 2, pp. 70–73, 1999.

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

[6] H. Wendt, N. Dobigeon, J.-Y. Tourneret, and P. Abry, "Bayesian estimation for the multifractality parameter," in *Proc. ICASSP*, Vancouver, Canada, May 2013.

[7] S. Combrexelle, 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.

[8] T. Lux, "The Markov-switching multifractal model of asset returns," *J. Business & Economic Stat.*, vol. 26, no. 2, pp. 194–210, 2008.

[9] O. Løvsletten and M. Rypdal, "Approximated maximum likelihood estimation in multifractal random walks," *Phys. Rev. E*, vol. 85, pp. 046705, 2012.

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

[11] S. Combrexelle, 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.

[12] O. Dikmen and A.T. Cemgil, "Gamma markov random fields for audio source modeling," *Audio, Speech, and Language Processing, IEEE Transactions on*, vol. 18, no. 3, pp. 589–601, March 2010.

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

[14] P. Whittle, "On stationary processes in the plane," *Biometrika*, vol. 41, pp. 434–449, 1954.

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