Multifractal Analysis of Multivariate Images Using Gamma Markov Random Field Priors*

Herwig Wendt†, Sébastien Combrexelle†, Yoann Altmann‡, Jean-Yves Tourneret†, Stephen McLaughlin‡, and Patrice Abry§

Abstract. Texture characterization of natural images using the mathematical framework of multifractal (MF) analysis, enables the study of the fluctuations in the regularity of image intensity. Although successfully applied in various contexts, the use of MF analysis has so far been limited to the independent analysis of a single image, while the data available in applications are increasingly multivariate. This paper addresses this limitation and proposes a joint Bayesian model and associated estimation procedure for MF parameters of multivariate images. It builds on a recently introduced generic statistical model that enabled the Bayesian estimation of MF parameters for a single image and relies on the following original key contributions: First, we develop a novel Fourier domain statistical model for a single image that permits the use of a likelihood that is separable in the MF parameters via data augmentation. Second, a joint Bayesian model for multivariate images is formulated in which prior models based on gamma Markov random fields encode the assumption of the smooth evolution of MF parameters between the image components. The design of the likelihood and of conjugate prior models is such that exploitation of the conjugacy between the likelihood and prior models enables an efficient estimation procedure that can handle a large number of data components. Numerical simulations conducted using sequences of multifractal images demonstrate that the proposed procedure significantly outperforms previous univariate benchmark formulations at a competitive computational cost.

Key words. Texture Analysis, Multifractal Analysis, Multivariate Images, Wavelet Leaders, Bayesian Estimation, Gamma Markov Random Field

AMS subject classifications. 65T60, 62H12, 62M40, 68U10, 62H35, 62F15, 62L20

1. Introduction.

1.1. Context. Texture is a perceptual attribute and different paradigms have been proposed in the literature for its characterization [24]. There is a growing body of work suggesting that a large class of natural images are modelled well by scale invariant processes [10, 55, 61], motivating the use of random fractals, scale invariance or self-similarity to characterize texture. These concepts can be linked to the degree of point-wise singular behavior or *regularity* (smoothness) of the image amplitude [37, 44]. From this point of view, texture models can

*Submitted to the editors 6 October 2017.

Funding: Work supported by Grant ANR-16-CE33-0020 MultiFracs and Grant ANR-12-BS03-003 Hypanema. S. Combrexelle was supported by the Direction Générale de l'Armement (DGA) under the DGA/DSTL collaborative PhD programme. S. McLaughlin acknowledges the support of EPSRC grant number EP/J015180/1 and the supply of the hyperspectral image data from DSTL.

†IRIT-INPT, CNRS, University of Toulouse, France (herwig.wendt@irit.fr, sebastian.combrexelle@gmail.com, jean-yves.tourneret@enseeiht.fr).

‡School of Engineering and Physical Sciences, Heriot-Watt University, Edinburgh, UK (y.altmann@hw.ac.uk, s.mclaughlin@hw.ac.uk).

§Laboratoire de Physique, Université de Lyon, ENS de Lyon, Université Claude Bernard, CNRS, F-69342 Lyon, France (patrice.abry@ens-lyon.fr).

be seen as densely interwoven sets of singularities of different regularity strengths, commonly measured by the Hölder exponent h [26, 44]. The suitable mathematical framework for the study of these models is *multifractal analysis* (MFA) [4, 32], which enables scale invariant image texture to be described via the spatial fluctuations of the point-wise regularity exponents of the image. More precisely, it provides a global characterization of the texture via the so-called *multifractal spectrum* $D(h)$, collecting the Hausdorff dimensions of the sets of positions that share the same regularity exponent. This paper is focused on one of the central parameters in MFA, the *multifractality* or *intermittency* parameter c_2, which is related to the width of the multifractal spectrum (hence, to the degree of fluctuations of regularity exponents in the image). It enables discrimination between the two major classes of scale invariant processes used in applications: (additive construction-based) self-similar processes for which $c_2 = 0$, and (multiplicative construction-based) processes constructed from multifractal multiplicative cascade (MMC) for which c_2 is strictly negative [56]. The parameter c_2 therefore highlights departures from Gaussian marginals as well as changes in the local transient dependence structure of texture. The corresponding practical analysis tools rely on the scale dependence of the statistics of certain multi-resolution quantities (such as increments, wavelet coefficients or, more recently, wavelet leaders [26], which will be used here). The main concepts of multifractal analysis are briefly recalled in Section 2, cf. [26, 27, 44] for details.

1.2. Multifractal analysis of multivariate images. MFA has been successfully used in a number of applications including texture classification [55, 61], biomedical imaging [5, 31, 33], physics [41, 48], biology [50], climate research [34] and art investigation [1, 13, 30]. Yet, its application remains so far conceptually limited to the analysis of single images with homogeneous scale invariant properties [29]. The main reason for this resides in the definition of the multifractal spectrum, which is intrinsically univariate. Although definitions of a multivariate multifractal spectrum have been studied, cf., e.g., [23, 38, 49] and the recent contributions [28, 57], these attempts remain essentially limited to pairs of time series and are of little relevance for M-tuples of data for $M \gg 2$. This limitation has become increasingly urgent in view of the number of recent applications in which the acquired images are *multivariate*, i.e., they consist of a set of images (multitemporal, multispectral, multimodal,...) or spatially organized collections of image patches. Such data can provide a rich resource for information, on condition that they are analyzed *jointly* rather than individually [42].

1.3. Related work: Estimation of c_2. The current existing estimation procedures for the multifractality parameter are limited to the independent processing of individual images. The standard estimator for c_2 is based on a simple linear regression of the sample variance of the logarithm of suitable multiresolution quantities over several analysis scales [9] (cf., (5)). This estimator is appealing for its simplicity and low computational cost, yet it has been reported to suffer from poor performance when applied to small-sized images [58]. A related approach, based on a wavelet scattering transform, was recently studied in [8]. Several parametric model-based approaches have also been proposed. These include maximum likelihood methods [6, 11, 39, 60] and the generalized method of moments [35]. Yet, their definitions are tied to specific instances of self-similar or multifractal processes and the use of fully parametric models is often too restrictive in practice. More recently, a Bayesian estimation framework was studied in [17]. It relied on a flexible semi-parametric model for the statistics of the

logarithm of wavelet leaders that is generically valid for self-similar and MMC processes. This model showed excellent estimation performance. However, its high computational cost made it difficult to apply to the analysis of multivariate data, and a first attempt at its direct application with a simultaneous autoregressive smoothing prior was reported in [18]. However, it was limited to the analysis of small image sequences for this reason.

1.4. Goals and contributions. The goal of the present work is to go beyond these existing, inherently univariate, estimation procedures for c_2 and propose the first operational approach for the multifractal analysis of multivariate images. Specifically, we propose to conduct the analysis within a multivariate Bayesian model that *jointly* describes the collection of multifractality parameters c_2 associated with the multifractal spectra of different individual data components. This Bayesian multivariate estimation framework builds upon ingredients of the model proposed in [17], which is recalled in Section 3.1. However, it also presents some essential new characteristics relying on the following two main contributions.

First, a novel Fourier domain statistical model for log-leaders is proposed. This model allows conjugate inverse-gamma ($\mathcal{IG}$) prior distributions to be defined. Consequently, it leads to efficient inference procedures that are appropriate for multivariate data. It dwells on three original key ingredients, described in Section 3.2: 1) The Whittle approximation [6, 53] is used to build an original Gaussian model for the Fourier coefficients of log-leaders of MMC processes, 2) The implicit joint constraint on the multifractal parameters is decomposed into independent positivity constraints through a suitable reparametrization, 3) Data augmentation is used in order to define an augmented distribution whose parameters are easier to estimate.

Second, we use this novel model in the formulation of a joint Bayesian model for multivariate images (cf. Section 4). Specifically, we propose to encode the prior belief that the multifractality parameter evolves slowly across time or spectral bands (for sequence of images) or throughout space (for image patches) through the design of a gamma Markov random field (GaMRF) prior. This GaMRF prior induces positive correlation between the multifractality parameters associated with different image components and hence regularizes estimation. This leads to simple conditional distributions for the parameters of the augmented Fourier domain model, namely $\mathcal{IG}$ distributions. As a result, the inference of the parameters of the associated posterior distribution (described in Section 5) can be conducted by a Gibbs sampler whose steps do not require accept-reject moves, leading to an efficient estimation procedure adapted to a large number of unknowns. This estimation framework elaborates on the approach described in [16], which was limited to the local estimation of c_2 in that it applied only to non-overlapping image patches and could be applied to a single image one at a time, two significant restrictions when application to real-world data is envisaged, as is the case here.

The performance of the method is assessed numerically in Section 6 using sequences of synthetic multifractal images with piece-wise constant (in time and space) multifractal properties. The proposed joint estimator yields significantly improved estimation performance compared to previous univariate formulations with a competitive computational cost (root-mean-squared error values are more than one order of magnitude smaller than those of standard linear regression, with a 4 times larger computational cost). Finally, Section 7 illustrates the benefits of the proposed procedure for the analysis of real-world data *via* two applications devoted to

hyperspectral and multitemporal imaging. Supplementary 3D animations associated with the results obtained in Sections 6 and 7, as well as MATLAB codes, are available online (see [14]).

2. Multifractal Analysis of images.

2.1. Local regularity and multifractal spectrum. Denote as $X(\mathbf{y}) : \mathbf{y} \in \mathbb{R}^2 \to X \in \mathbb{R}$ the image under analysis, where $\mathbf{y} = (y_1, y_2)$ stands for the spatial variable. The purpose of MFA is to characterize the fluctuations of the *point-wise regularity* of the image X. This regularity is most commonly measured using the *Hölder exponent* [26, 44], defined as follows. Assuming that X is locally bounded[1], X is said to belong to $C^\alpha(\mathbf{y}_0)$ at position $\mathbf{y}_0$ if there exist $\alpha > 0$ and a polynomial $P_{\mathbf{y}_0}$ of degree smaller than α such that $\|X(\mathbf{y}) - P_{\mathbf{y}_0}(\mathbf{y})\| \leq C\|\mathbf{y} - \mathbf{y}_0\|^\alpha$ for $\mathbf{y}$ sufficiently close to $\mathbf{y}_0$, with $\|\cdot\|$ denoting the Euclidian norm. The Hölder exponent is then defined as

$$h(\mathbf{y}_0) \triangleq \sup\{\alpha : X \in C^\alpha(\mathbf{y}_0)\} \tag{1}$$

with, qualitatively, the smaller (resp. larger) $h(\mathbf{y}_0)$, the rougher (resp. smoother) X at $\mathbf{y}_0$. MFA provides a *global* description of the fluctuations of $h(\mathbf{y})$ in space in terms of the *multifractal spectrum* $D(h)$. The multifractal spectrum is defined as the Hausdorff dimension of the sets of positions $\mathbf{y}$ that have identical Hölder exponent

$$\mathcal{D}(h) \triangleq \dim_H \{\mathbf{y} : h(\mathbf{y}) = h\}. \tag{2}$$

The estimation of $D(h)$ is the central goal of MFA (for further details on MFA see [26, 27, 44]). However, it cannot be conducted based on its formal definition (2). Instead, a so-called *multifractal formalism* is used, which allows for the assessment of $D(h)$ via the scale dependence of the statistical distributions of specifically tailored multiresolution coefficients (cf., Section 2.3 below). Several different multifractal formalisms have been proposed in the literature, relying on different multiresolution quantities, cf., e.g., [4, 26, 33, 48]. In this work, we make use of *wavelet leaders* which have been proven to possess the key theoretical properties for multifractal analysis purposes and have resulted in the current state-of-the-art multifractal formalism, see, e.g., [26, 27, 54].

2.2. Wavelet leaders. The 2D gray level digitized image $\{X(\boldsymbol{k}), \boldsymbol{k} = (k_1, k_2), k_i = 1, \ldots, N\}$ is first decomposed using a discrete wavelet transform (DWT) where a square image is considered here without loss of generality. Let $H_0(k)$ and $G_0(k)$ denote the low-pass and high-pass filters defining a 1D DWT relying on a mother wavelet ψ with $N_\psi > 0$ vanishing moments. A common way to obtain a 2D orthonormal DWT is to use four 2D filters $G^{(m)}(\boldsymbol{k})$, $m = 0, \ldots, 3$ defined as the tensor products of $H_0(k)$ and $G_0(k)$. The 2D low-pass filter is by convention $G^{(0)}(\boldsymbol{k}) \triangleq H_0(k_1)H_0(k_2)$, while the high-pass filters are defined by $G^{(1)}(\boldsymbol{k}) \triangleq H_0(k_1)G_0(k_2)$, $G^{(2)}(\boldsymbol{k}) \triangleq G_0(k_1)H_0(k_2)$ and $G^{(3)}(\boldsymbol{k}) \triangleq G_0(k_1)G_0(k_2)$. Let $D_X^{(0)}(j = 0, \boldsymbol{k}) \triangleq X(\boldsymbol{k})$. The 2D wavelet coefficients $D_X^{(m)}(j, \boldsymbol{k})$, $m = 1, \ldots, 3$ and approximation coefficients $D_X^{(0)}(j, \boldsymbol{k})$ for the analysis scales $j \geq 1$ are obtained by iterative convolution of $G^{(m)}$, $m =$

[1]The local boundedness condition on X is inherent to the definition of the Hölder exponent and assumed to hold and not further discussed in what follows (cf., [27, 54, 58] for details and practical procedures for assessing and circumventing this condition).

$0, \ldots, 3$, with the approximation coefficients $D_X^{(0)}(j-1,\cdot)$, followed by decimation. In the context of MFA, it is common to use L^1 normalized wavelet coefficients, i.e., $d_X^{(m)}(j,\boldsymbol{k}) \triangleq 2^{-j} D_X^{(m)}(j,\boldsymbol{k})$, $m = 1,2,3$ [4]. For details on wavelet transforms, cf., e.g., [3,36].

Denote as $\lambda_{j,\boldsymbol{k}} = \{[k_1 2^j, (k_1+1)2^j), [k_2 2^j, (k_2+1)2^j)\}$ the dyadic cube of side length 2^j based at $\boldsymbol{k}2^j$, and as $3\lambda_{j,\boldsymbol{k}} = \bigcup_{n_1,n_2\in\{-1,0,1\}} \lambda_{j,k_1+n_1,k_2+n_2}$ the union with its eight neighbors. The wavelet leaders are defined as the largest wavelet coefficient within this neighborhood over all finer scales [26]

$$l(j,\boldsymbol{k}) \triangleq \sup_{m\in(1,2,3),\ \lambda'\subset 3\lambda_{j,\boldsymbol{k}}} |d_X^{(m)}(\lambda')|. \tag{3}$$

2.3. Wavelet leader multifractal formalism. It can be shown that the qth order empirical moments of wavelet leaders $l(j,\boldsymbol{k})$ behave as power laws in the limit of fine scales, i.e.,

$$\frac{1}{n_j}\sum_{\boldsymbol{k}} l(j,\boldsymbol{k})^q \sim 2^{j\zeta(q)}, \quad 2^j \to 0$$

where n_j stands for the number of coefficients available at scale j. The Legendre transform of the so-named *scaling exponents* $\zeta(q)$ in this relation can be shown to provide an upper-bound estimate for the multifractal spectrum $D(h)$

$$\mathcal{D}(h) \leq D(h) \inf_q \big(2 + qh - \zeta(q)\big),$$

see, e.g., [26,44]. In practice, the estimate $D(h)$ is the only quantity that is accessible in a numerically stable way and is therefore commonly conflated with the multifractal spectrum. In the seminal contribution [9], it was moreover shown that $\zeta(q)$ can be expressed as

$$\zeta(q) = \sum_{p=1}^{\infty} c_p \frac{q^p}{p\,!} \tag{4}$$

where the coefficients c_p, $p \geq 1$, termed *log-cumulants*, are tied to the pth cumulant $C_p(j)$ of the logarithm of the wavelet leaders $l(j,\boldsymbol{k})$ via the relation

$$C_p(j) = c_p^0 + c_p \ln 2^j. \tag{5}$$

Computing the Legendre transform of (4), the multifractal spectrum $D(h)$ can be developed as

$$D(h) = 2 + \frac{c_2}{2!}\left(\frac{h-c_1}{c_2}\right)^2 + \frac{-c_3}{3!}\left(\frac{h-c_1}{c_2}\right)^3 + \ldots \tag{6}$$

when $c_2 < 0$, cf., e.g., [58] (when $c_2 = 0$, $D(h)$ reduces to a Delta function at $h = c_1$). The leading order coefficients c_p therefore capture most of the essential information on the multifractal properties of X and hence are often used in applications in place of the entire function $D(h)$ [9,27,54,58].

Relation (5) suggests that the coefficients c_p can be estimated by linear regression of sample cumulants of log-leaders across scales j. Specifically, for c_2, this relation underlies the definition of the current benchmark estimator for the multifractality parameter c_2, which consists of performing a linear regression of the sample variance, denoted by $\widehat{\mathrm{Var}}$, of $\ln l(j,\boldsymbol{k})$ over a range of scales $j \in [j_1, j_2]$

$$\hat{c}_2 = \frac{1}{\ln 2} \sum_{j=j_1}^{j_2} w_j \widehat{\mathrm{Var}}[\ln l(j,\cdot)] \tag{7}$$

where w_j are suitable regression weights, cf., [54, 58].

3. Data augmented Fourier domain model for log-leaders. The estimator (7) is known to yield poor performance (large variance) even for moderate image size and is attractive mainly for its low complexity. A Bayesian model for a single image that addresses this limitation was proposed in [17] and is recalled briefly in Section 3.1. This model leads to an improved estimation performance, yet it is not well suited for the design of joint priors in the analysis of multiple images. In Section 3.2, we propose a novel data augmented statistical model in the Fourier domain for a single image that is specifically designed to be applied to multivariate images, as explained in Section 4.

3.1. Direct 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 substraction of their average (since it does not convey any information on c_2).

3.1.1. Likelihood. It was recently shown that the statistics of the log-leaders $\boldsymbol{\ell}_j$ for MMC-based processes are well approximated by a multivariate Gaussian distribution whose covariance $\mathcal{C}_j(\boldsymbol{k}, \boldsymbol{\Delta k}) \triangleq \mathrm{Cov}[l(j,\boldsymbol{k}), l(j,\boldsymbol{k}+\boldsymbol{\Delta k})]$ is given by the radial symmetric function $\varrho_j(\boldsymbol{\Delta k}; \boldsymbol{c_2})$ defined as

$$\mathcal{C}_j(\boldsymbol{k}, \boldsymbol{\Delta k}) \approx \varrho_j(\boldsymbol{\Delta k}; \boldsymbol{c_2}) \triangleq \begin{cases} \varrho_j^0(||\boldsymbol{\Delta k}||; \boldsymbol{c_2}), & ||\boldsymbol{\Delta k}|| \leq 3 \\ \varrho_j^1(||\boldsymbol{\Delta k}||; \boldsymbol{c_2}), & 3 < ||\boldsymbol{\Delta k}|| \end{cases} \tag{8}$$

and parametrized only by

$$\boldsymbol{c_2} \triangleq (c_2, c_2^0)$$

see [17] for details. The functions ϱ_j^0 and ϱ_j^1 are given by $\varrho_j^0(r; \boldsymbol{c_2}) \triangleq a_j \ln(1+r) + c_2^0 + c_2 \ln 2^j$, where $a_j \triangleq (\varrho_j^1(3; \boldsymbol{c_2}) - c_2^0 - c_2 \ln 2^j)/\ln 4$, and $\varrho_j^1(r; \boldsymbol{c_2}) \triangleq c_2 \ln(r/r_j) \mathbb{I}_{[0,r_j]}(r)$, where $r_j = \lfloor \sqrt{n_j}/4 \rfloor$ ($\lfloor \cdot \rfloor$ truncates to integer values) and $\mathbb{I}_A$ is the indicator function of the set A, respectively [17]. With these assumptions, the likelihood of $\boldsymbol{\ell}_j$ is given by

$$p(\boldsymbol{\ell}_j | \boldsymbol{c_2}) \propto |\boldsymbol{\Sigma}_{j,\boldsymbol{c_2}}|^{-\frac{1}{2}} \exp\Big(-\frac{1}{2} \boldsymbol{\ell}_j^T \boldsymbol{\Sigma}_{j,\boldsymbol{c_2}}^{-1} \boldsymbol{\ell}_j \Big) \tag{9}$$

where the covariance matrix $\boldsymbol{\Sigma}_{j,\boldsymbol{c_2}}$ is defined element-wise by $[\boldsymbol{\Sigma}_{j,\boldsymbol{c_2}}]_{u,v} = \varrho_j(||\boldsymbol{k}_u - \boldsymbol{k}_v||; \boldsymbol{c_2})$, $|\cdot|$ denotes the determinant and ^{T} the transpose operator. Following [17], independence is assumed between log-leaders at different scales j, conditionally on $\boldsymbol{c_2}$. Dependence across scales is nevertheless partially accounted for in a hierarchical fashion because the parameters $\boldsymbol{c_2}$ are modeled using the relation (5), which precisely describes the evolution across scales for

the *distributions* of $\boldsymbol{\ell}_j$ that is imposed by the multifractal model. This assumption leads to the following likelihood for the vector $\boldsymbol{\ell} \triangleq [\boldsymbol{\ell}_{j_1}^T, ..., \boldsymbol{\ell}_{j_2}^T]^T$

$$p(\boldsymbol{\ell}|\boldsymbol{c_2}) \triangleq \prod_{j=j_1}^{j_2} p(\boldsymbol{\ell}_j|\boldsymbol{c_2}). \tag{10}$$

To ensure that (10) is a valid likelihood, the covariance matrices $\boldsymbol{\Sigma}_{j,\boldsymbol{c_2}}$, $j \in \{j_1, \ldots, j_2\}$, must be positive definite. This condition implicitly defines constraints on the parameter vector $\boldsymbol{c_2}$ that can be assessed only numerically.

3.1.2. Whittle approximation. The numerical evaluation of the likelihood (9) (and hence of (10)) is problematic even for images of small size since it requires the computation of the matrix inverses $\boldsymbol{\Sigma}_{j,\boldsymbol{c_2}}^{-1}$. To bypass this difficulty, it has been proposed in [17] to approximate (9) with the asymptotic Whittle likelihood [2, 6, 21, 59]

$$p^{\dagger}(\boldsymbol{\ell}_j|\boldsymbol{c_2}) = \exp\left(-\frac{1}{2} \sum_{\mathbf{m}\in J_j} \ln \phi_j(\boldsymbol{\omega}_\mathbf{m}; \boldsymbol{c_2}) \right) \exp\left(-\frac{1}{2} \sum_{\mathbf{m}\in J_j} \frac{y_j^*(\boldsymbol{\omega}_\mathbf{m}) y_j(\boldsymbol{\omega}_\mathbf{m})}{\phi_j(\boldsymbol{\omega}_\mathbf{m}; \boldsymbol{c_2})} \right) \tag{11}$$

where $y_j(\cdot)$ is the discrete Fourier transform (DFT) of $\ell(j, \cdot)$, i.e., $y_j^*(\boldsymbol{\omega}_\mathbf{m}) y_j(\boldsymbol{\omega}_\mathbf{m})$ corresponds to the periodogram of $\{\ell(j, \boldsymbol{k})\}$ at frequency $\boldsymbol{\omega}_\mathbf{m} = 2\pi\mathbf{m}/\sqrt{n_j}$, and * denotes the conjugation operator. The summation is taken over the grid of integers $J_j \triangleq [\![\lfloor(-\sqrt{n_j} - 1)/2\rfloor, \ldots, \sqrt{n_j} - \lfloor\sqrt{n_j}/2\rfloor]\!]\backslash\{0\}^2$, where the zero frequency is removed, making the model mean-independent. Moreover, $\phi_j(\boldsymbol{\omega}_\mathbf{m}; \boldsymbol{c_2})$ stands for the discretized parametric spectral density associated with the covariance model (8), i.e.,

$$\phi_j(\boldsymbol{\omega}_\mathbf{m}; \boldsymbol{c_2}) = \bar{\mathcal{FT}}\big[\varrho_j(\cdot; \boldsymbol{c_2})\big](\boldsymbol{\omega}_\mathbf{m}),$$

where $\bar{\mathcal{FT}}\big[\,\cdot\,\big]$ stands for a discretization of the continuous time Fourier transform that takes into account aliasing. It can be shown that $\phi_j(\boldsymbol{\omega}_\mathbf{m}; \boldsymbol{c_2})$ is of the form

$$\phi_j(\boldsymbol{\omega}_\mathbf{m}; \boldsymbol{c_2}) = c_2\, h_j^{(1)}(\boldsymbol{\omega}_\mathbf{m}) + c_2^0\, h_j^{(2)}(\boldsymbol{\omega}_\mathbf{m}) \tag{12}$$

where the vectors $h_j^{(1)}$ and $h_j^{(2)}$ do not depend on the parameter vector $\boldsymbol{c_2}$. They can hence be pre-calculated using DFT and stored.

3.2. Data augmented statistical model in the Fourier domain. The study of estimators for the model (10-12) revealed as a major practical limitation the difficulty of designing conjugate priors for the parameter vector $\boldsymbol{c_2}$ since it is encoded implicitly in $\boldsymbol{\Sigma}_{j,\boldsymbol{c_2}}^{-1}$, and its conditional distribution is thus not standard. Sampling the posterior distribution with a Markov chain Monte Carlo (MCMC) method was then achieved using accept/reject procedures, such as Metropolis-Hastings (MH) moves. In view of a model expansion to multiple images (and, hence, high-dimensional parameter vectors), such a procedure becomes quickly computationally impracticable. Instead, we propose an alternative model that leads to a more efficient algorithm (cf., [15] for a preliminary study) by interpreting (11) as a statistical model for the Fourier coefficients $\boldsymbol{y}$. Then a suitable reparametrization is introduced that enables us to formulate a data augmented model in which conjugate priors for $\boldsymbol{c_2}$ are available.

3.2.1. Fourier domain model. To see how (11) leads to a statistical model for $\boldsymbol{y}$, we first develop the expression using the following equivalences. First, because the log-leaders $\{\ell(j,\boldsymbol{k})\}$ are real-valued, their Fourier transform y_j has the central symmetry property $y_j(\boldsymbol{\omega_m}) = y_j^*(-\boldsymbol{\omega_m})$. Moreover, the parametric spectral density $\phi_j(\boldsymbol{\omega};\boldsymbol{c_2})$ has the same property. Due to this symmetry, only one half of the frequency plane J_j needs to be considered in the sum in (11), which we denote by $\underline{J}_j$. It can for instance contain the positive frequency half-plane, i.e., $\underline{J}_j \triangleq J_j \cap \big\{\{(i_1,i_2)\in\mathbb{Z}^2 | i_1 \geq 0\}\backslash\{(i_1,i_2)|i_1=0, i_2<0\}\big\}$. Second, it is straightforward to rewrite the first term in (11) as the inverse of the determinant of a diagonal matrix with entries $\phi_j(\boldsymbol{\omega_m};\boldsymbol{c_2})$. Third, we write the sum in the second term in (11) as a matrix-vector product by introducing the operator $\mathcal{F}_{\underline{J}_j}(\cdot)$ computes and vectorizes the DFT coefficients contained in the half-plane $\underline{J}_j$. With these ingredients, (11) can be developed into

$$p^\dagger(\boldsymbol{\ell}_j|\boldsymbol{c_2}) = \det(\boldsymbol{\Gamma}_{j,\boldsymbol{c_2}})^{-1} \exp\left(-\boldsymbol{y}_j^H \boldsymbol{\Gamma}_{j,\boldsymbol{c_2}}^{-1} \boldsymbol{y}_j\right),\ \boldsymbol{y}_j \triangleq \mathcal{F}_{\underline{J}_j}(\boldsymbol{\ell}_j) \tag{13}$$

where ^{H} is the conjugate transpose operator and $\boldsymbol{\Gamma}_{j,\boldsymbol{c_2}}$ is the diagonal matrix defined by

$$\begin{aligned} \boldsymbol{\Gamma}_{j,\boldsymbol{c_2}} &\triangleq c_2 \boldsymbol{F}_j + c_2^0 \boldsymbol{G}_j \\ \boldsymbol{F}_j &\triangleq \mathrm{diag}\,(\mathbf{f}_j) \ \text{ with } \ \mathbf{f}_j \triangleq \big(h_j^{(1)}(\boldsymbol{\omega_m})\big)_{\mathbf{m}\in\underline{J}_j} \\ \boldsymbol{G}_j &\triangleq \mathrm{diag}\,(\mathbf{g}_j) \ \text{ with } \ \mathbf{g}_j \triangleq \big(h_j^{(2)}(\boldsymbol{\omega_m})\big)_{\mathbf{m}\in\underline{J}_j}. \end{aligned} \tag{14}$$

By inspection of (13), modeling $\boldsymbol{\ell}_j$ using $p^\dagger(\boldsymbol{\ell}_j|\boldsymbol{c_2})$ is equivalent to modeling $\boldsymbol{y}_j = \mathcal{F}_{\underline{J}_j}(\boldsymbol{\ell}_j)$ by a random vector with a non-degenerate centered circular-symmetric complex Gaussian distribution $\mathcal{CN}(\mathbf{0},\boldsymbol{\Gamma}_{j,\boldsymbol{c_2}})$ [22, 43], provided the matrix $\boldsymbol{\Gamma}_{j,\boldsymbol{c_2}}$ is positive definite (PD). We therefore propose to consider the Fourier coefficients $\boldsymbol{y}_j = \mathcal{F}_{\underline{J}_j}(\boldsymbol{\ell}_j)$, $j = j_1,\ldots,j_2$, as the observed data with likelihood

$$p(\boldsymbol{y}_j|\boldsymbol{c_2}) = |\boldsymbol{\Gamma}_{j,\boldsymbol{c_2}}|^{-1} \exp\left(-\boldsymbol{y}_j^H \boldsymbol{\Gamma}_{j,\boldsymbol{c_2}}^{-1} \boldsymbol{y}_j\right) \tag{15}$$

instead of the log-leaders $\boldsymbol{\ell}_j$ with likelihood (9).

Assuming independence between scales j as in the direct model (10), the likelihood for the vector $\boldsymbol{y} \triangleq [\boldsymbol{y}_{j_1}^T,...,\boldsymbol{y}_{j_2}^T]^T$ replacing (10) is given by

$$p(\boldsymbol{y}|\boldsymbol{c_2}) \triangleq \prod_{j=j_1}^{j_2} p(\boldsymbol{y}_j|\boldsymbol{c_2}) \propto |\boldsymbol{\Gamma}_{\boldsymbol{c_2}}|^{-1} \exp\left(-\boldsymbol{y}^H \boldsymbol{\Gamma}_{\boldsymbol{c_2}}^{-1} \boldsymbol{y}\right) \tag{16}$$

where $\boldsymbol{\Gamma}_{\boldsymbol{c_2}}$ is the $N_Y \times N_Y$ diagonal covariance matrix, $N_Y \triangleq \mathrm{card}(\boldsymbol{y})$, defined as

$$\begin{aligned} \boldsymbol{\Gamma}_{\boldsymbol{c_2}} &\triangleq c_2 \boldsymbol{F} + c_2^0 \boldsymbol{G} \\ \boldsymbol{F} &\triangleq \mathrm{diag}\,(\mathbf{f}) \ \text{ with } \ \mathbf{f} \triangleq [\mathbf{f}_{j_1}^T,...,\mathbf{f}_{j_2}^T]^T \\ \boldsymbol{G} &\triangleq \mathrm{diag}\,(\mathbf{g}) \ \text{ with } \ \mathbf{g} \triangleq [\mathbf{g}_{j_1}^T,...,\mathbf{g}_{j_2}^T]^T. \end{aligned} \tag{17}$$

In contrast to (10), the admissible set for $\boldsymbol{c_2} = (c_2, c_2^0)$ that ensures that the matrix $\boldsymbol{\Gamma}_{\boldsymbol{c_2}}$ is PD (and hence (16) is a valid likelihood) can now be expressed explicitly as

$$\mathcal{A} = \{\boldsymbol{c_2} \in \mathbb{R}_{<0} \times \mathbb{R}_{<0} | c_2 \mathbf{f}(k) + c_2^0 \mathbf{g}(k) > 0, k = 1,\ldots,N_Y\}. \tag{18}$$

3.2.2. Reparametrization. The likelihood (16) and the constraints (18) are not separable in c_2 and c_2^0, which prevents the design of independent conjugate priors for c_2 and c_2^0. To circumvent this difficulty, we propose a suitable reparametrization that enables an augmented model, associated with a separable extended likelihood leading to $\mathcal{IG}$ conjugate priors for the parameters of interest. Note that, to this end, the covariance matrix $\boldsymbol{\Gamma}_{\boldsymbol{c_2}}$ must be decomposed as the sum of two PD diagonal matrices, which is not the case in (17) since $c_2\boldsymbol{F}$ is not PD for $c_2 < 0$ (because there always exists an integer k such that $\mathbf{f}(k) > 0$). Note that $c_2^0\boldsymbol{G}$ in (17) is in practice always PD, for $c_2^0 > 0$, since it can be checked that $g(k) > 0, \forall k$ for any reasonable image size. We thus propose to use a reparametrization defined by the mapping

$$\psi : \boldsymbol{c_2} \mapsto \tilde{\mathbf{c}}_{\boldsymbol{2}} = (\tilde{c}_{21}, \tilde{c}_{22}) \triangleq (-c_2, c_2^0/\gamma + c_2), \tag{19}$$

where $\gamma = \sup_k \mathbf{f}(k)/\mathbf{g}(k)$. It is easy to show that ψ is a one-to-one transformation from $\mathcal{A}$ to $\mathbb{R}^2_{>0}$ and hence maps the constraints (18) into independent positivity constraints, $\tilde{c}_{2i} \in \mathbb{R}_{>0},\ i = 1, 2$. Moreover, (16) expressed with $\tilde{\mathbf{c}}_{\boldsymbol{2}}$ reads

$$p(\boldsymbol{y}|\tilde{\mathbf{c}}_{\boldsymbol{2}}) \propto |\boldsymbol{\Gamma}_{\tilde{\mathbf{c}}_{\boldsymbol{2}}}|^{-1} \exp\left(-\boldsymbol{y}^H \boldsymbol{\Gamma}_{\tilde{\mathbf{c}}_{\boldsymbol{2}}}^{-1} \boldsymbol{y}\right) \tag{20}$$

$$\boldsymbol{\Gamma}_{\tilde{\mathbf{c}}_{\boldsymbol{2}}} = \tilde{c}_{21}\tilde{\boldsymbol{F}} + \tilde{c}_{22}\tilde{\boldsymbol{G}}, \quad \tilde{\boldsymbol{F}} = -\boldsymbol{F} + \boldsymbol{G}\gamma, \quad \tilde{\boldsymbol{G}} = \boldsymbol{G}\gamma \tag{21}$$

where, by construction, the two diagonal matrices $\tilde{c}_{21}\tilde{\boldsymbol{F}}$ and $\tilde{c}_{22}\tilde{\boldsymbol{G}}$ are now PD for $\tilde{\mathbf{c}}_{\boldsymbol{2}} \in \mathbb{R}^2_{>0}$.

3.2.3. Data augmentation. We can now introduce an $N_Y \times 1$ vector of latent variables $\boldsymbol{\mu}$ that enables the likelihood (20) to be augmented using the following model

$$\boldsymbol{y}|\boldsymbol{\mu}, \tilde{c}_{22} \sim \mathcal{CN}(\boldsymbol{\mu}, \tilde{c}_{22}\tilde{\boldsymbol{G}}), \quad \boldsymbol{\mu}|\tilde{c}_{21} \sim \mathcal{CN}(0, \tilde{c}_{21}\tilde{\boldsymbol{F}}) \tag{22}$$

which is associated with the extended likelihood [20, 52]

$$p(\boldsymbol{y}, \boldsymbol{\mu}|\tilde{\mathbf{c}}_{\boldsymbol{2}}) \propto \tilde{c}_{22}^{-N_Y} \exp\Big(-\frac{1}{\tilde{c}_{22}}(\boldsymbol{y}-\boldsymbol{\mu})^H \tilde{\boldsymbol{G}}^{-1}(\boldsymbol{y}-\boldsymbol{\mu})\Big) \times \tilde{c}_{21}^{-N_Y} \exp\Big(-\frac{1}{\tilde{c}_{21}}\boldsymbol{\mu}^H \tilde{\boldsymbol{F}}^{-1}\boldsymbol{\mu}\Big). \tag{23}$$

It can easily be verified that the likelihood (20) is recovered by marginalization of (23) with respect to the latent variables $\boldsymbol{\mu}$. Moreover, the extended likelihood (23) and the associated constraints are both separable in $(\tilde{c}_{21}, \tilde{c}_{22})$.

3.2.4. Prior and posterior distribution. When $\mathcal{IG}(\alpha_{0,i}, \beta_{0,i})$ distributions $p(\tilde{c}_{2i})$ are used as priors for $\tilde{c}_{2i} \in \mathbb{R}_{>0},\ i = 1, 2$, in (23), simple calculations show that the posterior distribution

$$p(\tilde{\mathbf{c}}_{\boldsymbol{2}}, \boldsymbol{\mu}|\boldsymbol{y}) \propto p(\boldsymbol{y}|\tilde{c}_{22}, \boldsymbol{\mu}) p(\boldsymbol{\mu}|\tilde{c}_{21}) p(\tilde{c}_{22}) p(\tilde{c}_{21}) \tag{24}$$

presents standard conditional distributions (detailed in (32a-32b) below) for the transformed multifractal parameters $\tilde{c}_{2i}$. A consequence of this property is that the computational cost of the sampler, which will be proposed to generate vectors distributed according to (24), will be significantly reduced.

4. Bayesian model for multivariate analysis. We are now ready to specify a Bayesian model addressing the estimation of c_2 for multivariate images by using the above statistical model and incorporating suitable joint priors.

4.1. Problem formulation.

4.1.1. Multivariate image scenario. We consider a (temporal, or spectral) sequence of images

$$\{\mathbf{X}_t\},\ \text{where } t \in \Omega^{(2)} \triangleq \{1,\dots,N_t\}$$

and divide each single image $\mathbf{X}_t$ into (non-overlapping) patches as illustrated in Fig. 1 (left)

$$\mathbf{X}_{(\boldsymbol{x},t)},\ \boldsymbol{x}=(x_1,x_2) \in \Omega^{(1)} \triangleq \{1,\dots,N_{x_1}\}\times\{1,\dots,N_{x_2}\}.$$

For notational convenience, we write $\boldsymbol{s}=(\boldsymbol{x},t) \in \Omega^{(1)}\times\Omega^{(2)}$. We investigate the situation where the (sequential or spatial) evolution of the multifractal properties between the elements $\mathbf{X}_{\boldsymbol{s}}$ is assumed to be *smooth*. The decomposition thus enables the assessment of the temporal / spectral evolution of the *spatially localized* (patch-wise) multifractal properties of the image. The distinction between the spatial and the temporal/spectral component in the decomposition is meaningful because $\boldsymbol{x}$ and t have different physical roles (since we deal with a sequence of images and not an isotropic 3D data cube). This formulation is also valid for sequences of images with spatial organization, rather than patches. Models for other multivariate image scenarios can be obtained in a straightforward manner from the developments detailed in the following sections. As an example, a model for non-overlapping patches of a single image was sketched in our preliminary work [16].

4.1.2. Likelihood. For each element $\mathbf{X}_{\boldsymbol{s}}$, we consider the extended model (23) and we gather the Fourier coefficients of the log-leaders $\boldsymbol{y}_{\boldsymbol{s}}$ and the latent variables $\boldsymbol{\mu}_{\boldsymbol{s}}$ for all elements $\boldsymbol{s}$ in matrices $\mathbf{Y}$ and $\mathbf{M}$, respectively. Moreover, let $\tilde{\mathbf{C}}_{\mathbf{2}}=(\tilde{\mathbf{C}}_{\mathbf{21}},\tilde{\mathbf{C}}_{\mathbf{22}})$ where $\tilde{\mathbf{C}}_{\mathbf{2}i}$ gathers the parameters $\tilde{c}_{2i,\boldsymbol{s}}$ for all $\boldsymbol{s}$. By assuming *a priori* independence between the elements $\mathbf{X}_{\boldsymbol{s}}$, the joint likelihood can be generically written as

(25)
$$p(\mathbf{Y},\mathbf{M}|\tilde{\mathbf{C}}_{\mathbf{2}}) \propto \prod_{\boldsymbol{s}} \tilde{c}_{21,\boldsymbol{s}}^{-N_Y}\exp\Big(-\frac{1}{\tilde{c}_{21,\boldsymbol{s}}}\boldsymbol{\mu}_{\boldsymbol{s}}^H\tilde{\boldsymbol{F}}^{-1}\boldsymbol{\mu}_{\boldsymbol{s}}\Big)\tilde{c}_{22,\boldsymbol{s}}^{-N_Y}\exp\Big(-\frac{1}{\tilde{c}_{22,\boldsymbol{s}}}(\boldsymbol{y}_{\boldsymbol{s}}-\boldsymbol{\mu}_{\boldsymbol{s}})^H\tilde{\boldsymbol{G}}^{-1}(\boldsymbol{y}_{\boldsymbol{s}}-\boldsymbol{\mu}_{\boldsymbol{s}})\Big).$$

4.2. Gamma Markov Random Field prior.

Inverse-gamma distributions $\mathcal{IG}(\alpha_{i,\boldsymbol{s}},\beta_{i,\boldsymbol{s}})$ are conjugate priors for the parameters $\tilde{c}_{2i,\boldsymbol{s}}$. A careful design of $(\alpha_{i,\boldsymbol{s}},\beta_{i,\boldsymbol{s}})$, rather than setting them a priori to constant values, can enforce the parameters of interest $\tilde{\mathbf{C}}_{\mathbf{2}i}$ to vary slowly in some privileged directions. To do so, we propose here to specify $(\alpha_{i,\boldsymbol{s}},\beta_{i,\boldsymbol{s}})$ such that the resulting prior for $\tilde{\mathbf{C}}_{\mathbf{2}i}$ is a hidden gamma Markov random field (GaMRF) [19]. The strategy for this prior relies on the use of a set of positive auxiliary variables $\mathbf{Z}_i$ to induce positive dependence between neighboring elements of $\tilde{\mathbf{C}}_{\mathbf{2}i}$ [19]. In what follows, we handle the spatial and temporal/spectral components with separate variables $\mathbf{Z}_i^{(1)}$ and $\mathbf{Z}_i^{(2)}$, respectively, and denote $\mathbf{Z}_i=(\mathbf{Z}_i^{(1)},\mathbf{Z}_i^{(2)})$. The distribution of the joint prior $(\tilde{\mathbf{C}}_{\mathbf{2}i},\mathbf{Z}_i)$ is associated with conditionals of the form

(26a)
$$\tilde{c}_{2i,\boldsymbol{s}}\,|\mathbf{Z}_i,\mathbf{a}_i \sim \mathcal{IG}\big(\alpha_{i,\boldsymbol{s}}(\mathbf{a}_i),\beta_{i,\boldsymbol{s}}(\mathbf{Z}_i,\mathbf{a}_i)\big)$$

(26b)
$$z_{i,\boldsymbol{s}}^{(m)}|\tilde{\mathbf{C}}_{\mathbf{2}i},\mathbf{a}_i \sim \mathcal{G}(\bar{\alpha}_{i,\boldsymbol{s}}^{(m)}(\mathbf{a}_i),\bar{\beta}_{i,\boldsymbol{s}}^{(m)}(\tilde{\mathbf{C}}_{\mathbf{2}i},\mathbf{a}_i))$$

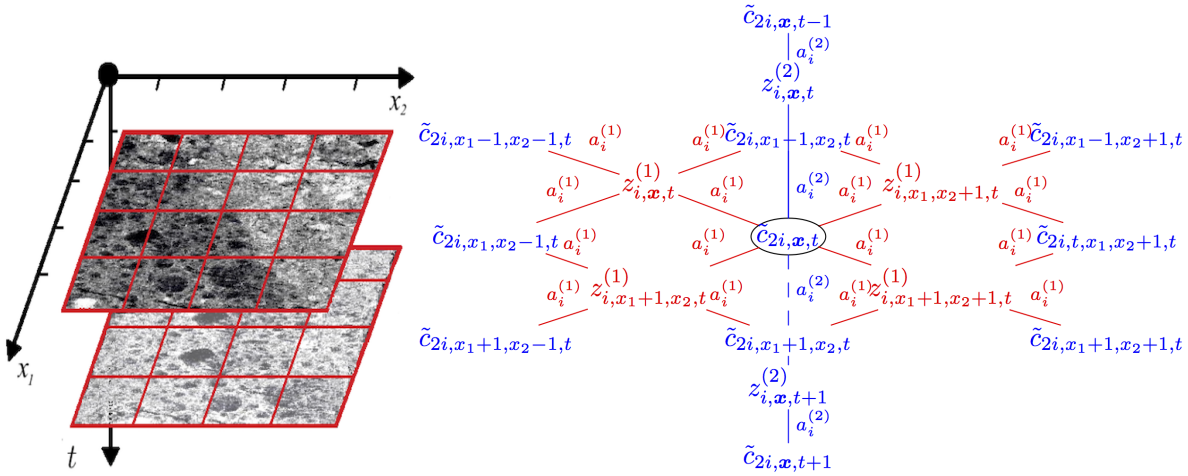


Figure 1: Illustration of multivariate image scenario (left). Spatial (red) and spectral (blue) components of the proposed bipartite conditional independence graphs between $\tilde{\mathbf{C}}_{\mathbf{2}i}$ and $\mathbf{Z}_i$ (right).

where $\mathcal{G}(\cdot,\cdot)$ stands for the gamma distribution and $\mathbf{a}_i \triangleq [a_i^{(1)}, a_i^{(2)}]$ is a regularization vector where $a_i^{(1)}$ and $a_i^{(2)}$ adjust the amount of spatial and temporal/spectral smoothness, respectively. The parameters $\alpha_{i,\boldsymbol{s}}, \beta_{i,\boldsymbol{s}}, \bar{\alpha}_{i,\boldsymbol{s}}^{(m)}$ and $\bar{\beta}_{i,\boldsymbol{s}}^{(m)}$ of the $\mathcal{IG}$ and $\mathcal{G}$ distributions in (26) are determined by the definition of a bipartite conditional independence graph between $\tilde{\mathbf{C}}_{\mathbf{2}i}$ and $\mathbf{Z}_i^{(m)}$. The design of the vertices and edges of the graph for the decomposition into a spatio-temporal/spectral collection $\{\mathbf{X}_{\boldsymbol{x},t}\}_{\boldsymbol{x},t\in\Omega^{(1)}\times\Omega^{(2)}}$ proposed here is sketched in Fig. 1: Correlation between the parameters of neighboring patches is induced by two sets of auxiliary variables $\mathbf{Z}_i^{(1)} = \{z_{i,\boldsymbol{x},t}^{(1)}\}_{\boldsymbol{x},t\in\Omega^{(1)}\times\Omega^{(2)}}$ and $\mathbf{Z}_i^{(2)} = \{z_{i,\boldsymbol{x},t}^{(2)}\}_{\boldsymbol{x},t\in\Omega^{(1)}\times\Omega^{(2)}}$, $z_{i,\boldsymbol{x},t}^{(1)} \in \mathbb{R}_{>0}$ and $z_{i,\boldsymbol{x},t}^{(2)} \in \mathbb{R}_{>0}$, using a graph that connects each $\tilde{c}_{2i,\boldsymbol{x},t}$ to its four natural spatial neighbors in $\mathbf{Z}_i^{(1)}$ and to its two natural temporal neighbors in $\mathbf{Z}_i^{(2)}$, and vice-versa. Notably, by doing so, $\tilde{c}_{2i,\boldsymbol{x},t}$ and $\tilde{c}_{2i,x_1,x_2-1,t}$ are connected via $z_{i,\boldsymbol{x},t}^{(1)}$ and $z_{i,x_1+1,x_2,t}^{(1)}$ over edges with weight $a_i^{(1)}$, and $\tilde{c}_{2i,\boldsymbol{x},t}$ are connected to $\tilde{c}_{2i,\boldsymbol{x},t-1}$ via $z_{i,\boldsymbol{x},t}^{(2)}$ over edges with weight $a_i^{(2)}$. The density of the resulting GaMRF prior for $(\tilde{\mathbf{C}}_{\mathbf{2}i}, \mathbf{Z}_i)$ is then given by [19]

$$
\begin{aligned}
p(\tilde{\mathbf{C}}_{\mathbf{2}i}, \mathbf{Z}_i|\mathbf{a}_i) &= \frac{1}{K(\mathbf{a}_i)} \prod_{\boldsymbol{x}\in\Omega^{(1)}} \prod_{t\in\Omega^{(2)}} \Phi_{\mathcal{I}}\big(\tilde{c}_{2i,\boldsymbol{x},t}; 4a_i^{(1)} + 2a_i^{(2)}\big) \\
&\times \prod_{\boldsymbol{x}\in\Omega^{(1)}} \prod_{t\in\Omega^{(2)}} \Phi_{\mathcal{G}}\big(z_{i,\boldsymbol{x},t}^{(1)}; 4a_i^{(1)}\big) \times \prod_{\boldsymbol{x}\in\Omega^{(1)}} \prod_{t\in\Omega^{(2)}} \Phi_{\mathcal{G}}\big(z_{i,\boldsymbol{x},t}^{(2)}; 2a_i^{(2)}\big) \\
&\times \prod_{\boldsymbol{x}\in\Omega^{(1)}} \prod_{t\in\Omega^{(2)}} \Phi\Big(a_i^{(2)} \sum_{t'\in\mathcal{V}_{\tilde{\mathbf{c}}_{\mathbf{2}}}^{(1)}(t)} z_{i,\boldsymbol{x},t'}^{(2)} + a_i^{(1)} \sum_{\boldsymbol{x}'\in\mathcal{V}_{\tilde{\mathbf{c}}_{\mathbf{2}}}^{(2)}(\boldsymbol{x})} z_{i,\boldsymbol{x}',t}^{(1)}, \tilde{c}_{2i,\boldsymbol{x},t}\Big)
\end{aligned} \tag{27}
$$

where we have introduced $\Phi_{\mathcal{I}}(\xi;\alpha) \triangleq \exp\left(-(\alpha+1)\log\xi\right)$, $\Phi_{\mathcal{G}}(\xi;\alpha) \triangleq \exp\left((\alpha-1)\log\xi\right)$ and $\Phi(\xi,\delta) \triangleq \exp(-\xi/\delta)$ for convenience of notation and where $K(\mathbf{a}_i)$ is a normalizing constant. The sums in (27) are taken over the neighborhoods

$$\mathcal{V}^{(1)}_{\tilde{\mathbf{c}}_{\mathbf{2}}}(\boldsymbol{x}) = \{(x_1,x_2),(x_1+1,x_2),(x_1,x_2+1),(x_1+1,x_2+1)\},$$
$$\mathcal{V}^{(1)}_{z}(\boldsymbol{x}) = \{(x_1-1,x_2-1),(x_1,x_2-1),(x_1-1,x_2),(x_1,x_2)\},$$
$$\mathcal{V}^{(2)}_{\tilde{\mathbf{c}}_{\mathbf{2}}}(t) = \{t,t+1\}, \qquad \mathcal{V}^{(2)}_{z}(t) = \{t-1,t\},$$

which are the indices of variables that define the neighborhoods of $\tilde{c}_{2i,\boldsymbol{x},t}$ and $z^{(1)}_{i,\boldsymbol{x},t}$ and $z^{(2)}_{i,\boldsymbol{x},t}$, respectively. The parameters of the associated conditionals (26a) and (26b) are

$$\alpha_{i,\boldsymbol{x},t} = 4a_i^{(1)} + 2a_i^{(2)}, \quad \beta_{i,\boldsymbol{x},t} = a_i^{(1)} \sum_{\boldsymbol{x}'\in\mathcal{V}^{(1)}_{\tilde{\mathbf{c}}_{\mathbf{2}}}(\boldsymbol{x})} z^{(1)}_{i,\boldsymbol{x}',t} + a_i^{(2)} \sum_{t'\in\mathcal{V}^{(2)}_{\tilde{\mathbf{c}}_{\mathbf{2}}}(t)} z^{(2)}_{i,\boldsymbol{x},t'} \tag{28a}$$

$$\bar{\alpha}^{(1)}_{i,\boldsymbol{x},t} = 4a_i^{(1)}, \quad \bar{\beta}^{(1)}_{i,\boldsymbol{x},t} = \Big(a_i^{(1)} \textstyle\sum_{\boldsymbol{x}'\in\mathcal{V}^{(1)}_{z}(\boldsymbol{x})} \tilde{c}^{-1}_{2i,\boldsymbol{x}',t}\Big)^{-1} \tag{28b}$$

$$\bar{\alpha}^{(2)}_{i,\boldsymbol{x},t} = 2a_i^{(2)}, \quad \bar{\beta}^{(2)}_{i,\boldsymbol{x},t} = \Big(a_i^{(2)} \textstyle\sum_{t'\in\mathcal{V}^{(2)}_{z}(t)} \tilde{c}^{-1}_{2i,\boldsymbol{x},t'}\Big)^{-1}. \tag{28c}$$

4.3. Joint posterior distribution.

Using Bayes' theorem and assuming prior independence between $(\tilde{\mathbf{C}}_{\mathbf{21}}, \mathbf{Z}_1)$ and $(\tilde{\mathbf{C}}_{\mathbf{22}}, \mathbf{M}, \mathbf{Z}_2)$, the joint posterior distribution associated with the proposed Bayesian model is

$$p(\tilde{\mathbf{C}}_{\mathbf{2}}, \mathbf{Z}, \mathbf{M}|\mathbf{Y}, \{\mathbf{a}_i\}) \propto p(\mathbf{Y}|\tilde{\mathbf{C}}_{\mathbf{22}}, \mathbf{M})\, p(\mathbf{M}|\tilde{\mathbf{C}}_{\mathbf{21}}) \times p(\tilde{\mathbf{C}}_{\mathbf{21}}, \mathbf{Z}_1|\mathbf{a}_1)\, p(\tilde{\mathbf{C}}_{\mathbf{22}}, \mathbf{Z}_2|\mathbf{a}_2) \tag{29}$$

with $\mathbf{Z} = (\mathbf{Z}_1, \mathbf{Z}_2)$ and where the regularization hyperparameters $\mathbf{a}_i$ are fixed a priori (the estimation of $\mathbf{a}_i$, given that the normalizing constant $K(\mathbf{a}_i)$ is intractable, is not considered here for the sake of clarity and focus, cf., e.g., [7, 25, 40, 45] for methods addressing such situations).

5. Bayesian Inference.

5.1. Bayesian estimators.

The knowledge of the unknown multifractal parameters $\tilde{\mathbf{C}}_{\mathbf{2}i}$ given the observed data and the prior information assigned to the different model parameters is summarized in the posterior distribution (29). We consider here the marginal posterior mean estimator for $\tilde{\mathbf{C}}_{\mathbf{2}i}$, denoted by the superscript MMSE for minimum mean square error estimator and defined as

$$\tilde{\mathbf{C}}^{\mathrm{MMSE}}_{\mathbf{2}i} \triangleq \mathbb{E}[\tilde{\mathbf{C}}_{\mathbf{2}i}|\boldsymbol{Y}, \mathbf{a}_i] \tag{30}$$

where the expectation is taken with respect to the marginal posterior density $p(\tilde{\mathbf{C}}_{\mathbf{2}i}|\boldsymbol{Y}, \mathbf{a}_i)$. The direct computation of (30) is intractable as it requires integrating the posterior (29) over all other unknown variables. However (30) can be approximated with an arbitrary precision by resorting to an MCMC algorithm [46]. Here we consider a Gibbs sampler drawing samples $(\tilde{\mathbf{C}}^{(q)}_{\mathbf{2}}, \boldsymbol{M}^{(q)}, \mathbf{Z}^{q})_{q=1}^{N_{mc}}$ that are asymptotically distributed according to the targeted joint posterior (29). These samples are used in turn to approximate the marginal posterior mean (30) by $\tilde{\mathbf{C}}^{\mathrm{MMSE}}_{\mathbf{2}i} \approx (N_{mc} - N_{bi})^{-1} \sum_{q=N_{bi}+1}^{N_{mc}} \tilde{\mathbf{C}}^{(q)}_{\mathbf{2}i}$ [46], where N_{bi} is the length of the burn-in period.

5.2. Gibbs sampler. The strategy of the Gibbs sampler consists of successively generating samples from the conditional distributions associated with the posterior. It is easy to show that the conditional distributions associated with the posterior (29) are given by

$$\boldsymbol{\mu_s}|\mathbf{Y}, \tilde{\mathbf{C}}_{\mathbf{2}} \sim \mathcal{CN}\Big(\tilde{c}_{21,\boldsymbol{s}}\tilde{\boldsymbol{F}}\boldsymbol{\Gamma}^{-1}_{\tilde{\mathbf{c}}_{\mathbf{2},\boldsymbol{s}}}\boldsymbol{y_s}, \Big((\tilde{c}_{21,\boldsymbol{s}}\tilde{\boldsymbol{F}})^{-1} + (\tilde{c}_{22,\boldsymbol{s}}\tilde{\boldsymbol{G}})^{-1}\Big)^{-1}\Big) \tag{31a}$$

$$\tilde{c}_{21,\boldsymbol{s}}|\mathbf{M}, \mathbf{Z}_1 \sim \mathcal{IG}\left(N_Y + \alpha_{1,\boldsymbol{s}}, ||\boldsymbol{\mu_s}||_{\tilde{\boldsymbol{F}}^{-1}} + \beta_{1,\boldsymbol{s}}\right) \tag{31b}$$

$$\tilde{c}_{22,\boldsymbol{s}}|\mathbf{Y}, \mathbf{M}, \mathbf{Z}_2 \sim \mathcal{IG}\left(N_Y + \alpha_{2,\boldsymbol{s}}, ||\boldsymbol{y_s} - \boldsymbol{\mu_s}||_{\tilde{\boldsymbol{G}}^{-1}} + \beta_{2,\boldsymbol{s}}\right) \tag{31c}$$

$$z^{(m)}_{i,\boldsymbol{s}}|\tilde{\mathbf{C}}_{\mathbf{2}i} \sim \mathcal{G}(\bar{\alpha}^{(m)}_{i,\boldsymbol{s}}, \bar{\beta}^{(m)}_{i,\boldsymbol{s}}) \tag{31d}$$

where $||\mathbf{x}||_{\mathbf{M}} \triangleq \mathbf{x}^H\mathbf{M}\mathbf{x}$. We recall that $\boldsymbol{\Gamma}_{\tilde{\mathbf{c}}_{\mathbf{2},\boldsymbol{s}}} \triangleq \tilde{c}_{21,\boldsymbol{s}}\tilde{\boldsymbol{F}} + \tilde{c}_{22,\boldsymbol{s}}\tilde{\boldsymbol{G}}$ and that $\alpha_{i,\boldsymbol{s}}, \beta_{i,\boldsymbol{s}}, \bar{\alpha}^{(m)}_{i,\boldsymbol{s}}$ and $\bar{\beta}^{(m)}_{i,\boldsymbol{s}}$ are defined in (28). Note that all conditional distributions are standard and thus can be sampled efficiently, without MH accept-reject steps. This property is a direct consequence of the proposed extended model (23) (involving $\mathcal{IG}$ conjugate priors) and enables the estimation of large unknown parameter vectors $\tilde{\mathbf{C}}_{\mathbf{2}i}$.

Finally, note that when each $\tilde{c}_{2i,\boldsymbol{s}}$ is modeled independently with the univariate model (24) (i.e., the parameters $\tilde{c}_{2i,\boldsymbol{s}}$ have inverse-gamma priors $\mathcal{IG}(\alpha_{0,i}, \beta_{0,i})$ with a priori fixed parameters $\alpha_{0,i}, \beta_{0,i}$ instead of the GaMRF priors of Section 4.2), the conditional distributions are given by (31a) and

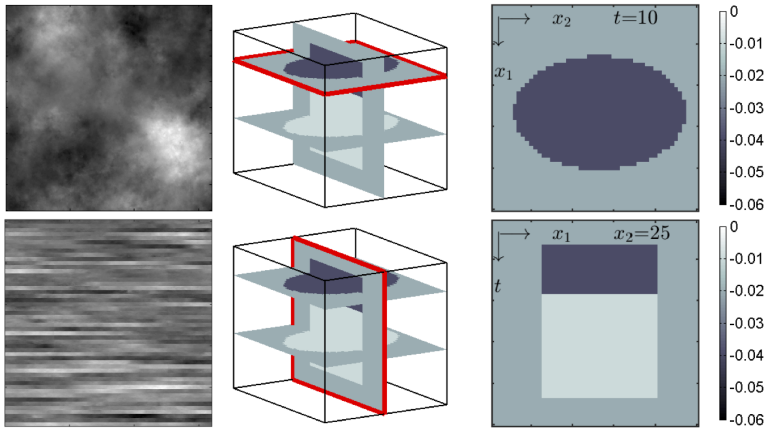
$$\tilde{c}_{21,\boldsymbol{s}}|\mathbf{M} \sim \mathcal{IG}\left(N_Y + \alpha_{0,1}, ||\boldsymbol{\mu_s}||_{\tilde{\boldsymbol{F}}^{-1}} + \beta_{0,1}\right) \tag{32a}$$

$$\tilde{c}_{22,\boldsymbol{s}}|\mathbf{Y}, \mathbf{M} \sim \mathcal{IG}\left(N_Y + \alpha_{0,2}, ||\boldsymbol{y_s} - \boldsymbol{\mu_s}||_{\tilde{\boldsymbol{G}}^{-1}} + \beta_{0,2}\right) \tag{32b}$$

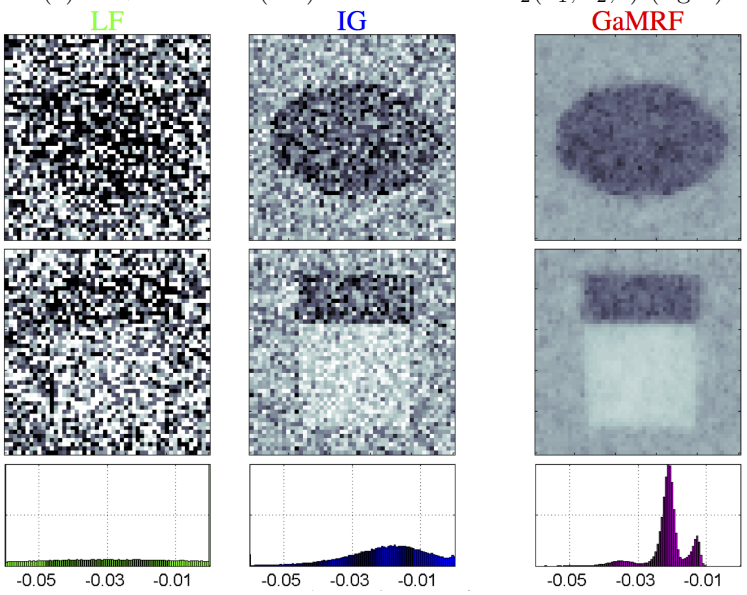
((31d) is discarded in this model). Moreover, it is easy to see that the model can be straightforwardly adapted to situations in which data components (i.e., elements of $\mathbf{Y}$) are corrupted or missing, in which case estimates for the affected parameters can be obtained by sampling from the corresponding components of the multivariate prior.

6. Numerical results for synthetic data. In this section, we compare the proposed Bayesian approach using joint GaMRF priors (i.e., the multivariate model leading to (31a-31d), denoted as GaMRF) to the method using independent $\mathcal{IG}$ priors (i.e., the novel univariate model (24) leading to (32a-32b), denoted as IG) and to the standard linear regression based method (using (7) and denoted as LF). The comparison is performed by running the different methods on a large number of independent realizations of sequences of heterogeneous synthetic multifractal images.

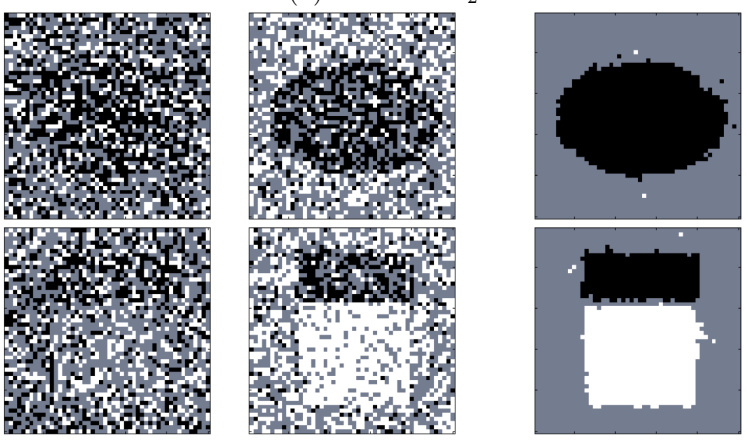
6.1. Synthetic multifractal image sequence. The scenario considered here is summarized in Fig. 2(a): Each realization of the synthetic data set consists of a sequence of 50 independent 2D multifractal random walks (MRW) of size 3200×3200. A MRW is chosen here because its multifractal properties mimic those of Mandelbrot's celebrated log-normal cascades [37] and for its ease of numerical synthesis. Its multifractal spectrum is given by (6) with $c_1 > 0.5$, $c_2 < 0$ and $c_p = 0$, $p \geq 3$ (cf., [47]). Each 2D MRW in the sequence, indexed by t, has two distinct multifractal regions whose geometry has been fixed for all t and comprises a back-



(a) 2D+1D MRW (left) and theoretical $c_2(x_1, x_2, t)$ (right)



(b) Estimates $\hat{c}_2$



54% 45% 3%

(c) K-means classification of $\hat{c}_2$ and misclassification rates

Figure 2: Estimation results for a temporal sequence of heterogeneous MRWs decomposed into $50 \times 50 \times 50$ patches of size $2^6 \times 2^6$: prescribed c_2 masks (a); estimates $\hat{c}_2$ for two different slices and overall histograms (b); classification labels obtained by histogram thresholding and misclassification rates (c).

ground with $c_2 = -0.02$ that includes an ellipse for which c_2 evolves with t according to a piece-wise constant profile. An example of a realization of heterogeneous 2D MRW (corresponding to frame $t = 30$) is displayed in Fig. 2(a) (top left). Note that the piece-wise constant evolution of c_2 (in space and along t) is intentionally chosen here as a limit test case for the robustness of the proposed approach (which assumes a smooth evolution of c_2 in the data).

6.2. Experimental setup. A Daubechies' mother wavelet with $N_\psi = 2$ vanishing moments is used in the 2D DWT. The linear regression weights w_j in (5) are chosen proportional to n_j, cf., e.g., [54, 58]. Following [17], the frequencies J_j in (11) (and, hence, $\underline{J}_j$ in (13)) are restricted to $0 < ||\boldsymbol{\omega}_\mathbf{m}|| < \pi/4$ for all Bayesian estimators. The values of the GaMRF parameters were set to $(a_i^{(1)}, a_i^{(2)}) = (10, 20)$ based on visual comparison of preliminary results obtained for a range of values for $(a_i^{(1)}, a_i^{(2)})$. The hyperparameters of the independent $\mathcal{IG}$ priors for IG were set to $(\alpha_{0,i}, \beta_{0,i}) = (10^{-3}, 10^{-3})$, which ensures that they closely resemble a non-informative Jeffreys' prior. The estimation is performed on a decomposition of the cube into $N_{x_1} \times N_{x_2} \times N_t = 50 \times 50 \times 50$ patches of size $2^6 \times 2^6$. The estimation performance is quantified as the average $\mathrm{m} \triangleq \widehat{\mathbb{E}}[\hat{c}_2]$, the standard deviation (STD) $\mathrm{s} = (\widehat{\mathrm{Var}}[\hat{c}_2])^{\frac{1}{2}}$ and the root mean squared error (RMSE) $\mathrm{rms} = \sqrt{(\mathrm{m} - c_2)^2 + \mathrm{s}^2}$, where $\widehat{\mathbb{E}}$ and $\widehat{\mathrm{Var}}$ stand for the sample mean and variance, respectively, evaluated over 100 independent realizations.

6.3. Results.

6.3.1. Illustration for a single realization. We first illustrate the performance of the different estimation methods for one single realization of the above described sequence of synthetic images. Fig. 2(b) plots estimates $\hat{c}_2$ for frame $t = 10$ (first row), and a slice along t for $x_2 = 25$ (second row) together with the histograms of estimates (third row) for LF, IG and GaMRF (left, center and right column, respectively). The corresponding theoretical values for c_2 are plotted in Fig. 2(a).

Clearly, LF exhibits strong spatial and temporal variability and fails to provide a smooth evolution of the multifractality in the dataset. The Bayesian estimator IG with non-informative prior improves the estimation accuracy with respect to LF and enables the visual identification (in time and space) of the zones with different multifractality, yet estimates obtained with IG still display strong variability and their histogram does not reveal the existence of three distinct zones of multifractality in the data. In contrast to these univariate estimators, the proposed GaMRF estimator provides more satisfactory results with increased spatial and temporal coherence and significantly reduced variability of the estimates. In particular, the estimates obtained with GaMRF lead to histograms in which the three different values for c_2 in the data are reflected as pronounced and well separated peaks.

A more quantitative analysis of these results is proposed in Fig. 2(c), which shows the results of a classification of the estimates, obtained by histogram thresholding using the k-means algorithm with 3 classes (the classes have been attributed in order to yield lowest misclassification error). The misclassification rates achieved by the different algorithms are 54% for LF and 45% for IG, but only 3% for GaMRF, thus showing the ability of the proposed procedure to perform a relevant multivariate estimation for c_2.

	LF	IG	GaMRF
$\lvert m - c_2 \rvert$	0.0057	0.0017	0.0023
s	0.038	0.011	0.0016
rmse	0.039	0.011	0.0029

Table 1: Absolute values of bias $\lvert m - c_2 \rvert$, standard deviations s and RMSE values rmse for the different estimators (results obtained for 100 independent realizations).

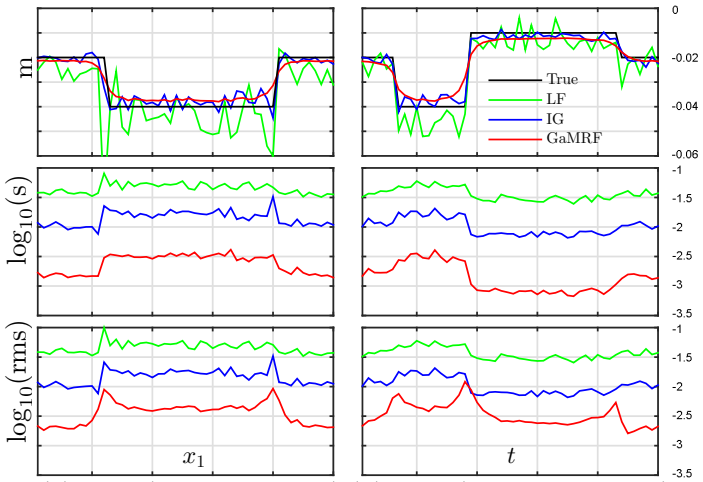
6.3.2. Estimation performance. Fig. 3 plots estimation performance, evaluated over 100 independent realizations, as a function of x (for $t = 10, x_2 = 25$) and of t (for $x_1 = 25, x_2 = 25$). The overall performance for the image sequence is given in Tab. 1. First, a comparison of the average of estimates for the Bayesian estimators leads to the conclusion that, despite the departure of the scenario considered from the assumption of slow evolution for c_2, the GaMRF estimator yields average profiles close to that of IG. Only close to sharp transitions for the value of c_2 does GaMRF introduce some bias due to the smoothing effect of the prior. Yet, this effect remains confined to ± 3 neighboring patches and has little impact on the overall bias reported in Tab. 1. Estimates obtained with LF are found to have the largest (by a factor of 3) bias. Second, while the Bayesian estimator IG with non-informative prior already yields a remarkable reduction of variability compared to LF (STD values are divided by 4), the proposed multivariate GaMRF estimator further and dramatically decreases STD to values that are more than one order of magnitude below those of LF. This is also reflected by the overall STD and RMSE values reported in Tab. 1, which are more than one order of magnitude better for GaMRF than for LF. Due to the bias introduced by GaMRF close to sharp transitions of the value of c_2, local RMSE values range from 25% (close to transitions) to only 4% (in homogeneous areas) of those of LF. Finally, note that these significant performance gains of GaMRF are achieved at very reasonable computational cost. As an example, the analysis of a $1024 \times 1024 \times 50$ data cube using patches of size 64×64 takes about 100 seconds for GaMRF on a standard desktop computer [2], which is only 4 times more than what the LF method requires. The cost of IG is similar to that of GaMRF. Note that the direct space-domain statistical model of Section 3.1 leads to two orders of magnitude larger computational cost (and, by construction, similar performance as IG) [17]. Overall, these results demonstrate the clear practical benefits of the proposed procedure for the multifractal analysis of multivariate images.

7. Application to real-world images.

Finally we illustrate the application of the proposed joint estimator for the multifractality parameter to two real-world multivariate remote sensing images of different natures: a hyperspectral (HS) image, and a multi-temporal (MT) image.

7.1. Application to a hyperspectral image. The HS image under study corresponds to a forested area near a city and was acquired by the Hyspex hyperspectral scanner during the Madonna project [51]. It contains 960×1952 pixels with a spatial resolution of 0.5 meters and 160 spectral bands ranging from the visible to near infrared. In our numerical experiment,

[2] Using Matlab, a 3.40 Ghz Intel Core i7 processor and 8GB RAM



(a) Line $(t = 10, x_2 = 25)$ (b) Line $(x_1 = 25, x_2 = 25)$

Figure 3: Estimation performance for heterogeneous 2D MRWs: mean (first row), standard deviation (second row) and root-mean square error (third row) in (a) spatial direction x_1 $(t = 10, x_2 = 25)$ and (b) temporal direction t $(x_1 = 25, x_2 = 25)$.

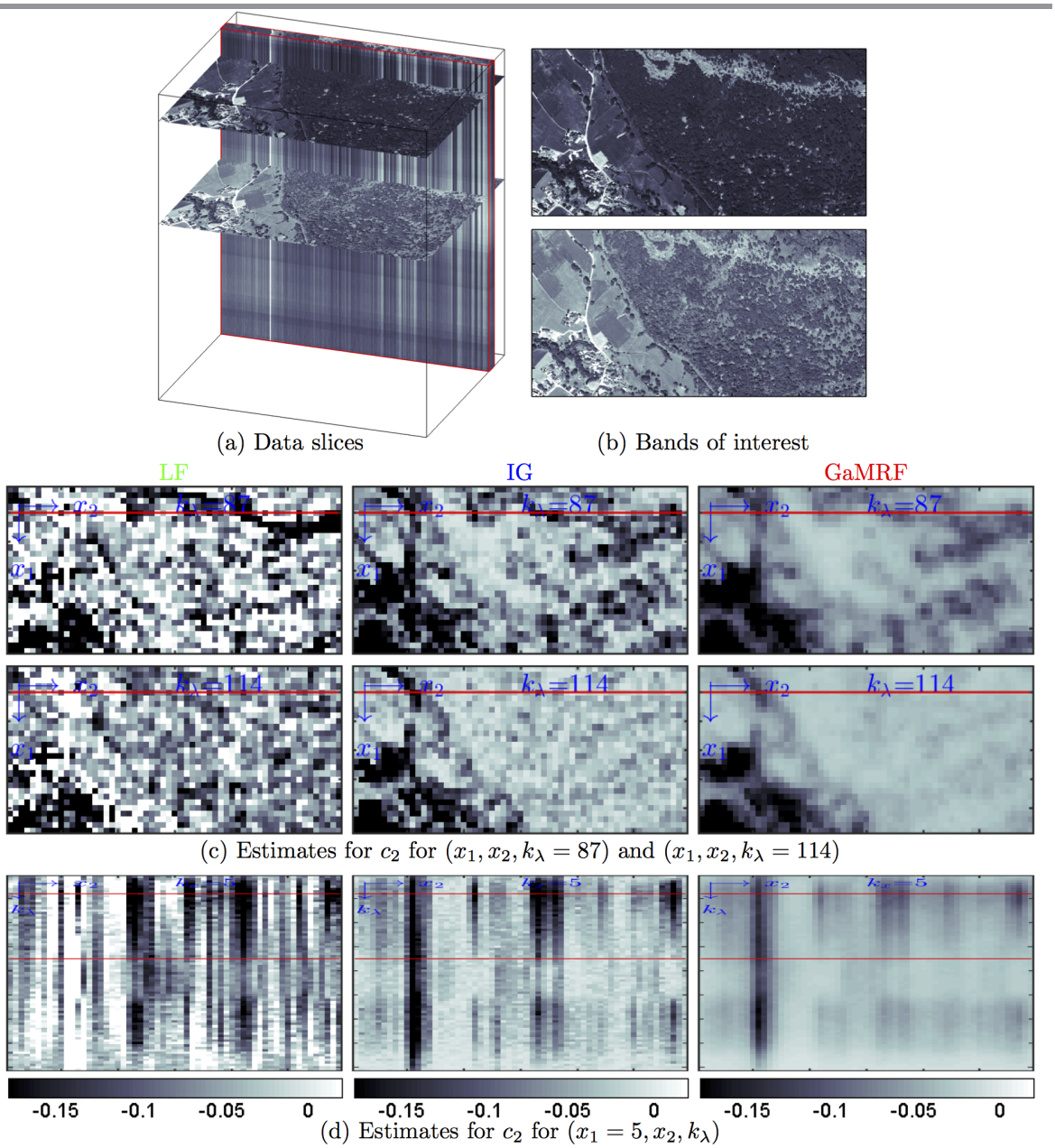
the 80 last bands are analyzed. Each band is decomposed into 29×60 patches of size 64×64 pixels, with 50% overlap, resulting in a decomposition into $29 \times 60 \times 80$ patches indexed by (x_1, x_2, k_λ), where k_λ stands for the spectral dimension. Overlapping patches are chosen here in order to increase the spatial resolution and to illustrate the robustness of the model (even if the independence assumption between patches is clearly violated when they overlap).

The subplots in Fig. 4 report the estimates for the multifractality parameter c_2 provided by LF, IG and GaMRF for two representative spectral bands (c) (the bands 87 and 114, which are plotted in (b)) as well as for a slice along the spectral dimension (d) (the corresponding $64 \times 1962 \times 80$ portion of the HS cube is indicated by a red frame in (a)). A visual inspection of the results for the bands 87 and 114 reveals that the strong spatial variability of the estimates obtained with LF prevents the identification of distinct regions in the image (with the exception perhaps of the city in the left bottom corner which yields clusters of strongly negative c_2). The estimator IG yields better spatial coherence and clearly improves over the estimates obtained with LF (see [17] for a similar experiment for one single spectral band using the model (10-11), leading to the same conclusions). Yet, the variability within visually homogeneous zones of the dataset (e.g., the forested region) is still important. In comparison with IG, and a fortiori with LF, the proposed GaMRF method further and dramatically reduces the variability within presumably homogeneously multifractal zones, inducing strong spatial coherence, which reinforces the contrast between regions of different multifractalities and visually sharpens their borders. As a result, these estimates can potentially reveal hidden underlying structures in the data set. Despite the absence of a ground truth for this data set we can, for instance, observe that the forested area in the right half of the image is homogeneous (with $c_2 \sim -0.05$) in the spectral band $k_\lambda = 114$, while it is composed of a background (where $c_2 \sim -0.05$) and scattered clusters (where $c_2 \sim -0.1$) in the spectral band $k_\lambda = 87$, which

could arguably indicate a physical change at this location (e.g., a lower tree density). Finally, the spectral evolution of c_2, plotted in Fig. 4(d), reveals the strong spectral redundancy that is generally observed for HS data, cf., e.g., [12]. Nevertheless, the proposed GaMRF estimator permits a reduction in the variance of estimates along the spectral dimension and yields a visually less noisy evolution of c_2 across the bands of the image, while maintaining its main features.

7.2. Application to a multi-temporal image. We further perform the multifractal analysis of an MT image with 11 time frames. Each frame consists of a 900×900 pixel image with a spatial resolution of about 0.2 meters, captured in 140 spectral band recorded from the visible to near infrared (415 to 990 nm; only one single band is analyzed here). The images were acquired over the same rural scene at Porton Down in the U.K. over a 2-day period (data courtesy of DSTL). The recorded scene comprises several roads and tracks between open fields, a few trees, as well as two man-made "targets" (visible as distinct white crosses). Two frames for $t = 2$ and $t = 8$ are plotted in Fig. 5 (a), showing that the image is subject to considerable changes in (partial) illumination across time. As in the previous section, each frame is decomposed in patches of size 64×64 pixels, with 50% overlap, resulting in a decomposition into $27 \times 27 \times 11$ patches, indexed by (x_1, x_2, t). Visual inspection of the estimates for c_2 provided by LF, IG and GaMRF for the frames $t = 2$ and $t = 8$, plotted in Fig. 5(b), first leads to conclusions similar to those obtained in Section 7.1: the strong variability of LF prevents from identifying any of the spatial image features; the IG estimates yield acceptable spatial coherence; the GaMRF estimator leads to the visually most satisfactory results and yields much less speckled estimates for the image regions that can be considered homogeneous (i.e., the fields) as well as sharp contrast for the artificial structures (roads, targets). In addition, it can be observed that the proposed GaMRF estimator permits better tracking of image features across time. This is, for instance, the case for the road indicated by red arrows in Fig. 5(b), which is at best partially visible in frame $t = 8$ for IG (and not at all for LF), probably due to insufficient illumination for this frame, but can be clearly and easily identified for the estimates obtained with GaMRF. The evolution of c_2 across time is further investigated in Fig. 5(c) and (d), where a longitudinal and a lateral slice of the multi-temporal estimates are plotted. Visual inspection leads us to conclude that GaMRF yields very consistent estimates across time. Since the multifractal parameters are invariant to (spatially smooth) changes in illumination, this is to be expected for a scene in which only the illumination varies but the scene itself does not vary. In contrast, IG and even more so LF display significant temporal noise. As a result, the proposed GaMRF estimator enables us to coherently render particular image features across time that are harder to identify for IG and LF, e.g., for the zone of strong multifractality ($c_2 \simeq -0.1$) in the vicinity of the trees in the upper left corner of the image frames, corresponding to the vertical strip of estimates for c_2, indicated by orange arrows in Fig. 5(d).

8. Discussion and Conclusion. The present paper introduced a novel Bayesian procedure for the joint estimation of the multifractality parameter for (patches of) multivariate images. It builds on two original key contributions. First, it is based on an extended likelihood for the Fourier coefficients of log-leaders of a single image, separable in the parameters of interest (c_2, c_2^0). Second, it models the collection of unknown (multifractal) parameters c_2 of a



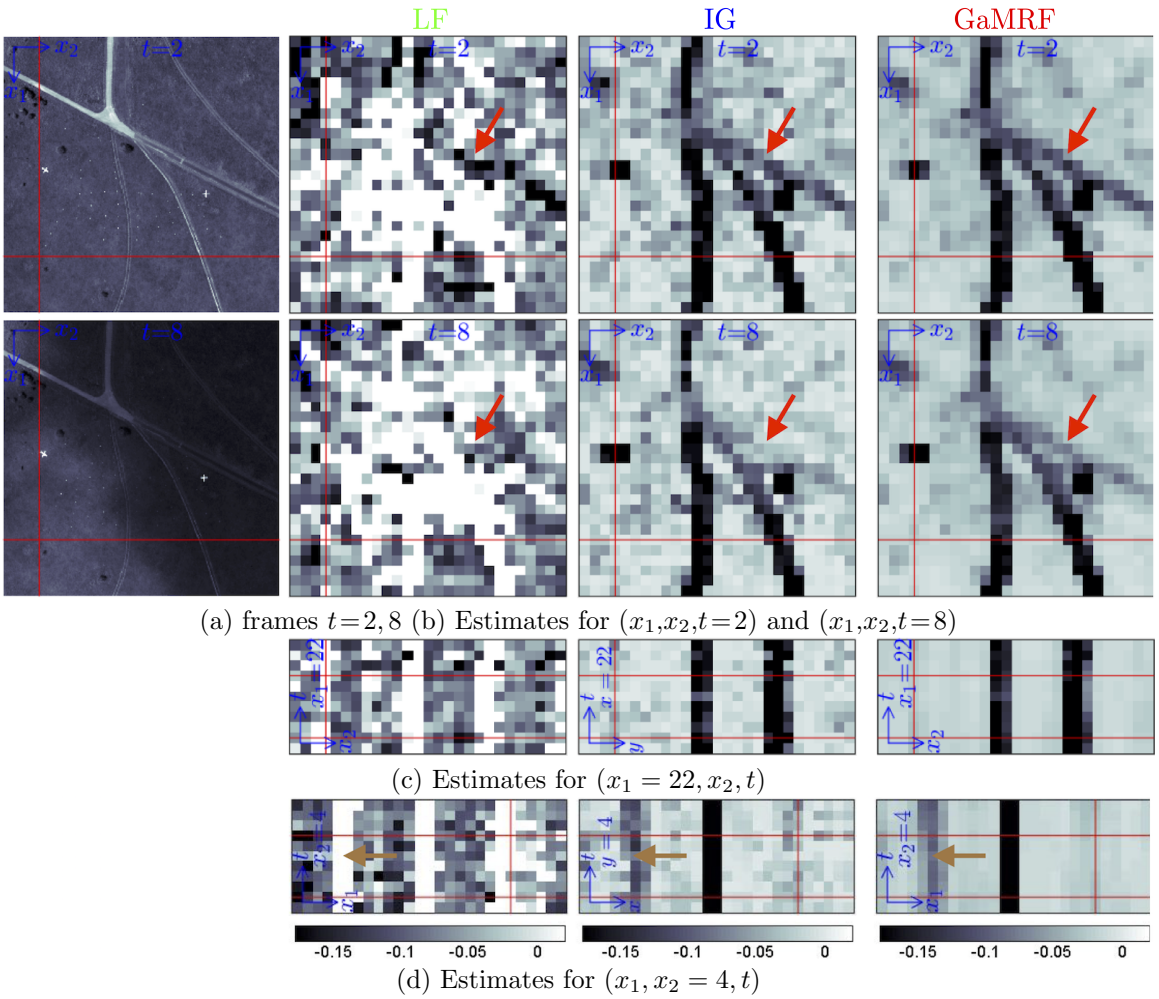
(a) Data slices (b) Bands of interest

(c) Estimates for c_2 for $(x_1, x_2, k_\lambda = 87)$ and $(x_1, x_2, k_\lambda = 114)$

(d) Estimates for c_2 for $(x_1 = 5, x_2, k_\lambda)$

Figure 4: Multifractal analysis of a hyperspectral image (bands $80 - 160$) from the Madonna project.

multivariate image by gamma Markov random field joint priors, accounting for the assumption that multifractal properties evolve smoothly in privileged (temporal/spectral, spatial) directions within the sequence of images and inducing regularization. Together, these two ingredients lead to a Bayesian model for the multifractality parameters of multivariate im-



(a) frames $t=2,8$ (b) Estimates for $(x_1,x_2,t=2)$ and $(x_1,x_2,t=8)$

(c) Estimates for $(x_1 = 22, x_2, t)$

(d) Estimates for $(x_1, x_2 = 4, t)$

Figure 5: Multifractal analysis of a multi-temporal image.

ages for which the associated Bayesian estimators can be approximated efficiently using a Gibbs sampler (without requiring any Metropolis-Hastings acceptance-reject move). To the best of our knowledge, the method constitutes the first operational multifractal analysis tool applicable to the joint analysis of multidimensional sets of images. Numerical experiments, conducted on synthetic data with heterogeneous multifractal properties as well as on two real-world hyperspectral and multi-temporal images, demonstrated the excellent performance of the Bayesian joint estimator, which significantly outperformed previously existing (univariate) methods. Moreover, the proposed Bayesian model results in a competitive computational cost, of the order of 4 times the cost of linear regression based estimation. Future work will include the incorporation of additional log-cumulants c_p in the statistical model and the investigation of procedures for estimating the GaMRF regularization hyperparameters.

REFERENCES

[1] P. Abry, S. Jaffard, and H. Wendt, *When Van Gogh meets Mandelbrot: Multifractal classification of painting's texture*, Signal Proces., 93 (2013), pp. 554–572.

[2] V. V. Anh and K. E. Lunney, *Parameter estimation of random fields with long-range dependence*, Math. Comput. Model., 21 (1995), pp. 67–77.

[3] J.-P. Antoine, R. Murenzi, P. Vandergheynst, and S. T. Ali, *Two-Dimensional Wavelets and their Relatives*, Cambridge University Press, 2004.

[4] A. Arneodo, N. Decoster, P. Kestener, and S. G. Roux, *A wavelet-based method for multifractal image analysis: from theoretical concepts to experimental applications*, in Adv. Imag. Electr. Phys., vol. 126, Academic Press, 2003, pp. 1–98.

[5] C. L. Benhamou, S. Poupon, E. Lespessailles, S. Loiseau, R. Jennane, V. Siroux, W. J. Ohley, and L. Pothuaud, *Fractal analysis of radiographic trabecular bone texture and bone mineral density: two complementary parameters related to osteoporotic fractures*, J. Bone Miner. Res., 16 (2001), pp. 697–704.

[6] J. Beran, *Statistics for Long-Memory Processes*, Chapman & Hall, New York, 1994.

[7] J. Besag, *Statistical analysis of non-lattice data*, The Statistician, (1975), pp. 179–195.

[8] J. Bruna, S. Mallat, E. Bacry, and J.-F. Muzy, *Intermittent process analysis with scattering moments*, Ann. Stat., 43 (2015), pp. 323–351.

[9] B. Castaing, Y. Gagne, and M. Marchand, *Log-similarity for turbulent flows?*, Physica D, 68 (1993), pp. 387–400, https://doi.org/http://dx.doi.org/10.1016/0167-2789(93)90132-K.

[10] P. Chainais, *Infinitely divisible cascades to model the statistics of natural images*, IEEE T. Pattern Anal. Mach. Intell., 29 (2007), pp. 2105–2119.

[11] N. H. Chan and W. Palma, *Estimation of long-memory time series models: A survey of different likelihood-based methods*, Adv. Econom., 20 (2006), pp. 89–121.

[12] C.-I. Chang, Q. Du, T.-L. Sun, and M. L. G. Althouse, *A joint band prioritization and band-decorrelation approach to band selection for hyperspectral image classification*, IEEE T. Geosc. Remote Sensing, 37 (1999), pp. 2631–2641.

[13] J. Coddington, J. Elton, D. Rockmore, and Y. Wang, *Multifractal analysis and authentication of Jackson Pollock paintings*, in Proc. SPIE 6810, 2008, p. 68100F.

[14] S. Combrexelle and H. Wendt, https://www.irit.fr/~Herwig.Wendt/MVMFA.html.

[15] 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. IEEE Int. Conf. Acoust., Speech, and Signal Process. (ICASSP), Shanghai, China, March 2016.

[16] S. Combrexelle, H. Wendt, Y. Altmann, J.-Y. Tourneret, S. McLaughlin, and P. Abry, *Bayesian joint estimation of the multifractality parameter of image patches using gamma markov random field priors*, in Proc. IEEE Int. Conf. Image Proc. (ICIP), Phoenix, AZ, USA, September 2016.

[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., 24 (2015), pp. 2540–2551.

[18] S. Combrexelle, H. Wendt, J.-Y. Tourneret, P. Abry, and S. McLaughlin, *Bayesian multifractal analysis of multi-temporal images using smooth priors*, in Proc. IEEE Workshop Statistical Signal Proces. (SSP), Palma de Mallorca, Spain, June 2016.

[19] O. Dikmen and A. Cemgil, *Gamma Markov random fields for audio source modeling*, IEEE T. Audio, Speech, and Language Proces., 18 (2010), pp. 589–601, https://doi.org/10.1109/TASL.2009.2031778.

[20] D. A. v. Dyk and X.-L. Meng, *The art of data augmentation*, J. of Comput. and Graphical Stat., 10 (2001), pp. pp. 1–50.

[21] M. Fuentese, *Approximate likelihood for large irregularly spaced spatial data*, J. Am. Statist. Assoc., 102 (2007), pp. 321–331.

[22] N. R. Goodman, *Statistical analysis based on a certain multivariate complex Gaussian distribution (an introduction)*, Ann. Math. Stat., 34 (1963), pp. pp. 152–177.

[23] T. C. Halsey, M. H. Jensen, L. P. Kadanoff, I. Procaccia, and B. I. Shraiman, *Fractal measures*

and their singularities: The characterization of strange sets, Phys. Rev. A, 33 (1986), pp. 1141–1151, https://doi.org/10.1103/PhysRevA.33.1141.

[24] R. M. Haralick, *Statistical and structural approaches to texture*, Proc. of the IEEE, 67 (1979), pp. 786–804.

[25] G. E. Hinton, *Training products of experts by minimizing contrastive divergence*, Neural computation, 14 (2002), pp. 1771–1800.

[26] S. Jaffard, *Wavelet techniques in multifractal analysis*, in Fractal Geometry and Applications: A Jubilee of Benoît Mandelbrot, Proc. Symp. Pure Math., M. Lapidus and M. van Frankenhuijsen, eds., vol. 72(2), AMS, 2004, pp. 91–152.

[27] 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 and N. Cohen, eds., World scientific publishing, Singapore, 2015, pp. 31–116.

[28] S. Jaffard, S. Seuret, H. Wendt, R. Leonarduzzi, S. Roux, and P. Abry, *Multivariate multifractal analysis*, Applied and Computational Harmonic Analysis, (2018). in press.

[29] H. Ji, X. Yang, H. Ling, and Y. Xu, *Wavelet domain multifractal analysis for static and dynamic texture classification*, IEEE T. Image Proces., 22 (2013), pp. 286–299.

[30] C. R. Johnson, P. Messier, W. Sethares, A. Klein, C. Brown, A. Do, P. Klausmeyer, P. Abry, S. Jaffard, H. Wendt, S. Roux, N. Pustelnik, N. van Noord, L. van der Maaten, E. Potsma, J. Coddington, L. Daffner, H. Murata, H. Wilhelm, S. Wood, and M. Messier, *Pursuing automated classification of historic photographic papers from raking light photomicrographs*, J. Amer. Inst. Conserv., 53 (2014), pp. 159–170.

[31] P. Kestener, J. Lina, P. Saint-Jean, and A. Arneodo, *Wavelet-based multifractal formalism to assist in diagnosis in digitized mammograms*, Image Analysis and Stereology, 20 (2001), pp. 169–175.

[32] J. Lévy-Véhel, P. Mignot, and J. Berroir, *Multifractals, texture and image analysis*, in Proc. IEEE Conf. Comp. Vis. Pattern Recognition (CVPR), Champaign, IL, USA, June 1992, pp. 661–664.

[33] R. Lopes and N. Betrouni, *Fractal and multifractal analysis: A review*, Medical Image Analysis, 13 (2009), pp. 634–649.

[34] S. Lovejoy and D. Schertzer, *The weather and climate: emergent laws and multifractal cascades*, Cambridge University Press, 2013.

[35] T. Lux, *Higher dimensional multifractal processes: A GMM approach*, J. Business & Economic Stat., 26 (2007), pp. 194–210.

[36] S. Mallat, *A Wavelet Tour of Signal Processing*, Academic Press, 3rd ed., 2008.

[37] B. B. Mandelbrot, *Intermittent turbulence in self-similar cascades: divergence of high moments and dimension of the carrier*, J. Fluid Mech., 62 (1974), pp. 331–358, https://doi.org/10.1017/S0022112074000711.

[38] C. Meneveau, K. Sreenivasan, P. Kailasnath, and M. Fan, *Joint multifractal measures: Theory and applications to turbulence*, Phys. Rev. A, 41 (1990), p. 894.

[39] E. Moulines, F. Roueff, and M. Taqqu, *A wavelet Whittle estimator of the memory parameter of a nonstationary Gaussian time series*, Ann. Stat., (2008), pp. 1925–1956.

[40] M. Pereyra, N. Whiteley, C. Andrieu, and J.-Y. Tourneret, *Maximum marginal likelihood estimation of the granularity coefficient of a Potts-Markov random field within an MCMC algorithm*, in Proc. IEEE Workshop on Statistical Signal Proces. (SSP), June 2014, pp. 121–124, https://doi.org/10.1109/SSP.2014.6884590.

[41] L. Ponson, D. Bonamy, H. Auradou, G. Mourot, S. Morel, E. Bouchaud, C. Guillot, and J. Hulin, *Anisotropic self-affine properties of experimental fracture surface*, J. Fracture, 140 (2006), pp. 27–36.

[42] J. Prats-Montalbn, A. de Juan, and A. Ferrer, *Multivariate image analysis: A review with applications*, Chemometrics and Intelligent Laboratory Systems, 107 (2011), pp. 1–23, https://doi.org/http://dx.doi.org/10.1016/j.chemolab.2011.03.002.

[43] T. S. Rao and R. Chandler, *A frequency domain approach for estimating parameters in point process models*, in Athens Conf. Applied Probability and Time Series Analysis, Springer, 1996, pp. 392–405.

[44] R. H. Riedi, *Multifractal processes*, in Theory and applications of long range dependence, P. Doukhan, G. Oppenheim, and M. Taqqu, eds., Birkhäuser, 2003, pp. 625–717.

[45] L. Risser, T. Vincent, F. Forbes, J. Idier, and P. Ciuciu, *Min-max extrapolation scheme for fast*

estimation of 3d potts field partition functions. application to the joint detection-estimation of brain activity in fMRI, J. Signal Proces. Syst., 65 (2011), pp. 325–338.

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

[47] R. Robert and V. Vargas, *Gaussian multiplicative chaos revisited*, Ann. Proba., 38 (2010), pp. 605–631.

[48] S. G. Roux, A. Arneodo, and N. Decoster, *A wavelet-based method for multifractal image analysis. III. Applications to high-resolution satellite images of cloud structure*, Eur. Phys. J. B, 15 (2000), pp. 765–786.

[49] F. Schmitt, D. Schertzer, S. Lovejoy, and Y. Brunet, *Multifractal temperature and flux of temperature variance in fully developed turbulence*, Europhys. Lett., 34 (1996), p. 195.

[50] F. G. Schmitt and L. Seuront, *Multifractal random walk in copepod behavior*, Physica A, 301 (2001), pp. 375–396.

[51] D. Sheeren, M. Fauvel, S. Ladet, A. Jacquin, G. Bertoni, and A. Gibon, *Mapping ash tree colonization in an agricultural mountain landscape: Investigating the potential of hyperspectral imagery*, in Proc. IEEE Int. Conf. Geosci. Remote Sens. (IGARSS), Vancouver, Canada, July 2011, pp. 3672–3675.

[52] M. A. Tanner and W. H. Wong, *The calculation of posterior distributions by data augmentation*, J. Am. Stat. Ass., 82 (1987), pp. pp. 528–540.

[53] C. Vacar, J.-F. Giovannelli, and Y. Berthoumieu, *Bayesian texture classification from indirect observations using fast sampling*, IEEE Transactions on Signal Processing, 64 (2016), pp. 146–159.

[54] H. Wendt, P. Abry, and S. Jaffard, *Bootstrap for empirical multifractal analysis*, IEEE Signal Proces. Mag., 24 (2007), pp. 38–48.

[55] H. Wendt, P. Abry, S. Jaffard, H. Ji, and Z. Shen, *Wavelet leader multifractal analysis for texture classification*, in Proc. IEEE Int. Conf. Image Proces. (ICIP), Cairo, Egypt, Nov. 2009.

[56] H. Wendt, S. Jaffard, and P. Abry, *Multifractal analysis of self-similar processes*, in Proc. IEEE Workshop on Statistical Signal Proces. (SSP), Ann Arbor, MI, USA, Aug. 2012, pp. 69–72, https://doi.org/10.1109/SSP.2012.6319798.

[57] 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.

[58] H. Wendt, S. G. Roux, S. Jaffard, and P. Abry, *Wavelet leaders and bootstrap for multifractal analysis of images*, Signal Proces., 89 (2009), pp. 1100–1114, https://doi.org/http://dx.doi.org/10.1016/j.sigpro.2008.12.015.

[59] P. Whittle, *On stationary processes in the plane*, Biometrika, 41 (1954), pp. 434–449.

[60] G. Wornell and A. V. Oppenheim, *Estimation of fractal signals from noisy measurements using wavelets*, IEEE T. Signal Proces., 40 (1992), pp. 611–623.

[61] Y. Xu, X. Yang, H. Ling, and H. Ji, *A new texture descriptor using multifractal analysis in multi-orientation wavelet pyramid*, in Proc. IEEE Conf. Comp. Vis. Pattern Recognition (CVPR), San Francisco, CA, USA, June 2010, pp. 161–168.