BAYESIAN MULTIFRACTAL ANALYSIS OF MULTI-TEMPORAL IMAGES USING SMOOTH PRIORS

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

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

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

[3] School of Engineering and Physical Sciences, Heriot-Watt University, Edinburgh, UK, s.mclaughlin@hw.ac.uk

ABSTRACT

Texture analysis can be conducted within the mathematical framework of multifractal analysis (MFA) via the study of the regularity fluctuations of image amplitudes. Successfully used in various applications, however MFA remains limited to the independent analysis of single images while, in an increasing number of applications, data are multi-temporal. The present contribution addresses this limitation and introduces a Bayesian framework that enables the joint estimation of multifractal parameters for multi-temporal images. It builds on a recently proposed Gaussian model for wavelet leaders parameterized by the multifractal attributes of interest. A joint Bayesian model is formulated by assigning a Gaussian prior to the second derivatives of time evolution of the multifractal attributes associated with multi-temporal images. This Gaussian prior ensures that the multifractal parameters have a smooth temporal evolution. The associated Bayesian estimators are then approximated using a Hamiltonian Monte-Carlo algorithm. The benefits of the proposed procedure are illustrated on synthetic data.

***Index Terms*—** Multifractal Analysis, Bayesian Estimation, Texture Analysis, Multivariate image, Hamiltonian Monte Carlo

1. INTRODUCTION

Context. In image processing, texture characterization is a standard problem for which different paradigms have been proposed. Multifractal analysis (MFA) is one such and characterizes a texture via the local fluctuations of the point-wise regularity of image amplitudes. More specifically, the texture is encoded in the so-called *multifractal spectrum* $D(h)$ collecting the Hausdorff dimensions of the sets of points sharing the same local regularity, classically measured by the Hölder exponent h, cf., e.g., [1–4] and references therein. Multifractal models are deeply tied to scale invariance properties. These properties are practically assessed via the exponents $\zeta(q)$ of the power laws, over a range of scales 2^j, of the sample moments of multiresolution quantities $T(j,\boldsymbol{k})$ of the image $\mathbf{X}$ at spatial position $\boldsymbol{k}$ and scale 2^j

$$S(q,j) \triangleq \frac{1}{n_j}\sum_{\boldsymbol{k}} |T(j,\boldsymbol{k})|^q \simeq (2^j)^{\zeta(q)},\ j_1 \le j \le j_2 \qquad (1)$$

where n_j is the number of multiresolution quantities at scale 2^j. The present work dwells on the *wavelet leaders* $l(j,\boldsymbol{k})$ that have been shown to match MFA purposes [1, 4] (their definition is recalled in Section 2). These so-called *scaling exponents* $\zeta(q)$ are intimately linked to the multifractal spectrum $D(h)$ via the Legendre transform $D(h) \le \mathcal{L}(h) \triangleq \inf_{q\in\mathbb{R}}[2+qh-\zeta(q)]$. This link enables the practical discrimination between multifractal models through the characterization of $\zeta(q)$. Notably, the two most prominent classes of processes used to model scale invariance in data are the self-similar processes [5], for which $\zeta(q)$ is linear in the vicinity of $q=0$, and multifractal multiplicative cascade (MMC) based processes [6], associated with a strictly concave function $\zeta(q)$. This discrimination can be efficiently handled via the polynomial expansion of $\zeta(q)$ (at $q=0$), $\zeta(q) = \sum_{m\ge 1} c_m q^m/m!$, by considering the coefficients c_m, termed log-cumulants [4, 7, 8]. It can be shown that the second log-cumulant c_2, referred to as intermittency or *multifractality parameter*, is strictly negative for multiplicative cascades whereas it is identically zero for self-similar processes, cf., e.g, [1,9]. The estimation of c_2 is thus of paramount importance in MFA since it enables the identification of the model that best fits the data. For more details on MFA, the reader is referred to, e.g., [1–4, 10].

Estimation of c_2. The log-cumulants c_m have been shown to be directly tied to the cumulants of the logarithm of the multiresolution quantities [7]. Specifically, for the multifractality parameter c_2

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

In view of (2), the classical estimation procedure for c_2 is defined as the linear regression of the sample variance $\widehat{\text{Var}}\,[\cdot]$ of the log-leaders with respect to scale j

$$\hat{c}_2 = \frac{1}{\ln 2}\sum_{j=j_1}^{j_2} w_j\, \widehat{\text{Var}}\,[\ln l(j,\boldsymbol{k})] \qquad (3)$$

where w_j are appropriate regression weights [4, 9, 11]. This estimator is widely used and known to provide relatively poor performance, in particular for small image size. Different attempts to improve estimation performance are reported in the literature. A generalized method of moments has been proposed [12]. This procedure relies on fully parametric models that are often too restrictive in real-world applications. More recently, Bayesian estimators for c_2 have been proposed [13, 14]. This approach builds on a semi-parametric model for the multivariate statistics of the log-leaders whose variance-covariance structure is controlled by the pair (c_2, c_2^0). The resulting Gaussian likelihood is numerically evaluated using a closed-form Whittle approximation, and the Bayesian inference is accomplished by a Markov chain Monte Carlo (MCMC) algorithm with Metropolis-Hasting within Gibbs (MHG) moves using random walk proposals. However, these approaches are all designed for processing single images only and cannot be used to take into account the information that is potentially jointly conveyed in a collection of several multi-temporal images.

Goals and contributions. The goal of the present work is to introduce a novel Bayesian framework suitable for the joint estimation of c_2 for image sequences. Elaborating on the statistical model

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.

formulated for a single image in previous work [13, 14] (recalled in Section 2), this is achieved through two original key contributions detailed in Section 3. First, the assumption that multifractality evolves smoothly across the collection of images is encoded via a Gaussian prior on the second derivatives of the evolution of the multifractal attributes. The degree of the induced smoothness is controlled by hyperparameters whose estimation is also embedded in a full Bayesian model. Second, the computation of Bayesian estimators associated with the proposed model is achieved via an MCMC algorithm relying on a Hamiltonian Monte Carlo scheme that permits the efficient exploration of the non-standard conditional distributions in high-dimensional variable space induced by the model. The performance of the proposed estimation procedure is analyzed in Section 4 for sequences of synthetic multifractal multi-temporal images with prescribed temporal evolutions of its parameters. The results indicate that the proposed method compares very favorably against the linear regression (3) as well as with previous Bayesian formulations [13, 14].

2. STATISTICAL MODEL FOR LOG-LEADERS

2.1. Wavelet leaders

Denote $\chi(x)$ and $\psi(x)$ a scaling function and a mother wavelet (characterized by $N_\psi \geq 1$ vanishing moments) defining a 1D wavelet transform. Two-dimensional wavelets can be constructed from the tensorial products $\psi^{(0)}(\mathbf{x}) = \chi(x_1)\chi(x_2)$, $\psi^{(1)}(\mathbf{x}) = \psi(x_1)\chi(x_2)$, $\psi^{(2)}(\mathbf{x}) = \chi(x_1)\psi(x_2)$, $\psi^{(3)}(\mathbf{x}) = \psi(x_1)\psi(x_2)$. Under certain admissibility conditions on ϕ, the collection of dilated (to scale 2^j) and translated (to location $\mathbf{x} = 2^j\boldsymbol{k}$) templates $\{\psi^{(m)}_{j,k}(\mathbf{x}) = 2^{-j}\psi^{(m)}(2^{-j}\mathbf{x} - k)\}$ forms a basis of L^2. The discrete wavelet transform coefficients of an image $\mathbf{X}$ are defined as the inner product $d^{(m)}_{\mathbf{X}}(j,k) = \langle \mathbf{X}, \psi^{(m)}_{j,k}\rangle$, $m = 0, \ldots, 3$ [15]. The wavelet leaders $l(j,\boldsymbol{k})$ are defined as the supremum of the wavelet modulus taken in a spatial neighborhood of $2^j\boldsymbol{k}$ over all finer scales $j' \leq j$

$$l(j,\boldsymbol{k}) \triangleq \sup_{m\in(1,2,3),\lambda'\subset 9\lambda_{j,\boldsymbol{k}}} |d^{(m)}_X(\lambda')| \tag{4}$$

where $9\lambda_{j,\boldsymbol{k}}$ denotes the 9 neighboring dyadic cubes of width 2^j centered around spatial location $\boldsymbol{k}2^j$, see [1, 4] for more details.

2.2. Statistical model

Recent results reported in [13, 14] suggest that the log-leaders $\ell(j,\boldsymbol{k}) \triangleq \ln l(j,\boldsymbol{k})$ can be modeled by a Gaussian random field whose covariance $\mathrm{Cov}[\ell(j,\boldsymbol{k}),\ell(j,\boldsymbol{k}+\boldsymbol{\Delta k})]$ is approximated by a radial symmetric function parametrized by $\boldsymbol{\theta} = [\theta_1,\theta_2]^T \triangleq [c_2, c_2^0]^T$, defined as follows

$$\varrho_j(\boldsymbol{\Delta k};\boldsymbol{\theta}) \triangleq \begin{cases} C_2(j) - \frac{c_2 \ln(2^j r_j^0/3)+c_2^0}{\ln 4}\ln(1+\|\boldsymbol{\Delta k}\|) & \|\boldsymbol{\Delta k}\| \leq 3 \\ c_2 \ln(\|\boldsymbol{\Delta k}\|/r_j^0) & 3 < \|\boldsymbol{\Delta k}\| \leq r_j^0 \\ 0 & r_j^0 < \|\boldsymbol{\Delta k}\| \end{cases} \tag{5}$$

with $r_j^0 = \sqrt{n_j}/4$. The mean of the log-leaders is linked only to c_1 and is empirically removed from the model, $\bar{\ell}(j,\boldsymbol{k}) \triangleq \ell(j,\boldsymbol{k}) - \widehat{\mathbb{E}}[\ell(j,.)]$, and the centered log-leaders $\bar{\ell}$ are gathered in the vector $\boldsymbol{\ell}_j$, where $\widehat{\mathbb{E}}[\cdot]$ is the sample mean. By assuming moreover independence between scales, the likelihood for the vector $\boldsymbol{\ell} = [\boldsymbol{\ell}_{j_1}^T, ..., \boldsymbol{\ell}_{j_2}^T]^T$ can be expressed as

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

where $\boldsymbol{\Sigma}_j(\boldsymbol{\theta})$ is the covariance matrix induced by ϱ_j, $|\cdot|$ denotes the determinant and ^{T} is the transpose operator.

2.3. Whittle approximation

The direct evaluation of the likelihood (6) requires computing the covariance matrix inverses $\boldsymbol{\Sigma}_j(\boldsymbol{\theta})^{-1}$. As proposed in [13, 14], this can be circumvented by replacing (6) with the numerically robust and efficient Whittle approximation [16–18]

$$p(\boldsymbol{\ell}_j|\boldsymbol{\theta}) \propto \exp\left(-\frac{1}{2}\sum_{\mathbf{m}\in J_j} \ln\phi_j(\mathbf{m};\boldsymbol{\theta}) + \frac{I_j(\mathbf{m};\boldsymbol{\ell}_j)}{\phi_j(\mathbf{m};\boldsymbol{\theta})}\right) \tag{7}$$

that evaluates the fit between the periodogram of the log-leaders $I_j(\mathbf{m},\boldsymbol{\ell}_j)$ and the spectral density $\phi_j(\mathbf{m};\boldsymbol{\theta})$ associated with the covariance model over the frequencies $\boldsymbol{\omega}_{\mathbf{m}} = 2\pi\mathbf{m}/n_j$ with $\mathbf{m} \in J_j \triangleq [\![\lfloor(-\sqrt{n_j}-1)/2\rfloor, \ldots, \sqrt{n_j} - \lfloor\sqrt{n_j}/2\rfloor]\!]^2_{\setminus 0}$. A closed-form expression of the spectral density, $\phi_j(\mathbf{m};\boldsymbol{\theta}) = \theta_1 f_{1,j}(\mathbf{m}) + \theta_2 f_{2,j}(\mathbf{m})$, has been derived in [14], where the functions $\{f_{i,j}\}$ do not depend on $\boldsymbol{\theta}$, hence enabling precomputation and storage of the vectors $\mathbf{f}_{i,j} \triangleq (f_{i,j}(\mathbf{m}))_{\mathbf{m}\in J_j}$), see [14] for details. By substitution of (7) in (6), the likelihood (6) can be approximated as

$$p(\boldsymbol{\ell}|\boldsymbol{\theta}) \propto \exp\left(-\frac{1}{2}\mathbf{1}^T\big(\ln\boldsymbol{\phi}(\boldsymbol{\theta}) + \mathbf{I}(\boldsymbol{\ell}) \oslash \boldsymbol{\phi}(\boldsymbol{\theta})\big)\right) \tag{8}$$

where $\boldsymbol{\phi}(\boldsymbol{\theta}) \triangleq \theta_1\mathbf{f}_1 + \theta_2\mathbf{f}_2$, with $\mathbf{f}_i \triangleq [\mathbf{f}_{i,j_1}^T, \ldots, \mathbf{f}_{i,j_2}^T]^T$, $\mathbf{I}(\boldsymbol{\ell}) \triangleq [(I_{j_1}(\mathbf{m},\boldsymbol{\ell}_{j_1})^T_{\mathbf{m}\in J_{j_1}}, \ldots, I_{j_2}(\mathbf{m},\boldsymbol{\ell}_{j_2})^T_{\mathbf{m}\in J_{j_2}}]^T$, $\mathbf{1}$ is an $M\times 1$ vector of ones and $\oslash$ denotes the component-wise division operator.

3. SMOOTH BAYESIAN ESTIMATION

Based on the above model and its approximation (8), we specify a Bayesian model for a sequence of images $\{\mathbf{X}_t\}_{t=1}^M$.

3.1. Bayesian model

Likelihood. We write $\boldsymbol{\ell}^t$ and $\boldsymbol{\theta}^t \triangleq [\theta_1^t, \theta_2^t]^T$ for the centered log-leader and parameter vectors associated with the image $\mathbf{X}_t$, respectively, and use the notations $\mathbf{L} \triangleq \{\boldsymbol{\ell}^t\}_{t=1}^M$ and $\boldsymbol{\Theta} = \{\boldsymbol{\theta}_1, \boldsymbol{\theta}_2\}$ with $\boldsymbol{\theta}_i = \{\theta_i^t\}_{t=1}^M$. Assuming independence between leaders of different images, the likelihood of $\mathbf{L}$ is given by

$$p(\mathbf{L}|\boldsymbol{\Theta}) \propto \prod_{t=1}^M p(\boldsymbol{\ell}^t|\boldsymbol{\theta}^t). \tag{9}$$

Priors. A smooth evolution of the parameters $\boldsymbol{\theta}^t$ is enforced by assigning a Gaussian distribution to the second order differences of $\boldsymbol{\theta}_i$

$$p(\boldsymbol{\theta}_i|\epsilon_i^2) \propto \left(\frac{1}{\epsilon_i^2}\right)^{\frac{M-1}{2}} \exp\left(-\frac{1}{2\epsilon_i^2}\|\boldsymbol{D}\boldsymbol{\theta}_i\|^2\right) \tag{10}$$

where $\boldsymbol{D}$ is the Laplacian operator. This prior models the parameters $\boldsymbol{\theta}_i$ by a simultaneous autoregression (SAR) and has been considered in various applications, e.g., for image deconvolution [19]. It constrains the second derivative of parameters of interest to be small and hence promotes smoothness.

Hyperpriors. The degree of smoothing depends on the values of hyperparameters ϵ_i^2. We adopt here a full Bayesian strategy by including them in the unknown parameter vector and assigning them a

non-informative Jeffreys' prior $p(\epsilon_i^2) = (\epsilon_i^2)^{-1}\mathbb{I}_{\mathbb{R}^+}(\epsilon_i^2)$, with $\mathbb{I}_{\mathbb{R}^+}(\cdot)$ the indicator function of the set $\mathbb{R}^+$.

Posterior distribution. Assuming a priori independence between $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$, the Bayes' theorem yields the following posterior

$$p(\boldsymbol{\Theta}, \boldsymbol{\epsilon}|\mathbf{L}) \propto p(\mathbf{L}|\boldsymbol{\Theta}) \prod_{i=1}^{2} p(\boldsymbol{\theta}_i|\epsilon_i^2) p(\epsilon_i^2), \quad \boldsymbol{\epsilon} \triangleq [\epsilon_1^2, \epsilon_2^2]^T. \quad (11)$$

Bayesian estimators. In the context of MFA, only the parameters $\boldsymbol{\theta}_i$ are of interest. We therefore consider here the marginal posterior mean estimator associated with (11), denoted MMSE (minimum mean square error estimator) and defined by

$$\boldsymbol{\theta}_i^{\text{MMSE}} \triangleq \mathbb{E}[\boldsymbol{\theta}_i|\mathbf{L}] \quad (12)$$

where the expectation is taken with respect to the marginal posterior distribution $p(\boldsymbol{\theta}_i|\mathbf{L})$. Since (12) involves integrating over the posterior (11), its direct computation is intractable. Instead, the inference is performed by using a Gibbs sampler (GS) generating the collection of samples $\{\boldsymbol{\Theta}^{(k)}, \boldsymbol{\epsilon}^{(k)}\}_{k=1}^{N_{mc}}$ that are asymptotically distributed according to (11). These samples are used in turn to approximate the marginal posterior mean estimator by

$$\boldsymbol{\theta}_i^{\text{MMSE}} \approx (N_{mc} - N_{bi})^{-1} \sum_{k=N_{bi}+1}^{N_{mc}} \boldsymbol{\theta}_i^{(k)} \quad (13)$$

where N_{bi} is the length of the burn-in period [20].

3.2. Gibbs sampler

The GS considered here successively generates samples according to the conditional distributions associated with the posterior distribution $p(\boldsymbol{\Theta}, \boldsymbol{\epsilon}|\mathbf{L})$ [20].

Multifractal parameters. It can be shown that the conditional distribution $p(\boldsymbol{\theta}_i|\mathbf{L}, \boldsymbol{\theta}_{i'\neq i}, \boldsymbol{\epsilon})$ is not standard. Due to the high dimension of $\boldsymbol{\theta}_i$ ($M \gg 1$), implementing an MHG procedure (with a random walk proposal as in [13, 14]) would yield a poor exploration of the target distribution. Instead, we resort to a Hamiltonian Monte-Carlo algorithm (HMC) [21] whose strategy is recalled in the next subsection.

Hyperparameters. The conditional distributions for hyperparameters ϵ_i^2 are inverse-gamma ($\mathcal{IG}$) distributions that are easy to sample

$$\epsilon_i^2|\boldsymbol{\theta}_i \sim \mathcal{IG}\left(\frac{M-1}{2}, \frac{\|\boldsymbol{D}\boldsymbol{\theta}_i\|^2}{2}\right). \quad (14)$$

3.3. Hamiltonian Monte-Carlo

Hamiltonian system. The HMC algorithm is a sampling scheme inspired by Hamiltonian dynamics [21]. The target distribution, here $p(\boldsymbol{\theta}_i|\mathbf{L}, \boldsymbol{\theta}_{i'\neq i}, \boldsymbol{\epsilon})$, is associated with a potential energy $E(\boldsymbol{q}) = -\ln p(\boldsymbol{q}|\mathbf{L}, \boldsymbol{\theta}_{i'\neq i}, \boldsymbol{\epsilon})$ with $\boldsymbol{q} \triangleq \boldsymbol{\theta}_i$. Moreover $\boldsymbol{p} \in \mathbb{R}^M$ auxiliary momentum variables are introduced and associated with a kinetic energy $K(\boldsymbol{p}) = \boldsymbol{p}^T\boldsymbol{p}/2$. The Hamiltonian $H(\boldsymbol{q}, \boldsymbol{p}) = E(\boldsymbol{q}) + K(\boldsymbol{p})$ defines trajectories $(\boldsymbol{q}(\tau), \boldsymbol{p}(\tau))$, in continuous time τ, with constant total energy $H(\boldsymbol{q}, \boldsymbol{p})$ via the system of equations $\frac{d\boldsymbol{q}}{d\tau} = \frac{\partial H}{\partial \boldsymbol{p}}(\boldsymbol{q}, \boldsymbol{p})$, $\frac{d\boldsymbol{p}}{d\tau} = -\frac{\partial H}{\partial \boldsymbol{q}}(\boldsymbol{q}, \boldsymbol{p})$ [21].

Sampling. In an HMC sampling scheme, the proposal of a candidate is achieved through the discrete evaluation of the Hamiltonian equations. More precisely, at the iteration k of the GS, starting from the initial state $\boldsymbol{q}_0 \triangleq \boldsymbol{\theta}_i^{(k)}$ and $\boldsymbol{p}_0 \sim \mathcal{N}(\mathbf{0}_M, \boldsymbol{I}_M)$, the system of Hamiltonian equations is numerically integrated for a time interval of length $\delta \cdot L$ using the leap-frog method, which yields the candidate $(\boldsymbol{q}^\star, \boldsymbol{p}^\star)$. This candidate is then accepted with an acceptance rate $\alpha = \min(1, \rho)$ with $\rho = \exp[H(\boldsymbol{q}_0, \boldsymbol{p}_0) - H(\boldsymbol{q}^\star, \boldsymbol{p}^\star)]$. The leap-frog method, detailed in Algo. 1, is composed of L steps associated with an increment δ that is tuned during the N_{bi} first iterations of the GS such that $\alpha \in [0.5, 0.8]$. Note that this scheme requires the computation of the derivatives of the potential energy. For the model proposed in this work, the derivatives can be calculated analytically and are given by the closed-form expression

$$\frac{\partial E(\boldsymbol{\theta}_i)}{\partial \theta_i^t} = \left[\frac{D^T D \boldsymbol{\theta}_i}{\epsilon_i^2}\right]_t + \frac{1}{2}\left(\mathbf{f}_i \oslash \boldsymbol{\phi}(\boldsymbol{\theta}^t)\right)^T \left(\mathbf{1} - \mathbf{I}(\boldsymbol{\ell}^t) \oslash \boldsymbol{\phi}(\boldsymbol{\theta}^t)\right) \quad (15)$$

where $[\cdot]_t$ stands for the t-th element of a vector.

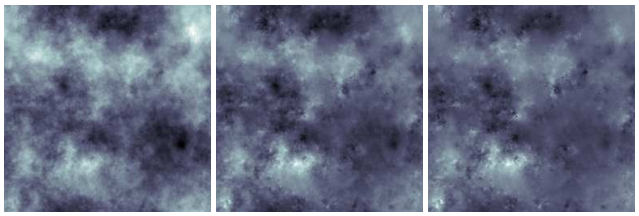
Algorithm 1 Sampling scheme via HMC

Set $\boldsymbol{q}_0 = \boldsymbol{\theta}_i^{(k)}$ and draw $\boldsymbol{p}_0 \sim \mathcal{N}(\mathbf{0}_M, \boldsymbol{I}_M)$
Leap-frog method
for $n = 0 : L-1$
  $\boldsymbol{p}_{n+1/2} = \boldsymbol{p}_n - \frac{\delta}{2}\frac{\partial E}{\partial \boldsymbol{q}^T}(\boldsymbol{q}_n)$
  $\boldsymbol{q}_{n+1} = \boldsymbol{q}_n + \delta \boldsymbol{p}_{n+1/2}$
  $\boldsymbol{p}_{n+1} = \boldsymbol{p}_{n+1/2} - \frac{\delta}{2}\frac{\partial E}{\partial \boldsymbol{q}^T}(\boldsymbol{q}_{n+1})$
end for
Set $(\boldsymbol{q}^\star, \boldsymbol{p}^\star) = (\boldsymbol{q}_L, \boldsymbol{p}_L)$ and draw $u \sim U_{[0,1]}$
Acceptance-reject
  Compute $\alpha = \min(1, \exp[H(\boldsymbol{q}_0, \boldsymbol{p}_0) - H(\boldsymbol{q}^\star, \boldsymbol{p}^\star)])$
  Set $\boldsymbol{\theta}_i^{(k+1)} = \boldsymbol{q}^\star$ if $u < \alpha$, otherwise $\boldsymbol{\theta}_i^{(k+1)} = \boldsymbol{\theta}_i^{(k)}$

4. NUMERICAL EXPERIMENTS

Scenario. We compare the proposed multivariate Bayesian approach (denoted as mB) to the univariate Bayesian method proposed in [13] (denoted as uB) and to the linear regression (3) (denoted as LF) by applying them to a large number of independent realizations of a sequence of $M = 100$ multi-temporal images defined as 2D multifractal random walks (MRWs) of size $2^7 \times 2^7$. An MRW has multifractal properties mimicking those of Mandelbrot's log-normal cascades, with scaling exponents $\zeta(q) = (H - c_2)q + c_2q^2$ [22]. Typical realizations of MRW images are plotted in Fig. 1 for three different values of c_2. We use a Daubechies wavelet with $N_\psi = 2$ vanishing moments and scales $(j_1, j_2) = (1, 3)$ in the analysis and set $H = 0.72$. Four different evolutions of c_2 across the image sequences are studied, with values for c_2 ranging from -0.04 to -0.12: a slow sinusoidal profile, a fast sinusoidal profile, a chirp profile including both slow and fast evolutions (as a limit benchmark case violating the slow evolution assumption) and a discontinuous evolution (cf., Fig. 2, top row, left to right columns, respectively).

Estimation performance. Estimation performance is quantified via the average, the standard deviation (STD) and the root mean squared error (RMSE) computed for MC $= 100$ independent series and defined by $m = \widehat{\mathbb{E}}[\hat{c}_2]$, $s = (\widehat{\text{Var}}[\hat{c}_2])^{\frac{1}{2}}$ and $rms = \sqrt{(m - c_2)^2 + s^2}$. Results are summarized in Fig. 2 (a) for the four different evolutions of c_2 and yield the following conclusions. All estimators succeed in reproducing on average the prescribed values of c_2 for the smooth evolutions (i.e., they have small bias). For the discontinuous evolution of c_2, mB provides a smooth estimate, as expected, and hence introduces a small bias in the vicinity of the discontinuities. More importantly, however, it can



$c_2 = -0.01$ $c_2 = -0.07$ $c_2 = -0.12$

Fig. 1. Realizations of 2D MRW with different values for c_2.

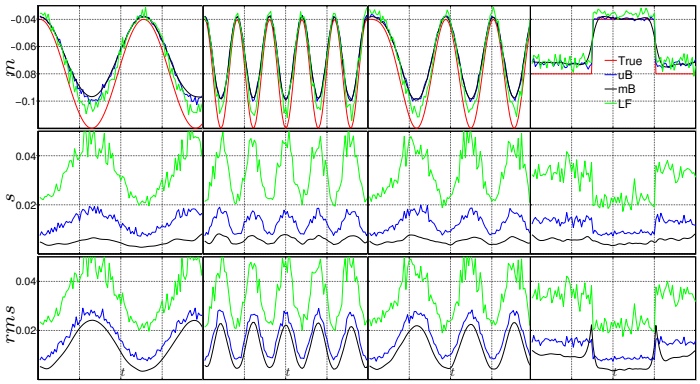
be observed that the proposed joint Bayesian estimator mB consistently yields a significant reduction of STD values as compared to LF (STD divided by up to 6.5) and also to the univariate Bayesian estimator uB (STD divided by up to 2.5). This demonstrates the clear benefits of the proposed Bayesian procedure for the joint estimation of c_2 and is also reflected in RMSE values which are, except at the locations of discontinuities, $50-70\%$ smaller than those of uB and LF for mB.

The observed improvements come at the price of increased yet reasonable computational time (for a sequence of 100 $2^7 \times 2^7$ images, $\sim$ 8s for mB instead of $\sim$ 1s for LF or $\sim$ 3s for uB).

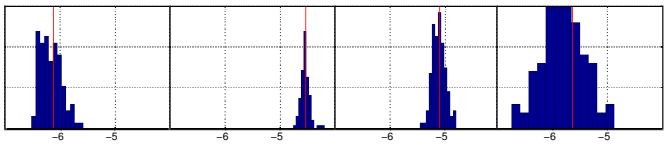
Estimation of hyperparameters ϵ_i^2. In order to study the effectiveness of the automatic tuning of the hyperparameters ϵ_1^2, Fig. 2 (b) reports histograms of MMSE estimates $\epsilon_1^{2\text{MMSE}} \approx (N_{mc} - N_{bi})^{-1} \sum_{k=N_{bi}+1}^{N_{mc}} \epsilon_1^{2(k)}$ for the four different evolutions of c_2. Note that since ϵ_1^2 corresponds to the variance in the prior (10), the smaller ϵ_1^2, the smoother the solution. For the sinusoidal evolutions, the average value of ϵ_1^2 (indicated by a thin red line in Fig. 2 (b)) is $10^{-6.2}$ and $10^{-4.5}$ for slow and fast evolution, respectively, thus reflecting the degree of smoothness in the evolution. For the chirp, the average value is 10^{-5} and thus sightly below that of the fast sinusoid (indeed, stronger smoothing would introduce bias and would hence be highly penalizing). The values of ϵ_1^2 for the discontinuous case are centered at $10^{-5.8}$ and close to the slow sinusoid case, indicating that larger bias at the two discontinuities is traded off for small variability within the segments with constant c_2. These results clearly demonstrate that the model succeeds in adjusting the hyperparameter to an appropriate smoothing level for the data.

5. CONCLUSIONS AND PERSPECTIVES

This work introduced a Bayesian procedure that enables, for the first time, the joint estimation of the multifractality parameter c_2 for sequences of multi-temporal images. Building on a recent statistical model for the multivariate statistics of log-leaders of MMC based processes, the procedure relies on two mains contributions. First, a smoothness assumption for parameter values across the collection of images was encoded via a Gaussian prior on the second derivatives of the multifractal attributes. The resulting Bayesian model also enables the estimation of the hyperparameters controlling the amount of smoothness. Second, to bypass the difficulties resulting from non-standard conditional distributions during the Bayesian inference, a Hamiltonian Monte-Carlo scheme was proposed. Closed-form expressions for the derivatives required in this scheme were obtained. Our numerical results demonstrate that the proposed procedure yields excellent estimation performance and significantly improves over previous (univariate) formulations. Future work will include the extension of the proposed approach to spatial multivariate data, i.e., the joint estimation of parameters for patches of heterogeneous multifractal images.



(a) Estimation performance



(b) Histograms $\log_{10} \epsilon_1^2$

Fig. 2. Numerical experiments: (a) estimation performance (from top to bottom: mean, STD and RMSE) assessed on 100 independent realizations for different evolutions of c_2; (b) histograms of MMSE estimates $\epsilon_1^{2\text{MMSE}}$.

6. REFERENCES

[1] 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. 2004, vol. 72(2), pp. 91–152, AMS.

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

[3] J. F. Muzy, E. Bacry, and A. Arneodo, "The multifractal formalism revisited with wavelets," *Int. J. of Bifurcation and Chaos*, vol. 4, pp. 245–302, 1994.

[4] H. Wendt, S. G. Roux, S. Jaffard, and P. Abry, "Wavelet leaders and bootstrap for multifractal analysis of images," *Signal Proces.*, vol. 89, no. 6, pp. 1100 – 1114, 2009.

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

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

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

[8] H. Wendt, S. Jaffard, and P. Abry, "Multifractal analysis of self-similar processes," in *Proc. IEEE Workshop Statistical Signal Proces. (SSP)*, Ann Arbor, USA, 2012.

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

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

[11] P. Abry, R. Baraniuk, P. Flandrin, R. Riedi, and D. Veitch, "Multiscale nature of network traffic," *IEEE Signal Proces. Mag.*, vol. 19, no. 3, pp. 28–46, 2002.

[12] T. Lux, "Higher dimensional multifractal processes: A GMM approach," *J. Business Econ. Stat.*, vol. 26, pp. 194–210, 2007.

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

[14] S. Combrexelle, H. Wendt, J.-Y. Tourneret, P. Abry, and S. McLaughlin, "Bayesian estimation of the multifractality parameter for images via a closed-form Whittle likelihood," in *Proc. 23rd European Signal Proces. Conf. (EUSIPCO)*, Nice, France, 2015.

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

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

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

[18] V. V. Anh and K. E. Lunney, "Parameter estimation of random fields with long-range dependence," *Math. Comput. Model.*, vol. 21, no. 9, pp. 67–77, 1995.

[19] P. Campisi and K. Egiazarian, *Blind image deconvolution: theory and applications*, CRC press, 2007.

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

[21] R. M. Neal, "MCMC using Hamiltonian dynamics," *Handbook of Markov Chain Monte Carlo*, vol. 54, pp. 113–162, 2010.

[22] R. Robert and V. Vargas, "Gaussian multiplicative chaos revisited," *Ann. Proba.*, vol. 38, no. 2, pp. 605–631, 2010.