BAYESIAN ESTIMATION OF THE MULTIFRACTALITY PARAMETER FOR IMAGES VIA A CLOSED-FORM WHITTLE LIKELIHOOD

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 is central in many image processing problems. It can be conducted by studying the local regularity fluctuations of image amplitudes, and multifractal analysis provides a theoretical and practical framework for such a characterization. Yet, due to the non Gaussian nature and intricate dependence structure of multifractal models, accurate parameter estimation is challenging: standard estimators yield modest performance, and alternative (semi-)parametric estimators exhibit prohibitive computational cost for large images. This present contribution addresses these difficulties and proposes a Bayesian procedure for the estimation of the multifractality parameter c_2 for images. It relies on a recently proposed semi-parametric model for the multivariate statistics of log-wavelet leaders and on a Whittle approximation that enables its numerical evaluation. The key result is a closed-form expression for the Whittle likelihood. Numerical simulations indicate the excellent performance of the method, significantly improving estimation performance over standard estimators and computational efficiency over previously proposed Bayesian estimators.

***Index Terms*—** Multifractal analysis, Bayesian estimation, Hankel transform, Whittle likelihood, Texture analysis

1. INTRODUCTION

Context. Texture constitutes one of the central features in images, and its characterization plays an important role in a variety of image processing applications. Many different mathematical models for texture have been developped. It has been recognized that texture characterization can be efficiently conducted within the mathematical framework of multifractal analysis, which provides a standard signal and image processing tool that has been used in a large variety of applications, cf., e.g., [1, 2] and references therein. Multifractal analysis is a specific instance of scale invariance analysis which enables the fluctuations of the pointwise smoothness of image amplitudes to be studied: The texture of an image $X(\boldsymbol{k})$ is characterized by means of the so-called *multifractal spectrum* $D(h)$ which is defined as the Hausdorff dimension of the sets of points that have the same pointwise regularity h (commonly measured with the so-called Hölder exponent, cf. [3]). In practice, multifractal models translate into the power law behaviors of the sample moments of appropriate multiresolution quantities $T_X(a, \boldsymbol{k})$ (i.e., quantities that depend jointly on scale $a = 2^j$ and spatial position $\boldsymbol{k}$) of $X(\boldsymbol{k})$,

$$S(q, j) \equiv \frac{1}{n_j} \sum_{k=1}^{n_j} |T_X(j, k)|^q \simeq a^{\zeta(q)},\ a_m \leq a \leq a_M. \quad (1)$$

Here, wavelet leaders $\ell(j, \boldsymbol{k})$ will be used as multiresolution quantities $T_X(a, \boldsymbol{k})$, which are considered to be the benchmark multiresolution quantities for multifractal analysis and defined in Section 2.1 below [1, 3]. The exponents $\zeta(q)$ of the power laws in (1), termed *scaling exponents*, are intimately tied to the local regularity fluctuations of the image amplitudes, measured by $D(h)$, via a Legendre transform, $D(h) \leq \mathcal{L}(h) := \inf_{q\in\mathbb{R}}[1+qh-\zeta(q)]$. Notably, they enable the formal discrimination between the two most prominent classes of scale invariance models: self-similar processes, for which $\zeta(q)$ is a linear function of q in the neighborhood of $q = 0$ [4], and multifractal multiplicative cascade based processes, for which $\zeta(q)$ is a strictly concave function [5]. The decision whether a class or the other better models real-world data is fundamental in applications because they imply completely different data production mechanisms: additive for the former, and multiplicative for the latter. In practice, it can be cast into testing the linearity of $\zeta(q)$ at $q = 0$ [1, 6] by considering the development of $\zeta(q)$ as a polynomial at $q = 0$, $\zeta(q) = \sum_{m\geq 1} c_m q^m/m!$. One can show that $c_2 < 0$ for multiplicative cascades while $c_2 \equiv 0$ for self-similar processes (cf., e.g, [6]) and $c_2 \equiv 0$ implies that $c_m \equiv 0,\ \forall m \geq 3$ [3]. The estimation of c_2, termed the intermittency or *multifractality* parameter, is therefore central in multifractal analysis.

Estimation of c_2. In the seminal contribution [7], it was shown that the coefficients c_m are related to the cumulants of the logarithm of the multiresolution quantities (here, the wavelet leaders $\ell_X(j, \boldsymbol{k})$). In particular, the variance of the log of wavelet leaders is given by the expression

$$C_2(j) \equiv \text{ Var } [\ln \ell_X(j, \boldsymbol{k})] = c_2^0 + c_2 \ln 2^j. \quad (2)$$

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.

978-0-9928626-3-3/15/$31.00 ©2015 IEEE

The parameter c_2 can thus be estimated by linear regression of the sample variance $\widehat{\text{Var}}$ of the log-leaders with respect to scale j

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

where w_j are suitable regression weights. The main limitation of (3) is that it requires a sufficient number of scales to be available and hence sufficiently large images (in practice, images of size $N \times N \gtrsim 512 \times 512$) to yield satisfactory performance. As an alternative to (3), a generalized method of moments has been proposed [8], relying on fully parametric models that are often too restrictive in applications.

To overcome such limitations, it has recently been proposed to conduct the estimation of c_2 in a Bayesian framework [2,9]. The method relies on the use of a semi-parametric Gaussian likelihood model for the multivariate statistics of the log-leaders. The model is generically valid for multifractal multiplicative cascade processes and imposes minimal assumptions (essentially, (2)) on data. To enable the evaluation of the likelihood for images, a Whittle approximation was used in [2] that expresses the likelihood in the spectral domain. The Bayesian estimators associated with the model were then approximated using a Markov chain Monte Carlo (MCMC) algorithm. The method is robust and significantly improves estimation performance when compared to (3). Yet, the variance-covariance model does not lead to a closed-form expression of the Whittle likelihood and its successive evaluations in the MCMC algorithm are thus costly, practically prohibiting its application to very large images.

Goals and contributions. This present contribution proposes a Bayesian estimation procedure for c_2 for images that is effective and efficient for both small and large images. The proposed procedure extends the work presented in [2] by developing an original formulation of the multivariate Gaussian model leading to a more efficient algorithm.

The key contribution resides in the derivation of a closed-form expression for the spectral density associated with the multivariate Gaussian model for log-leaders. This contribution is specific for the analysis of images. First, the model is expressed in continuous time and its radial symmetry is exploited for expressing the associated spectral density as a Hankel transform [10]. Second, by evaluation of the Hankel transform, we obtain closed-form expressions for the radial component of the spectral density model in which the parameters are made explicit. Finally, this closed-form model is discretized and substituted in the Whittle approximation.

The performance of the proposed Bayesian estimation procedure for the parameter c_2 of images are assessed with Monte Carlo simulations for synthetic multiplicative cascade based multifractal processes, showing the clear benefits of the proposed method: it strongly outperforms (3) in estimation performance and significantly reduces computation time when compared to the solution of [2].

2. BAYESIAN MODEL AND ESTIMATION

2.1. Wavelet coefficients and leaders.

Let $\phi(x)$ and $\psi(x)$ denote the scaling function and mother wavelet defining a 1D multiresolution analysis [11]. 2D wavelets can be defined as: $\psi^{(0)}(\mathbf{x}) = \phi(x_1)\phi(x_2)$, $\psi^{(1)}(\mathbf{x}) = \psi(x_1)\phi(x_2)$, $\psi^{(2)}(\mathbf{x}) = \phi(x_1)\psi(x_2)$, $\psi^{(3)}(\mathbf{x}) = \psi(x_1)\psi(x_2)$. The collections $\psi^{(m)}_{j,\boldsymbol{k}}(\mathbf{x}) = 2^{-j}\psi^{(m)}(2^{-j}\mathbf{x} - \boldsymbol{k})$ of templates of ψ_0, dilated to scales $a = 2^j$ and translated to space positions $\mathbf{x} = 2^j\boldsymbol{k}$, form a basis of $L^2(\mathbb{R}^2)$ for a well chosen ψ. The (L^1-normalized) discrete wavelet transform coefficients of the image X are defined as $d^{(m)}_X(j,k) = \langle X, \psi^{(m)}_{j,k} \rangle$, $m = 0, \ldots, 3$ [11]. Let $\lambda_{j,\boldsymbol{k}}$ denote the dyadic cube of side length 2^j centered at $\boldsymbol{k}2^j$ and $3\lambda_{j,\boldsymbol{k}} = \bigcup_{n_1,n_2=\{-1,0,1\}} \lambda_{j,k_1+n_1,k_2+n_2}$ the union of this cube with its eight neighbors. The wavelet leaders are defined as the supremum of the wavelet coefficients in this spatial neighborhood over all finer scales [1, 3]

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

2.2. Statistics of log-wavelet leaders for images

Extensive numerical simulations in [2, 9] have shown that the statistics of the log-leaders, $l(j,\cdot) \triangleq \ln \boldsymbol{\ell}(j,\cdot)$, of multiplicative cascade based multifractal processes can be well approximated by multivariate Gaussian distributions. In [2], a model for the covariance $\mathcal{C}_j(\boldsymbol{k}, \boldsymbol{\Delta k}) \triangleq \text{Cov}[l(j,\boldsymbol{k}), l(j,\boldsymbol{k}+\boldsymbol{\Delta k})]$ at scale j for images has been proposed. It is inspired by the asymptotic covariance properties that are induced by the multiplicative cascade construction [5], can be parametrized by the parameter vector $\boldsymbol{\theta} \triangleq [c_2, c_2^0]^T$ and is given by

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

where $|\cdot|$ is the Euclidian norm. The model (5) is radial symmetric, with radial component defined by the functions ϱ^0_j and ϱ^1_j. These functions are linear functions of $\ln(r+1)$ and $\ln r$, respectively, and are defined as

$$\varrho^0_j(r; \boldsymbol{\theta}) \triangleq a_j \ln(1+r) + c_2^0 + c_2 \ln 2^j \tag{6}$$

where $a_j \triangleq (\varrho^1_j(3; \boldsymbol{\theta}) - c_2^0 - c_2 \ln 2^j)/\ln 4$ and

$$\varrho^1_j(r; \boldsymbol{\theta}) \triangleq c_2 \ln(r/r_j) \mathbb{I}_{[0,r_j]}(r) \tag{7}$$

where $r_j = \lfloor \sqrt{n_j}/4 \rfloor$, $\lfloor \cdot \rfloor$ truncates to integer values, $n_j \approx \lfloor N^2/2^{2j} \rfloor$ is the number of wavelet leaders at scale j and $\mathbb{I}_A(r)$ is the indicator function of the set A. Note that $\varrho^0_j(r = 0; \boldsymbol{\theta})$ reduces to the variance given in (2).

2.3. Likelihood, prior and posterior distributions

We consider the estimation of c_2 here and therefore do not consider the mean of the log-leaders in the model. The log-leaders at scale j are centered, $\bar{l}(j, \boldsymbol{k}) \triangleq l(j, \boldsymbol{k}) - \widehat{\mathbb{E}}[l_X(j,.)]$,

where $\widehat{\mathbb{E}}[\cdot]$ is the sample mean, and stacked in the vector $\boldsymbol{\ell}_j$ according to the lexicographic order $\boldsymbol{k}_i$, $i = 1, \cdots, n_j^2$.

Likelihood. For a given scale j, the statistical model and notations above straightforwardly lead to the likelihood of $\boldsymbol{\ell}_j$

$$p(\boldsymbol{\ell}_j|\boldsymbol{\theta}) \triangleq \left((2\pi)^{n_j} \det \boldsymbol{\Sigma}_j(\boldsymbol{\theta})\right)^{-\frac{1}{2}} \exp\left(-\frac{1}{2}\boldsymbol{\ell}_j^T \boldsymbol{\Sigma}_j^{-1}(\boldsymbol{\theta})\boldsymbol{\ell}_j\right) \quad (8)$$

where $\boldsymbol{\Sigma}_j(\boldsymbol{\theta})$ is the covariance matrix with elements given by $[\boldsymbol{\Sigma}_j(\boldsymbol{\theta})]_{u,v} = \varrho_j(|\boldsymbol{k}_u - \boldsymbol{k}_v|; \boldsymbol{\theta})$. The log-leaders at different scales are assumed independent. The likelihood for all centered log-leaders, $\mathbf{L} = [\boldsymbol{\ell}_{j_1}^T, ..., \boldsymbol{\ell}_{j_2}]^T$, is thus given by

$$p(\mathbf{L}|\boldsymbol{\theta}) = \prod_{j=j_1}^{j_2} p(\boldsymbol{\ell}_j|\boldsymbol{\theta}). \quad (9)$$

Prior distribution. The parameter vector $\boldsymbol{\theta}$ must be chosen such that the variances $C_2(j)$ are positive for $j = j_1, \ldots, j_2$, which is ensured if $\boldsymbol{\theta}$ belongs to the admissible set $\mathcal{A} = (\mathcal{A}^+ \cup \mathcal{A}^-) \cap \mathcal{A}^M$ with $\mathcal{A}^- = \{(c_2, c_2^0) \in \mathbb{R}^2 \mid c_2 < 0 \text{ and } c_2^0 + c_2\, j_2 \ln 2 > 0\}$, $\mathcal{A}^+ = \{(c_2, c_2^0) \in \mathbb{R}^2 \mid c_2 > 0 \text{ and } c_2^0 + c_2\, j_1 \ln 2 > 0\}$ and $\mathcal{A}^m = \{(c_2, c_2^0) \in \mathbb{R}^2 \mid |c_2^0| < c_2^{0,M}, |c_2| < c_2^M\}$, where c_2^M and $c_2^{0,M}$ are the largest admissible values for c_2 and c_2^0, respectively. When no additional prior information is available, a uniform prior distribution on the set $\mathcal{A}$ is assigned to $\boldsymbol{\theta}$, i.e., $\pi(\boldsymbol{\theta}) = U_{\mathcal{A}}(\boldsymbol{\theta}) \propto \mathbb{I}_{\mathcal{A}}(\boldsymbol{\theta})$.

Posterior distribution and Bayesian estimators. The posterior distribution of $\boldsymbol{\theta}$ follows from Bayes rule

$$p(\boldsymbol{\theta}|\mathbf{L}) \propto p(\mathbf{L}|\boldsymbol{\theta})\, \pi(\boldsymbol{\theta}). \quad (10)$$

It is used to define the minimum mean squared error (MMSE) and maximum a posteriori (MAP) estimators in (11).

2.4. Gibbs sampler

Since the Bayesian estimators associated with (10) are difficult to compute, we investigate a Gibbs sampling strategy to generate samples $\{\boldsymbol{\theta}^{(t)}\}_{t=1}^{N_{mc}}$ that are asymptotically distributed according to the posterior distribution (10). It relies on successive sampling according to the conditional distributions associated with $p(\boldsymbol{\theta}|\mathbf{L})$. Since these conditional distributions do not correspond with standard laws, they are sampled using a Metropolis-within-Gibbs procedure, defined by random walks with Gaussian proposal distributions. More precisely, given the state $\boldsymbol{\theta}^{(t-1)}$, two steps are computed.

Sampling according to $p(c_2^{(t)}|c_2^{0,(t-1)}, \mathbf{L})$. A candidate c_2^* is drawn from the proposal distribution $p_{c_2}(c_2^*|c_2^{(t-1)}) = \mathcal{N}(c_2^{(t-1)}, \sigma_{c_2}^2)$. It is accepted ($c_2^{(t)} = c_2^*$) or rejected ($c_2^{(t)} = c_2^{(t-1)}$) with the Metropolis-Hastings ratio r_{c_2}.

Sampling according to $p(c_2^{0,(t)}|c_2^{(t)}, \mathbf{L})$. A candidate $c_2^{0,*}$ is drawn from the proposal distribution $p_{c_2^0}(c_2^{0,*}|c_2^{0,(t-1)}) = \mathcal{N}(c_2^{0,(t-1)}, \sigma_{c_2^0}^2)$ and accepted ($c_2^{0,(t)} = c_2^{0,*}$) or rejected ($c_2^{0,(t)} = c_2^{0,(t-1)}$) with the Metropolis-Hastings ratio $r_{c_2^0}$.

The Metropolis-Hastings acceptance ratios are defined by $r_\theta = \frac{p(\boldsymbol{\theta}^*|\mathbf{L})}{p(\boldsymbol{\theta}^{(t-1)}|\mathbf{L})} \frac{p_\theta(\theta^{(t-1)}|\theta^*)}{p_\theta(\theta^*|\theta^{(t-1)})}$. The variances $\sigma_{(\cdot)}^2$ of the proposal distributions are chosen to yield acceptance ratios $r_\theta \in [0.4, 0.6]$ (see [12] for details on MCMC methods).

Bayesian estimators. The generated samples (after a burn-in of N_{bi} samples) are used to approximate the Bayesian estimators of $\boldsymbol{\theta}$ as follows

$$\hat{\boldsymbol{\theta}}^{\text{MMSE}} \approx \frac{1}{N_{mc} - N_{bi}} \sum_{N_{bi}}^{N_{mc}} \boldsymbol{\theta}^{(t)},\ \hat{\boldsymbol{\theta}}^{\text{MAP}} \approx \underset{t \geq N_{bi}}{\operatorname{argmax}}\, p(\boldsymbol{\theta}^{(t)}|\mathbf{L}). \quad (11)$$

3. FAST COMPUTATION USING A WHITTLE APPROXIMATION AND A HANKEL TRANSFORM

Whittle likelihood. The inversion of $\boldsymbol{\Sigma}_j(\boldsymbol{\theta})$ in (8) at each iteration of the Gibbs sampler is computationally intensive and numerically problematic (large condition number) even for small images. Therefore, it was proposed in [2] to replace the exact likelihood (8) by a Whittle approximation [13] whose evaluation in the spectral domain is efficient and numerically robust. Under the assumptions of Section 2.2, the Whittle approximation of (8) is given by [13, 14]

$$p(\boldsymbol{\ell}_j|\boldsymbol{\theta}) \propto p^\dagger(\boldsymbol{\ell}_j|\boldsymbol{\theta}) \triangleq \exp -\frac{1}{2} \sum_{\mathbf{m} \in J_j} \ln \phi_j(\boldsymbol{\omega_m}; \boldsymbol{\theta}) + \frac{I_j(\boldsymbol{\omega_m})}{\phi_j(\boldsymbol{\omega_m}; \boldsymbol{\theta})} \quad (12)$$

where $I_j(\boldsymbol{\omega_m}) \triangleq |\sum_{\boldsymbol{k}} \bar{l}(j, \boldsymbol{k}) \exp(-\mathrm{i}\boldsymbol{k}^T \boldsymbol{\omega_m})|^2 / n_j$ is the periodogram of $\{\bar{l}(j, \boldsymbol{k})\}$, $\boldsymbol{\omega_m} = 2\pi \mathbf{m}/n_j$ and $\mathbf{m} \in J_j \triangleq [\![\lfloor(-\sqrt{n_j}-1)/2\rfloor : \sqrt{n_j} - \lfloor\sqrt{n_j}/2\rfloor]\!]^2$. The function $\phi_j(\boldsymbol{\omega_m}; \boldsymbol{\theta})$ is the parametric spectral density associated with the covariance model (5). Finally, the approximation $p^\dagger(\mathbf{L}|\boldsymbol{\theta}) \triangleq \prod_{j=j_1}^{j_2} p^\dagger(\boldsymbol{\ell}_j|\boldsymbol{\theta})$ replaces (9) in (10).

Numerical spectral density model. In [2], it was proposed to compute the spectral density $\phi_j(\boldsymbol{\omega_m}; \boldsymbol{\theta})$ in (12) numerically by using the discrete Fourier transform (DFT)

$$\phi_j^{DFT}(\boldsymbol{\omega_m}; \boldsymbol{\theta}) = \Big| \sum_{\boldsymbol{\Delta k}} \varrho_j(|\boldsymbol{\Delta k}|; \boldsymbol{\theta}) \exp(-\mathrm{i}\boldsymbol{\Delta k}^T \boldsymbol{\omega_m}) \Big|. \quad (13)$$

Practically, this means that at each step of the Gibbs sampler a 2D DFT must be computed.

Closed-form parametric spectral density. In this work, we derive a closed-form parametric expression for $\phi_j(\boldsymbol{\omega_m}; \boldsymbol{\theta})$ that avoids the costly evaluations of 2D DFTs in (12) implied by (13). The continuous spectral density associated with the covariance model $\varrho_j(r; \boldsymbol{\theta})$ is given by Bochner's theorem

$$\tilde{\phi}_j(\boldsymbol{\omega}; \boldsymbol{\theta}) = \int_{\mathbb{R}^2} e^{-\mathrm{i}(\mathbf{x}^T \boldsymbol{\omega})} \varrho_j(|\mathbf{x}|; \boldsymbol{\theta})\, d\mathbf{x}. \quad (14)$$

Note that because $\varrho_j(|\mathbf{x}|; \boldsymbol{\theta})$ is a radial symmetric function, its Fourier transform $\tilde{\phi}_j(\boldsymbol{\omega}; \boldsymbol{\theta})$ is also radial symmetric. It can therefore be expressed as a Hankel transform [10], given by

$$\tilde{\phi}_j(\boldsymbol{\omega}; \boldsymbol{\theta}) = \tilde{\phi}_j^H(|\boldsymbol{\omega}|; \boldsymbol{\theta}) = 2\pi \int_0^\infty r \varrho_j(r; \boldsymbol{\theta}) J_0(r|\boldsymbol{\omega}|)\, dr \quad (15)$$

where $J_n(\cdot)$ is the n-th order Bessel function. To evaluate the integral (15), we make use of the identities: (i) $\rho \int_0^R r J_0(r\rho) dr = R J_1(R\rho)$ and (ii) $\rho^2 \int_0^R r \ln(r/R) J_0(r\rho) dr$

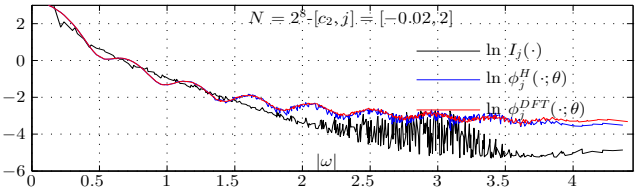

Fig. 1. Radial evolution of the periodogram of MRW (black) and of models (13) and (18) (red and blue, respectively).

$= -(1 - J_0(R\rho))$ which are valid for $R > 0$, $\rho > 0$, cf. [10, Tab. 17.1]. The function $\varrho_j^1(r;\boldsymbol{\theta})$ is affine in $\ln(r)$ as in (ii), and we thus break up the integration according to the range of validity of ϱ_j^0 and ϱ_j^1 in (5), i.e., $\tilde{\phi}_j^H(|\boldsymbol{\omega}|;\boldsymbol{\theta}) = A_j(|\boldsymbol{\omega}|)+B_j(|\boldsymbol{\omega}|)$ with $A_j(|\boldsymbol{\omega}|)=2\pi\int_0^3 r\varrho_j^0(r;\boldsymbol{\theta})J_0(r|\boldsymbol{\omega}|)dr$ and $B_j(|\boldsymbol{\omega}|) = 2\pi\int_3^\infty r\varrho_j^1(r;\boldsymbol{\theta})J_0(r|\boldsymbol{\omega}|)\,dr$. This yields

$$\frac{A_j(|\boldsymbol{\omega}|)}{2\pi} = a_j\frac{\ln 4}{2\pi}\mathcal{I}(|\boldsymbol{\omega}|)+(c_2^0+c_2\ln 2^j)\frac{3J_1(3|\boldsymbol{\omega}|)}{|\boldsymbol{\omega}|}$$

$$\frac{B_j(|\boldsymbol{\omega}|)}{2\pi c_2} = \tilde{B}_j(|\boldsymbol{\omega}|)=\frac{J_0(r_j|\boldsymbol{\omega}|)-J_0(3|\boldsymbol{\omega}|)}{|\boldsymbol{\omega}|^2}+\frac{3\ln(\frac{r_j}{3})J_1(3|\boldsymbol{\omega}|)}{|\boldsymbol{\omega}|}$$

where $\mathcal{I}(|\boldsymbol{\omega}|) \triangleq 2\pi\int_0^3 r\ln(1+r)J_0(r|\boldsymbol{\omega}|)\,dr/\ln 4$. Grouping terms in c_2 and c_2^0 leads to the expression

$$\tilde{\phi}_j^H(|\boldsymbol{\omega}|;\boldsymbol{\theta}) = c_2\, f_j(|\boldsymbol{\omega}|) + c_2^0\, g_j(|\boldsymbol{\omega}|) \qquad (16)$$

for (15), where

$$f_j(|\boldsymbol{\omega}|) \triangleq 2\pi\tilde{B}_j(|\boldsymbol{\omega}|) + 3\ln 2^j\frac{J_1(3|\boldsymbol{\omega}|)}{|\boldsymbol{\omega}|} - \mathcal{I}(|\boldsymbol{\omega}|)\ln(\frac{r_j 2^j}{3})$$

$$g_j(|\boldsymbol{\omega}|) \triangleq 6\pi\frac{J_1(3|\boldsymbol{\omega}|)}{|\boldsymbol{\omega}|} - \mathcal{I}(|\boldsymbol{\omega}|).$$

At last, (16) is discretized using spectral aliasing [14]

$$\phi_j^H(\boldsymbol{\omega_m};\boldsymbol{\theta}) \triangleq \sum_{\mathbf{p}\in\mathbb{Z}^2}\tilde{\phi}_j^H(|\boldsymbol{\omega_m}+2\pi\mathbf{p}|;\boldsymbol{\theta}). \qquad (17)$$

Computation of $\phi_j^H(\boldsymbol{\omega_m};\boldsymbol{\theta})$. The infinite summation in (17) is truncated to $\mathbf{p}\in[\![-K:K]\!]^2$, yielding the approximation

$$\phi_j^H(\boldsymbol{\omega_m};\boldsymbol{\theta})\approx c_2\sum_{\mathbf{p}\in[\![-K:K]\!]^2} f_j(|\boldsymbol{\omega_{m,p}}|)+c_2^0\sum_{\mathbf{p}\in[\![-K:K]\!]^2} g_j(|\boldsymbol{\omega_{m,p}}|) \qquad (18)$$

where $\boldsymbol{\omega_{m,p}} \triangleq \boldsymbol{\omega_m}+2\pi\mathbf{p}$. Note that the two functions f_j and g_j do not depend on the parameters c_2 and c_2^0. The two partial sums can thus be pre-calculated and stored for the discrete set of frequencies $\boldsymbol{\omega_m}$, using a quadrature rule for the computation of the integral $\mathcal{I}$. The evaluation of (12) using (18) in the Gibbs sampler then only requires updating with the parameter candidates (c_2, c_2^0) in (18) at each iteration.

4. NUMERICAL EXPERIMENTS

The proposed Bayesian estimator (using (18) in (12), denoted H) is compared to the Bayesian estimator in [2] (using (13) in (12), denoted DFT) and to the standard linear regression based estimator (3) (denoted LF) by applying them to a large number of independent realizations of 2D multifractal random walk (MRW) [15]. MRW is a non Gaussian process whose multifractal properties mimic those of the Mandelbrot's multiplicative log-normal cascades and its scaling exponents are given by $\zeta(q) = (\mathcal{H} - c_2)q + c_2q^2$ (the reader is referred to [15] for precise definitions and details).

Experimental setup. The parameters of MRW are set to $\mathcal{H} = 0.72$ and $c_2 \in \{-0.01, -0.02, \ldots, -0.1\}$. For the 2D DWT, a Daubechies's mother wavelet with $N_\psi = 2$ was used. The weights w_j in (3) are chosen proportional to n_j (see, e.g., [1]). The summation in (12) is restricted to low frequencies $|\boldsymbol{\omega_m}| \leq \pi\sqrt{\eta}$ with $\eta = 0.25$ as in [2] and $K = 3$ in (18). The parameters of the Gibbs sampler are set to $N_{mc} = 2000$ and $N_{bi} = 1000$. Performance are quantified using the sample mean $\mathrm{m}=\widehat{\mathbb{E}}[\hat{c}_2]$, standard deviation $\mathrm{s}=(\widehat{\mathrm{Var}}[\hat{c}_2])^{\frac{1}{2}}$ (STD) and root mean squared error $\mathrm{rms}=\sqrt{(m-c_2)^2+s^2}$ (RMSE) of the estimates for 100 independent realizations of MRW.

Spectral density models. The radial components of the spectral density models (13) and (18) are plotted in Fig. 1 (blue and red lines, respectively) and compared to those of the periodogram of MRW (black curve, mean over 100 realizations). As expected, the DFT based expression (13) and the closed-form expression (18) proposed in this work are numerically close. Both yield good fits for the periodogram for low frequencies. The benefits of the proposed model (18) will become clear in the next paragraph.

Computational complexity. Fig. 2 investigates the computational time T for the estimators LF, DFT and H for different sample sizes N. The computational time includes the 2D DWT for all methods. The pre-calculation of partial sums in (18) is performed offline and not taken into account in T. The estimator LF unsurprisingly exhibits the lowest computational cost. Among the Bayesian estimators, the proposed estimator H yields a significant reduction of the computational time as compared to estimator DFT of [2], by a factor ranging from 6 (small images) to 12 (large images). Neglecting the computation of the 2D DWT and the 2D periodogram I_j in (12), the order of the reduction factor can be estimated as $\sum_{j=j_1}^{j_2} n_j\ln(n_j)/\sum_{j=j_1}^{j_2} n_j$ since computing $\phi_j(\boldsymbol{\omega_m};\boldsymbol{\theta})$ is of complexity $O(n_j\ln(n_j))$ when using (13) and $O(n_j)$ when using the proposed expression (18). As a result, LF is only 30 times faster than H but 400 times faster than DFT for $N=2^{11}$.

Estimation of c_2. Fig. 3 reports estimation performance as a function of c_2 for sample sizes $N = 2^7$ (top) and $N = 2^8$ (bottom). The performance for the MAP estimators are similar to the MMSE estimators and not reported for space reasons. Clearly, the Bayesian estimators H and DFT both yield excellent estimates for c_2. They significantly improve estimation quality as compared to LF, both in terms of bias and standard deviation. As a result, RMSE values are reduced by a factor up to 4. The bias and RMSE reduction yielded by H and DFT is particularly pronounced for small images. It is interesting to note that STD approaches 0 faster with decreasing $|c_2|$ for the Bayesian approach than for LF, resulting in larger STD gains for H and DFT for small values of $|c_2|$.

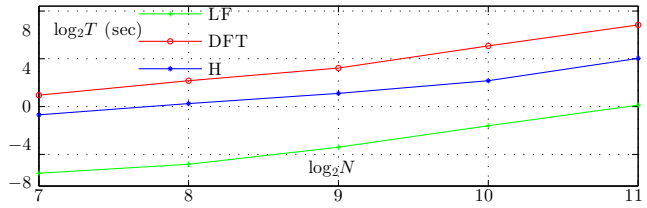

Fig. 2. Computational time T versus $\log_2$ of sample size N with $j_1 = 2$ and $j_2 = \log_2 N - 4$ for all methods.

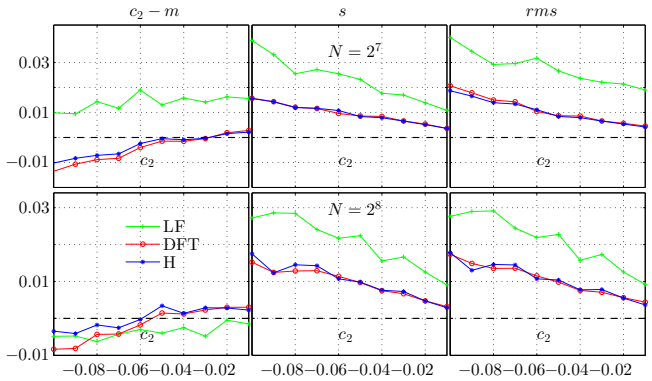

Fig. 3. Estimation performance for c_2 for sample size $N = \{2^7, 2^8\}$ using scales $(j_1, j_2) = \{(1, 3), (2, 4)\}$ (top and bottom plots, respectively): bias, STD, RMSE (from left to right)

5. CONCLUSIONS AND PERSPECTIVES

A Bayesian approach for the estimation of the multifractality parameter c_2 for images was proposed. It relies on a semi-parametric model for the multivariate statistics of log-leaders of multiplicative cascade based multifractal processes. An MCMC algorithm was devised for approximating the Bayesian estimators associated with the posterior distribution. The computation of the estimators was enabled through the use of a Whittle approximation in the Bayesian model. The key contribution resides in the derivation of a closed-form Whittle approximation via a Hankel transform, which significantly reduces computational cost. The proposed method realizes, to our knowledge, the first Bayesian estimator for the parameter c_2 that can actually be applied to real-world images of both small and large size. Numerical simulations conducted on synthetic multifractal processes highlight its excellent estimation performance and its computational efficiency as compared to previous formulations.

The efficiency of the method could be further improved by performing computations entirely in polar coordinates, effectively reducing most 2D calculations to 1D calculations due to radial symmetry. Furthermore, the closed-form Whittle likelihood enables the use of more efficient Hamiltonian Monte Carlo algorithms in the Metropolis-within-Gibbs procedure for the analysis of multivariate images (multi-band, multi-temporal) and large numbers of image patches in remote sensing and biomedical applications. These aspects are currently under investigation.

REFERENCES

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

[2] 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.

[3] 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.

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

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

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

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

[9] H. Wendt, N. Dobigeon, J.-Y. Tourneret, and P. Abry, "Bayesian estimation for the multifractality parameter," in *Proc. IEEE Int. Conf. Acoust., Speech, and Signal Process. (ICASSP)*, Vancouver, Canada, May 2013.

[10] A. D. Poularikas, "The Hankel transform," in *The Handbook of formulas and Tables for Signal Processing*, A. D. Poularikas, Ed. CRC Press LLC, 1999.

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

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

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

[14] J. Guinness and M. Fuentes, "Circulant embedding of approximate covariances for inference from Gaussian data on large lattices," *NC State Department of Statistics Technical Reports, no. 2657*, 2014.

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