BAYESIAN ESTIMATION FOR THE MULTIFRACTALITY PARAMETER

Herwig Wendt[1], *Nicolas Dobigeon*[1], *Jean-Yves Tourneret*[1], *Patrice Abry*[2]

[1,2] University of Toulouse, UMR CNRS 5505, INP-ENSEEIHT, F-31062 Toulouse, France, firstname.lastname@irit.fr
[2] ENS Lyon, Physics Dept., UMR CNRS 5672, F-69007 Lyon, France, patrice.abry@ens-lyon.fr

ABSTRACT

Multifractal analysis has matured into a widely used signal and image processing tool. Due to the statistical nature of multifractal processes (strongly non-Gaussian and intricate dependence) the accurate estimation of multifractal parameters is very challenging in situations where the sample size is small (notably including a range of biomedical applications) and currently available estimators need to be improved. To overcome such limitations, the present contribution proposes a Bayesian estimation procedure for the multifractality (or intermittence) parameter. Its originality is threefold: First, the use of wavelet leaders, a recently introduced multiresolution quantity that has been shown to yield significant benefits for multifractal analysis; Second, the construction of a simple yet generic semi-parametric model for the marginals and covariance structure of wavelet leaders for the large class of multiplicative cascade based multifractal processes; Third, the construction of original Bayesian estimators associated with the model and the constraints imposed by multifractal theory. Performance are numerically assessed and illustrated for synthetic multifractal processes for a range of multifractal parameter values. The proposed procedure yields significantly improved estimation performance for small sample sizes.

***Index Terms*—** multifractal analysis, Bayesian estimation, wavelet leaders, multiplicative cascade processes, log-cumulants

1. MOTIVATIONS, RELATED WORKS, CONTRIBUTIONS

Motivations. Multifractal analysis has become a standard tool for signal and image processing, focussing on the characterization of local regularity fluctuations and scale invariance properties. It has been successfully used in a variety of applications of very different natures, including biomedical (heart rate variability [1], fMRI [2]), physics (turbulence [3]), geophysics (rainfalls [4]), finance [5], Internet traffic [6], to name but a few. A recently introduced powerful formalism for multifractal analysis relies on wavelet leaders $L_X(j,k)$ [13, 11, 16], which are constructed from wavelet coefficients. It assumes that the time averages of the $q-$th powers of $L_X(j,k)$ at given analysis scales $a=2^j$ behave as power-laws over a wide range of scales $a \in [a_m, a_M]$, i.e.,

$$S(q,j) \equiv \frac{1}{n_j}\sum_{k=1}^{n_j} L_X^q(j,k) \simeq a^{\zeta(q)},\ a_m \le a \le a_M. \qquad (1)$$

The so-called *scaling exponents* $\zeta(q)$ fully characterize the scaling properties and local regularity fluctuations. It is known from multifractal theory that the analysis of the full scaling properties of data requires the use of both positive and negative values of q.

Two major classes of processes commonly serve as models for the scaling properties observed in real-world data: self-similar processes, for which $\zeta(q) = qH$ in a neighborhood of $q=0$, fractional Brownian motion (fBm) [7] being the emblematic member of this class, multiplicative cascade-based processes, for which $\zeta(q)$ is a strictly concave function, fBm in multifractal time (MF-fBm) [8, 9, 10] being a well-known member of this class. Deciding which class better models real-world data is of crucial importance in applications since the underlying construction mechanisms are of fundamentally different natures: Additive for self-similar processes, multiplicative for cascade-based processes. Practically, this amounts to testing whether the estimated $\zeta(q)$ are linear or strictly concave [11].

In a seminal contribution [12], B. Castaing suggested the use of the polynomial expansion $\zeta(q) = \sum_{p\ge 1} c_p q^p/p!$ and showed that the coefficients c_p are related to the cumulants of the logarithm of the multiresolution quantities used for the analysis (here, the wavelet leaders $L_X(j,k)$) $C_p(j) =$ Cum$_p \ln L_X(j,k)$ independently of k. Notably, $C_1(j) \equiv \mathbb{E}[\ln L_X(j,k)] = c_1^0 + c_1 \ln 2^j$ and

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

It can be shown theoretically that $c_2 \equiv 0$ implies that $\forall p \ge 3, c_p \equiv 0$ [13]. Estimating c_2, referred to as the *intermittence* or *multifractality* parameter, is thus of prime importance in multifractal analysis since it measures the departure from linearity of $\zeta(q)$ around $q=0$.

Related Works: Estimation of c_2. Historically, scaling and multifractal analysis used to be based either on increments, oscillations or wavelet coefficients [14]. It has later been observed that it should be based on the modulus maxima (or skeleton) of the continuous wavelet transform (CWT) [15]. Recently, it has been shown that the wavelet leader based formulation of multifractal analysis benefits both from a better theoretical grounding and from it being based on the discrete wavelet transform (DWT) [13, 11], thus enabling fast and efficient numerical implementations as well as straightforward extensions to higher dimensions (images notably) [16].

Deciding which class of processes better describes data was classically performed by estimating scaling exponents $\zeta(q)$ for a collection of values of q and testing a posteriori whether $\zeta(q)$ is linear or not (cf., e.g., [15, 11]). Formalizing the test is, however, very difficult because the $S(q,j)$ for different q are, by nature, strongly dependent. This motivated the estimation of c_2 as an alternative [12] and testing a posteriori whether[1] $c_2 \equiv 0$ or $c_2 < 0$ [11].

Estimation in multifractal analysis has been most commonly performed by means of linear regressions of $\log_2 S(q,j)$ versus $\log_2 2^j = j$ for $\zeta(q)$ and of $C_p(j)$ versus $\ln 2^j$ for c_p (cf., e.g., [12, 15, 11]). The use of ordinary versus weighted linear regressions has been documented in [11]. Multifractal analysis was first employed in the context of hydrodynamic turbulence, where experimental data can be collected for long periods of time, yielding very long time series of tens of thousand of samples (this is also the case for Internet traffic monitoring). Then, linear regressions based on Eq. (1) are useful tools: DWT and linear regressions induce

[1] Note also that for self-similar processes $c_2 \equiv 0$ while linearity of $\zeta(q)$ generally holds only in a neighborhood of $q=0$, see, e.g., [17] for details.

very low computational cost and can thus be applied to very long time series. They furthermore yield very satisfactory performance (unbiased estimations with rapidly decreasing variance). However, in numerous other applications where multifractal analysis is commonly used, notably in biomedical applications such as fMRI or heart rate variability, sample size is drastically limited and can be as small as a few hundreds of samples only. For such small sample size, it has been documented that estimators of c_2 based on DWT coefficients are unbiased but their variance is too large for their use in most applications, while wavelet leaders (or skeleton of CWT) have better variance at the price though of a bias increase (cf., [11]).

Attempts to overcome this limitation for small sample size are given by generalized moment approaches. They do, however, heavily depend on fully parametric models for the data and achieve, to the best of our knowledge, only limited actual benefits [18]. The Bayesian framework, classical in parameter inference, has been applied to the specific case of fBm, either in the wavelet domain [19], the frequency domain [20] or directly in the time (or space) domain [21]. Indeed, fBm is a jointly Gaussian process with fully parametric covariance structure and thus fits well in a Bayesian framework. Yet, Bayesian estimation has never been performed for the multifractality parameter c_2. This is essentially due to the statistical properties of scale invariant processes with strictly negative c_2 which strongly depart from Gaussian and exhibit intricate dependence structures that are not fully studied.

Contributions. In real-world applications, the use of fully parametric models is often very restrictive. The challenge addressed in the present contribution thus consists of proposing a Bayesian procedure for the estimation of c_2 for small sample sizes that assumes as little information as possible (essentially the simple relations (1-2)) on data. To this end, it is first shown that for multiplicative cascade based processes the distributions of $\ln L_X(j,k)$ are at each scale $a = 2^j$ well approximated by Gaussian laws whose covariances can be efficiently modeled with few parameters, including the desired c_2 (cf. Section 2). From this generic modeling, valid for all members of the class of multiplicative cascade-based processes, a Bayesian procedure for the estimation of parameter c_2 is devised in Section 3. An appropriate prior distribution is assigned to the multifractality parameter c_2 to ensure relevant constraints inherent to the model (e.g., positivity of the variance of the coefficients $\ln L_X(j,k)$). This prior allows a large class of covariance structures to be efficiently handled. The Bayesian estimators associated with the resulting posterior are then approximated by Monte Carlo sampling. Due to the constraints imposed on the multifractality parameter, a suitable Markov chain Monte Carlo (MCMC) algorithm is designed to sample according to the posterior distribution of interest. Specifically, the admissible set of values for c_2 is explored through a random-walk Metropolis-Hastings scheme that ensures the required positivity constraint. The performance of the Bayesian estimation of parameter c_2 is then assessed by means of Monte Carlo simulations and compared to the one obtained for linear regressions, for various c_2 and different short sample sizes, demonstrating the clear benefits and potentials of the Bayesian approach (cf. Section 4).

2. MODELING WAVELET LEADER STATISTICS

Multifractal processes. For the class of multiplicative cascade based multifractal processes, characterized by a strictly negative c_2, it is well-known that the marginal distributions depart from Gaussianity and that the dependence has a long range structure (cf., e.g., [22]). The statistics of these processes are not known exactly in general except for the (power law) scaling behaviors made explicit in (1) or (2). Departures from Gaussianity and long-range dependence also hold for wavelet coefficients and leaders. The fact that the statistics of such processes and of the corresponding wavelet coefficients and leaders are not known exactly is the key reason that has precluded the use of Bayesian approaches for estimation.

Yet, we show below that the marginals and intra-scale covariance of the logarithm of wavelet leaders associated with multiplicative cascade based processes can be well approximated by a generic semi-parametric model, which will in turn allow us to devise a Bayesian estimation procedure for c_2. A prominent model for this class of process, multifractal random walk (MRW), is chosen here for illustrations since it is easy to simulate and c_1, c_2 are easy to prescribe. It has been verified that equivalent results are obtained for other multiplicative cascade based processes, specifically for MF-fBm. MRW has been introduced in [23] as a non Gaussian process with stationary increments whose multifractal properties mimic those of the celebrated Mandelbrot's multiplicative log-normal cascades. The process is defined as $X(k) = \sum_{k=1}^{n} G_H(k) e^{\omega(k)}$, where $G_H(k)$ consists of the increments of fBm with parameter H. The process ω is independent of G_H, Gaussian, with non trivial covariance: $\mathrm{Cov}[\omega(k_1), \omega(k_2)] = c_2 \ln\left(\frac{L}{|k_1-k_2|+1}\right)$ when $|k_1-k_2| < L$ and 0 otherwise. MRW has scaling properties as in Eq. (1) for $q \in \left[-\sqrt{2/c_2}, \sqrt{2/c_2}\right]$, with $\zeta(q) = (H + c_2)q - c_2 q^2/2$.

Wavelet coefficients and leaders. Let ψ denote the oscillating reference pattern referred to as the mother wavelet. It is characterized by its number of vanishing moments N_ψ, a strictly positive integer, defined as: $\forall n = 0, \ldots, N_\psi - 1$, $\int_{\mathbb{R}} t^k \psi(t) dt \equiv 0$ and $\int_{\mathbb{R}} t^{N_\psi} \psi(t) dt \neq 0$. Further, ψ is chosen such that its dilated and translated templates $\psi_{j,k}(t) = 2^{-j}\psi(2^{-j}t - k)$ form an orthonormal basis of $\mathcal{L}^2(\mathbb{R})$. The ($L^1$-normalized) discrete wavelet transform coefficients $d_X(j,k)$ of X are defined as $d_X(j,k) = 2^{-j/2}\langle \psi_{j,k} | X \rangle$. Readers are referred to, e.g., [24] for detailed introduction to wavelets. *Wavelet leaders* $L_X(j,k)$ are defined as the local supremum of wavelet coefficients taken within a neighborhood over all finer scales [13, 11]: $L_X(j,k) = \sup_{\lambda' \subset 3\lambda_{j,k}} |d_X(\lambda')|$, where $\lambda_{j,k} = [k2^j, (k+1)2^j)$ and $3\lambda_{j,k} = \bigcup_{m\{-1,0,1\}} \lambda_{j,k+m}$.

Modeling the marginal distribution of wavelet leaders. The statistics of wavelet coefficients and – a forteriori – leaders of multiplicative cascade multifractal processes strongly depart from Gaussianity. Numerical simulations reveal, however, that the *logarithm* of wavelet leaders $l_X(j,k) = \ln L_X(j,k)$ (which enters relation (2)) of multiplicative cascade based processes has a distribution very well modeled by a Gaussian. This is illustrated in Fig. 1 (top row) for MRW with small to large $|c_2|$ (weak to strong multifractality).

Modeling the intra-scale covariance of wavelet leaders. The model is motivated by results in [25] which show that the asymptotic covariance of $\ln |d_X(j,k)|$ in random wavelet cascades (a specific multiplicative process directly defined on wavelet coefficients) behaves linearly in $\log_2(\Delta k)$ coordinates. Numerical simulations indicate that the covariance of $l_X(j,k) = \ln L_X(j,k)$ for multiplicative cascade based multifractal processes is well described by

$$\mathrm{Cov}[l_X(j,k), l_X(j,k+\Delta k)] \approx$$
$$\approx \Gamma(j, \Delta k; c_2) = \gamma(c_2) + c_2(\log_2(\Delta k/N) + j)\ln 2 \quad (3)$$

for $3 < \Delta k \ll 2^{-j}N$, where N is the sample size, and for a wide range of the multifractality parameter. We combine (3) with the variance relation (2) to form a piecewise linear (in $\log_2(\Delta k)$ coordinates) model $\Sigma(j, \Delta k; c_2, c_2^0)$ for the full intra-scale covariance

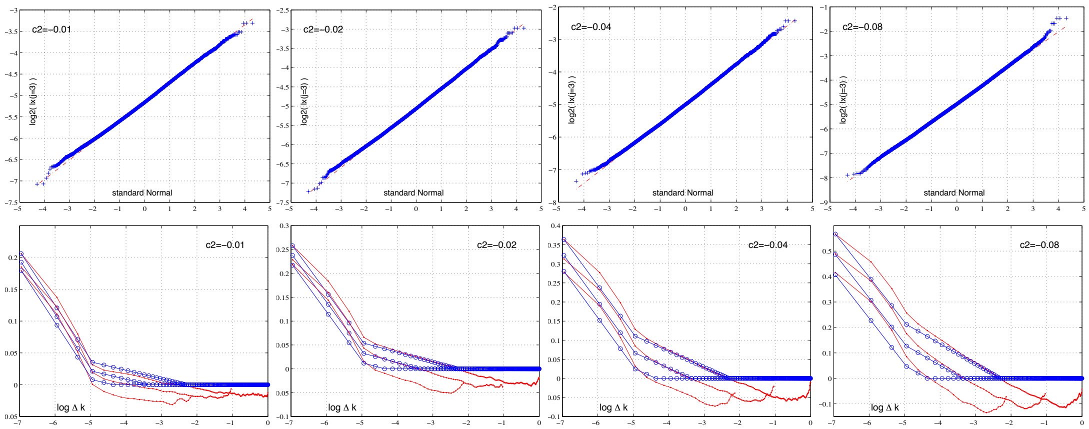

Fig. 1. Top: Quantile plots of $\log_2(l_X(j=3,\cdot))$ against standard Normal. Bottom: covariance estimates (in red) and model (in blue) as functions of $\log_2(\Delta_k/N2^{j_1})$ for scales $j_1=2$ to $j_2=4$ (1000 realizations of length $N=2^9$; $\ln(2)$ factors are absorbed in the axes).

of $l_X(j,k)$:
a. $\Sigma(j,\Delta k;c_2,c_2^0)=c_2^0+jc_2\ln 2$ for $\Delta k=0$
b. $\Sigma(j,\Delta k;c_2,c_2^0)=\max(0,\Gamma(j,\Delta k;c_2))$ for $3\le\Delta k\le 2^{-j}N$
c. $\Sigma(j,0;c_2,c_2^0)$ and $\Sigma(j,3;c_2,c_2^0)$ are connected with a line segment (in $\log_2(\Delta k/N)$ coordinates).
In b.), only non-negative values are admitted for numerical reasons (conditioning number of the covariance matrices $\mathbf{\Sigma}$ used in Section 3). The linear part c.) models the short-term correlations of log wavelet leaders. The parameter $\gamma(c_2)$ in (3) is obtained from the heuristic condition $\text{Cov}[l_X(j,k),l_X(j,k+\Delta k=\lfloor N2^{-j}/5\rfloor]=0$ ($\lfloor\cdot\rfloor$ truncates to integer values). It could in principle be estimated as a third free parameter together with c_2, c_2^0 using the procedure described in Section 3. Fig. 1 (bottom row) plots the intra-scale covariances of $l_X(j,k)$ and the model $\Sigma(j,\Delta k;c_2,c_2^0)$ for MRW with small to large $|c_2|$. Note that the proposed model provides excellent fits for the positive portion of the observed covariances.

3. BAYESIAN ESTIMATION

3.1. Bayesian model

Let j_1 (resp., j_2) denote the finest (resp., coarsest) scale used in the estimation and $l_X(j,k)$ the logarithm of the kth wavelet leader ($k=1,\ldots,n_j$) in the jth scale ($j=j_1,\ldots,j_2$). These coefficients are scale-wise centered, rearranged and stacked in a unique $n\times 1$-vector $\boldsymbol{\ell}_X=[\ell_X(1),\ldots,\ell_X(n)]^T$, with $n=\sum_{j=j_1}^{j_2}n_j$. This vector is assumed to be distributed according to a zero-mean Gaussian distribution with covariance matrix $\mathbf{\Sigma}(\boldsymbol{\gamma}_2)=\mathbb{E}[\boldsymbol{\ell}_X\boldsymbol{\ell}_X^T]$, where $\boldsymbol{\gamma}_2\triangleq\left[c_2,c_2^0\right]^T$ are the parameters to be estimated. Note that, for clarity, the dependence of the covariance matrix on the parameters c_2 and c_2^0 has been explicitly mentioned by denoting $\mathbf{\Sigma}(\boldsymbol{\gamma}_2)$. We propose to estimate the set of parameters $\boldsymbol{\gamma}_2$ in a Bayesian setting. The likelihood function and prior distributions for the unknown parameters required to build the Bayesian model are introduced in the following paragraphs.

Likelihood. The statistical properties of the wavelet leaders introduced in Section 2 yield the following likelihood function for $\boldsymbol{\ell}_X$:

$$f(\boldsymbol{\ell}_X|\boldsymbol{\gamma}_2)=(2\pi)^{-\frac{n}{2}}\left[\det\mathbf{\Sigma}(\boldsymbol{\gamma}_2)\right]^{-1/2}\times\exp\left[-\frac{1}{2}\boldsymbol{\ell}_X^T\mathbf{\Sigma}(\boldsymbol{\gamma}_2)^{-1}\boldsymbol{\ell}_X\right].\quad(4)$$

Prior for $\boldsymbol{\gamma}_2$. To ensure positivity of the variance $C_2(j)=[\mathbf{\Sigma}(\boldsymbol{\gamma}_2)]_{j,j}$ ($j=j_1,\ldots,j_2$) in (2), the parameters c_2 and c_2^0 must belong to the admissible set $\mathcal{C}_2=(\mathcal{C}_2^-\cup\mathcal{C}_2^+)\cap\mathcal{C}_2^m$ with $\mathcal{C}_2^-=\{(c_2,c_2^0)\in\mathbb{R}^2|c_2<0\text{ and }c_2j_2+c_2^0>0\}$, $\mathcal{C}_2^+=\{(c_2,c_2^0)\in\mathbb{R}^2|c_2>0\text{ and }c_2j_1+c_2^0>0\}$ and $\mathcal{C}_2^m=\{(c_2,c_2^0)\in\mathbb{R}^2||c_2|<c_2^m,|c_2^0|<c_2^{0,m}\}$, with $(c_2^m,c_2^{0,m})$ the largest admissible (c_2,c_2^0). In absence of additional prior knowledge regarding (c_2,c_2^0), a uniform prior distribution on $\mathcal{C}_2$ is assigned to $\boldsymbol{\gamma}_2$:

$$f(\boldsymbol{\gamma}_2)\propto\mathbf{1}_{\mathcal{C}_2}(\boldsymbol{\gamma}_2).\quad(5)$$

Posterior distribution. The posterior distribution of $\boldsymbol{\gamma}_2$ can be computed from the Bayes rule:

$$f(\boldsymbol{\gamma}_2|\boldsymbol{\ell}_X)\propto f(\boldsymbol{\ell}_X|\boldsymbol{\gamma}_2)\,f(\boldsymbol{\gamma}_2).\quad(6)$$

Due to the non-trivial dependence of $f(\boldsymbol{\gamma}_2|\boldsymbol{\ell}_X)$ upon the parameters c_2 and c_2^0, computing the Bayesian estimators (e.g., the maximum *a posteriori* (MAP) and the minimum mean square error (MMSE) estimators) associated with (6) is not straightforward. To alleviate the difficulty, it is common to resort to a Markov chain Monte Carlo (MCMC) algorithm to generate samples distributed according to $f(\boldsymbol{\gamma}_2|\boldsymbol{\ell}_X)$ (denoted as $\cdot^{(t)}$, $t=1,\ldots,N_{\text{mc}}$) that are used to approximate the estimators. The proposed algorithm is described next.

3.2. Gibbs sampler

This section describes the Gibbs sampling strategy that allows samples $\left\{c_2^{(t)},c_2^{0(t)}\right\}_{t=1}^{N_{\text{mc}}}$ to be generated according to the posterior (6). This algorithm is divided into two successive steps that consist of sampling according to the conditional distributions associated with the joint distribution $f\left(c_2,c_2^0|\boldsymbol{\ell}_X\right)$. The reader is invited to consult [26] for more details regarding MCMC methods.

Sampling according to $f\left(c_2|c_2^0,\boldsymbol{\ell}_X\right)$. To sample according to the conditional distribution $f\left(c_2|c_2^0,\boldsymbol{\ell}_X\right)$, a Metropolis-within-Gibbs procedure is proposed. Precisely, we use a random-walk algorithm with a normal distribution as the instrumental distribution. Let denote as $\boldsymbol{\gamma}_2^{(t)}=[c_2^{(t)},c_2^{0(t)}]^T$ the current state vector at iteration t of the sampler. A candidate $c_2^{(\star)}$ is drawn according to a proposal distribution $q(c_2^{(\star)}|c_2^{(t)})$ chosen as the Gaussian distribution $\mathcal{N}(c_2^{(t)},\eta^2)$ where η^2 is a given variance (to ensure good mixing properties). Then the proposed state vector $\boldsymbol{\gamma}_2^{(\star)}=[c_2^{(\star)},c_2^{0(t)}]^T$ is accepted with the probability $p_{c_2} = \min(1, \rho_{c_2})$ where ρ_{c_2} is the Metropolis-Hasting acceptance rate

$$\begin{aligned} \rho_{c_2} = & \frac{f\left(\boldsymbol{\gamma}_2^{(\star)}|\boldsymbol{\ell}_X\right)}{f\left(\boldsymbol{\gamma}_2^{(t)}|\boldsymbol{\ell}_X\right)} \frac{q\left(\boldsymbol{\gamma}_2^{(t)}|\boldsymbol{\gamma}_2^{(\star)}\right)}{q\left(\boldsymbol{\gamma}_2^{(\star)}|\boldsymbol{\gamma}_2^{(t)}\right)} \\ = & \left[\frac{\det \boldsymbol{\Sigma}\left(\boldsymbol{\gamma}_2^{(t)}\right)}{\det \boldsymbol{\Sigma}\left(\boldsymbol{\gamma}_2^{(\star)}\right)}\right]^{1/2} \mathbf{1}_{\mathcal{C}_2}\left(\boldsymbol{\gamma}_2^{(\star)}\right) \\ & \times \exp\left[-\frac{1}{2}\boldsymbol{\ell}_X^T \left(\boldsymbol{\Sigma}^{-1}\left(\boldsymbol{\gamma}_2^{(\star)}\right) - \boldsymbol{\Sigma}^{-1}\left(\boldsymbol{\gamma}_2^{(t)}\right)\right)\boldsymbol{\ell}_X\right] \end{aligned} \quad (7)$$

Finally, the current vector $\boldsymbol{\gamma}_2^{(t)}$ is updated as $\boldsymbol{\gamma}_2^{(t+\frac{1}{2})} = \boldsymbol{\gamma}_2^{(\star)}$ or as $\boldsymbol{\gamma}_2^{(t+\frac{1}{2})} = \boldsymbol{\gamma}_2^{(t)}$ with probabilities p_{c_2} and $1 - p_{c_2}$, respectively.

Sampling according to $f\left(c_2^0|c_2, \boldsymbol{\ell}_X\right)$. In a similar fashion, to sample according to $f\left(c_2^0|c_2, \boldsymbol{\ell}_X\right)$ a random-walk Metropolis-Hastings step is used to update the current vector $\boldsymbol{\gamma}_2^{(t+\frac{1}{2})} = [c_2^{(t+\frac{1}{2})}, c_2^{0(t+\frac{1}{2})}]^T$. At iteration $t+\frac{1}{2}$, a candidate $c_2^{0(\star)}$ is proposed according to a Gaussian instrumental distribution $\mathcal{N}(c_2^{0(t+\frac{1}{2})}, \eta_0^2)$, leading to the candidate $\boldsymbol{\gamma}_2^{(\star)} = [c_2^{(t+\frac{1}{2})}, c_2^{0(\star)}]^T$. The current state vector $\boldsymbol{\gamma}_2^{(t+\frac{1}{2})}$ is updated either as $\boldsymbol{\gamma}_2^{(t+1)} = \boldsymbol{\gamma}_2^{(\star)}$ with probability $p_{c_2^0}$ or as $\boldsymbol{\gamma}_2^{(t+1)} = \boldsymbol{\gamma}_2^{(t+\frac{1}{2})}$ with probability $1 - p_{c_2^0}$, where $p_{c_2^0} = \min(1, \rho_{c_2^0})$ and $\rho_{c_2^0}$ is computed as in (7).

3.3. Approximating the Bayesian estimators

The proposed Gibbs sampler enables us to generate N_{mc} samples $\{\boldsymbol{\gamma_2}^{(t)}\}_{t=1}^{N_{\text{mc}}}$ which are asymptotically distributed according to the distribution (6). After a short burn-in of N_{bi} iterations, these samples can be used to approximate the Bayesian estimators, i.e.,

$$\hat{\gamma}_2^{\text{MMSE}} \approx \frac{1}{N_r} \sum_{t=N_{\text{bi}+1}}^{N_{\text{mc}}} \boldsymbol{\gamma_2}^{(t)}; \; \hat{\gamma}_2^{\text{MAP}} \approx \underset{t=1,\ldots,N_{\text{mc}}}{\text{argmax}} \; f\left(\boldsymbol{\gamma_2}^{(t)}|\boldsymbol{\ell}_X\right).$$

4. ESTIMATION PERFORMANCE

We analyze the estimation performance for 200 realizations of MRW defined as follows: sample size $N = 256$ or $N = 512$ and parameter c_2 varying from weak ($c_2 = -0.01$) to strong ($c_2 = -0.08$) multifractality. We use Daubechies' wavelet with $N_\psi = 2$ vanishing moments, scaling range $j_1 = 2$ and $j_2 = 4$ ($N = 256$) and $j_2 = 5$ ($N = 512$), respectively. The Gibbs sampler is run with $N_{mc} = 700$ and $N_{bi} = 350$. Table 1 summarizes the mean, bias and (root) mean square error (RMSE) of the weighted linear regression (LF) and Bayesian MMSE and MAP estimators for c_2. The results clearly indicate that the proposed semi-parametric Bayesian estimation procedure significantly improves the quality of c_2 estimates for the small sample sizes considered here: Compared to weighted linear regression, the Bayesian estimators have RMSEs systematically and strongly reduced by a factor ranging from 3 (for $|c_2|$ small) to 4 (for $|c_2|$ large). This drastic improvement of estimation quality is mostly due to the significant reduction of variance of the Bayesian estimators (indeed, standard deviations are reduced by a factor 3 to 4 for small and large $|c_2|$, respectively) while the bias plays a minor role: linear regression and Bayesian estimators display similar bias for large $|c_2|$ and linear fits have slightly smaller bias for small $|c_2|$. This slight advantage in terms of bias is, however, strongly outweighed by the severely larger variability of linear regression based estimators. Furthermore, note that the increase in variance when

$\mathbf{N = 2^8}$	c_2	-0.01	-0.02	-0.03	-0.04	-0.06	-0.08
mean	LF	-0.019	-0.023	-0.037	-0.044	-0.072	-0.094
mean	MMSE	-0.023	-0.031	-0.039	-0.045	-0.058	-0.073
mean	MAP	-0.019	-0.028	-0.039	-0.045	-0.061	-0.076
std	LF	0.055	0.064	0.073	0.077	0.097	0.106
std	MMSE	0.015	0.018	0.020	0.020	0.021	0.024
std	MAP	0.016	0.020	0.022	0.022	0.022	0.025
rmse	LF	0.056	0.064	0.073	0.077	0.097	0.107
rmse	MMSE	0.020	0.021	0.022	0.021	0.021	0.025
rmse	MAP	0.018	0.021	0.023	0.023	0.022	0.025
$\mathbf{N = 2^9}$	c_2	-0.01	-0.02	-0.03	-0.04	-0.06	-0.08
mean	LF	-0.010	-0.021	-0.033	-0.037	-0.068	-0.077
mean	MMSE	-0.019	-0.029	-0.037	-0.047	-0.062	-0.076
mean	MAP	-0.017	-0.028	-0.036	-0.047	-0.063	-0.078
std	LF	0.032	0.037	0.044	0.056	0.052	0.066
std	MMSE	0.011	0.012	0.014	0.016	0.017	0.015
std	MAP	0.011	0.014	0.015	0.016	0.018	0.017
rmse	LF	0.032	0.037	0.045	0.056	0.052	0.066
rmse	MMSE	0.014	0.015	0.016	0.017	0.017	0.016
rmse	MAP	0.013	0.016	0.016	0.018	0.018	0.017

Table 1. Mean, standard deviation and root mean square error of estimators of c_2 for $N = 256$ (top) and $N = 512$ (bottom).

increasing $|c_2|$ (due to stronger variability of the data) is less pronounced for the Bayesian estimators. Finally, when comparing the two Bayesian estimators, MMSE is slightly advantageous in terms of bias, MAP in terms of variance, and both yield equivalent RMSEs. The gain in estimation performance is a direct consequence of the covariance structure included in the proposed Bayesian model. It furthermore demonstrates the relevance of the proposed model for the statistics of the logarithm of wavelet leaders. Similar results are obtained for MF-fBm and are not presented here for space reasons.

5. CONCLUSIONS AND PERSPECTIVES

We have, to the best of our knowledge, devised the first operational Bayesian estimation procedure for the multifractality parameter c_2. The procedure is designed for the large class of multiplicative cascade based multifractal processes. Its versatility results from the proposition of a simple yet accurate and generic statistical model for the logarithm of wavelet leaders that incorporates the marginal distributions, the covariance at each scale, and the power law scaling of the variance across scales. An MCMC algorithm is proposed to sample according to the joint posterior distribution of the multifractal parameters, ensuring inherent constraints for the multifractal paradigm. The procedure enables the reliable estimation of the multifractality parameter c_2 in applications where sample size is small and the variance of commonly used linear regression based estimators is prohibitively large. Indeed, the proposed Bayesian estimators yield a decrease in variance and MSE of up to a factor 4 (at the price though of increasing computation time by orders of magnitude). The performance could be further improved by incorporating (application dependent) prior information (here, vague priors have been used). The Bayesian framework also enables the construction of confidence intervals and hypothesis tests for c_2. Their performance are currently under study. The procedure is currently being applied to the analysis of fMRI and heart rate variability data. Future work includes the definition of a generic model for the joint time-scale covariance of wavelet leaders as well as extensions of the proposed procedure to 2D images.

6. REFERENCES

[1] A. L. Goldberger, L. A. Amaral, J. M. Hausdorff, P. C. Ivanov, C. K. Peng, and S. H. E., "Fractal dynamics in physiology: alterations with disease and aging," *Proc. Natl. Acad. Sci. USA*, vol. 99, no. Suppl 1, pp. 2466–2472, 2002.

[2] P. Ciuciu, G. Varoquaux, P. Abry, S. Sadaghiani, and A. Kleinschmidt, "Scale-free and multifractal dynamic properties of fMRI signals during rest and task," *Frontiers in Physiology*, vol. 3, no. 186, June 2012.

[3] B. Mandelbrot, "Intermittent turbulence in self-similar cascades: divergence of high moments and dimension of the carrier," *J. Fluid Mech.*, vol. 62, pp. 331–358, 1974.

[4] E. Foufoula-Georgiou and P. Kumar, Eds., *Wavelets in Geophysics*. San Diego: Academic Press, 1994.

[5] B. Mandelbrot, *Fractals and scaling in finance*, ser. Selected Works of Benoit B. Mandelbrot. New York: Springer-Verlag, 1997, discontinuity, concentration, risk, Selecta Volume E, With a foreword by R. E. Gomory.

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

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

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

[9] E. Bacry and J. Muzy, "Multifractal stationary random measures and multifractal random walks with log-infinitely divisible scaling laws," *Phys. Rev. E*, vol. 66, 2002.

[10] P. Chainais, R. Riedi, and P. Abry, "On non scale invariant infinitely divisible cascades," *IEEE Trans. Inform. Theory*, vol. 51, no. 3, March 2005.

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

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

[13] S. Jaffard, "Wavelet techniques in multifractal analysis," in *Fractal Geometry and Applications: A Jubilee of Benoît Mandelbrot, M. Lapidus and M. van Frankenhuijsen Eds., Proceedings of Symposia in Pure Mathematics*, M. Lapidus and M. van Frankenhuijsen, Eds., vol. 72, no. 2. AMS, 2004, pp. 91–152.

[14] U. Frisch, *Turbulence, the Legacy of A.N. Kolmogorov*. Cambridge University Press, 1995.

[15] J. Muzy, E. Bacry, and A. Arneodo, "Multifractal formalism for fractal signals: The structure function approach versus the wavelet transform modulus-maxima method," *Phys. Rev. E*, vol. 47, no. 2, pp. 875–884, 1993.

[16] H. Wendt, S. G. Roux, P. Abry, and S. Jaffard, "Wavelet leaders and bootstrap for multifractal analysis of images," *Signal Proces.*, vol. 89, pp. 1100–1114, 2009.

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

[18] T. Lux, "Higher dimensional multifractal processes: A GMM approach," *Journal of Business and Economic Statistics*, vol. 26, pp. 194–210, 2007.

[19] G. Wornell and A. V. Oppenheim, "Estimation of fractal signals from noisy measurements using wavelets," *IEEE Trans. on Signal Processing*, vol. 40, no. 3, pp. 611–623, March 1992.

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

[21] T. T. Lundhal, W. J. Ohley, S. M. Kay, and R. Siert, "Fractional brownian motion: an ML estimator and its application to image texture," *IEEE Trans. on Medical Imaging*, vol. MI-5, no. 3, pp. 152–161, sep 1986.

[22] B. Vedel, H. Wendt, P. Abry, and S. Jaffard, *On the impact of the number of vanishing moments on the dependence structures of compound Poisson motion and fractional Brownian motion in multifractal time*, ser. Lecture Notes in Statistics. Springer, Oct. 2010, pp. 71–102.

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

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

[25] A. Arneodo, E. Bacry, and J. Muzy, "Random cascades on wavelet dyadic trees," *J. Math. Phys.*, vol. 39, no. 8, pp. 4142–4164, 1998.

[26] C. P. Robert and G. Casella, *Monte Carlo Statistical Methods*, 2nd ed. New York, NY, USA: Springer, 2004.