Open Archive TOULOUSE Archive Ouverte (OATAO)

OATAO is an open access repository that collects the work of Toulouse researchers and makes it freely available over the web where possible.

This is an author-deposited version published in: http://oatao.univ-toulouse.fr/ Eprints ID: 12435

To link to this article: DOI: 10.1109/ICASSP.2013.6638929
URL: http://dx.doi.org/10.1109/ICASSP.2013.6638929


Any correspondence concerning this service should be sent to the repository administrator: staff-oatao@listes-diff.inp-toulouse.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 multisolutions 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.,

\[
S(q,j) \equiv \frac{1}{n_j} \sum_{k=1}^{n_j} L_X^q(j,k) \approx \alpha^q, \quad a_m \leq a \leq a_M.
\]  

(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 \(\zeta(q) = \sum_{p=1}^{P} c_p q^p / p!\) 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 \(L_X(j,k)\)) \(\zeta_p(j) = \text{Cum}_p \ln L_X(j,k)\) independently of \(k\). Notably, \(\zeta_1(j) \equiv \text{E} \ln L_X(j,k)\) for \(\ln 2^j\) and

\[
C_2(j) \equiv \text{Var} \ln L_X(j,k) = c_0^2 + c_2 2^j. \quad (2)
\]

It can be shown theoretically that \(c_2 \equiv 0\) implies that \(Vp \geq 3, c_p \equiv 0\) [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 \(c_2 = 0\) or \(c_2 < 0\) [11].

Estimation in multifractal analysis has been most commonly performed by means of linear regressions of \(\log_3 S(q,j)\) 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

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
(unsampled estimations with rapidly decreasing variance). However,
in numerous other applications where multifractal analysis is com-
monly 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, heav-
ily 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 ap-
plicated 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 para-
metric models is often very restrictive. The challenge addressed in
the present contribution thus consists of proposing a Bayesian pro-
dure 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 \( \ln L_X(j, k) \) 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 pro-
cedure for the estimation of parameter \( c_2 \) is devised in Section 3. An
appropriate prior distribution is assigned to the multifractality pa-
rameter \( c_2 \) to ensure relevant constraints inherent to the model (e.g.,
positivity of the variance of the coefficients in \( \ln L_X(j, k) \)). 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 as-
sessed 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 Gaus-
sianity and that the dependence has a long range structure (cf., e.g.,
[22]). The statistics of these processes are not known exactly in gen-
eral 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 covari-
ance of the logarithm of wavelet leaders associated with multiplica-
tive cascade based processes can be well approximated by a generic
semi-parametric model, which will then 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
\( MF^2 \) 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 cas-
cades. The process is defined as \( X(k) = \sum_{n=1}^{N} G_H(k) e^{\lambda X} \),
where \( G_H(k) \) consists of the increments of fBm with parameter \( H \).
The process \( \omega \) is independent of \( G_H \), Gaussian, with non
trivial covariance: \( \text{Cov}[\omega(k_1), \omega(k_2)] = c_2 \ln \left( \frac{1}{|k_1-k_2|+\epsilon} \right) \) 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 \( \ln(\mu) = (H+c_2)q-c_2 q^2/2 \).

Wavelet coefficients and leaders. Let \( \psi \) denote the oscillating
reference pattern referred to as the mother wavelet. It is charac-
terized by its number of vanishing moments \( N_\psi \), a strictly posi-
tive integer, defined as: \( \forall n = 0, \ldots, N_\psi - 1, \int_{R} \psi(t)^n dt \equiv 0 \)
and \( \int_{R} t^n \psi(t) dt \neq 0 \). Further, \( \psi \) is chosen such that its dilated
and translated templates \( \{ \psi_0,\psi_{k,\Delta} \} \) form an or-
thonormal basis of \( L^2(R) \). The (\( L^2 \)-normalized) discrete wavelet
transform coefficients \( d_X(j, k) \) of \( X \) are defined as \( d_X(j, k) = 2^{-j/2} \sum_{\Delta} \psi_{j,\Delta} X \). Readers are referred to, e.g., [24] for detailed in-
roduction 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]: \( L_X(j, k) = \sup_{\Delta \subset \Lambda(j,k)} |d_X(j,j)| \),
where \( \Lambda(j,k) = [k2^j, (k+1)2^j) \) and \( \Lambda(j,k) = \bigcup_{\Delta \subset \Lambda^{-1}(j, k)} \Lambda_\Delta \).

Modeling the marginal distribution of wavelet leaders. The
statistics of wavelet coefficients and – a fortiori – leaders of mul-
plicative cascade multifractal processes strongly depart from Gaus-
sianity. Numerical simulations reveal, however, that the logarithm
\( \text{logarithm} \) 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 asym-
ptotic covariance of \( \ln|d_X(j, k)| \) in random wavelet cascades (a spe-
cific 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 mul-
plicative cascade based multifractal processes is well described by

\[
\text{Cov}[l_X(j, k), l_X(j, k + \Delta k)] \approx \Gamma(\Delta k, c_2) \\
\approx \Gamma(\Delta k, c_2) + c_2 \log_2(\Delta k/N) + j \ln 2
\]
of \( l_X(j, k) \):
\( \Sigma \) in Section 2 yield the following likelihood function for \( l_X \):
\[
f(l_X | \gamma_2) = (2\pi)^{-n/2} \frac{1}{\sqrt{\det \Sigma}} \exp \left( -\frac{1}{2} (l_X - \mu)^T \Sigma^{-1} (l_X - \mu) \right).
\]

Prior for \( \gamma_2 \). To ensure positivity of the variance \( C_2(j) = \Sigma(\gamma_2)_j,j \) in (2), the parameters \( c_2 \) and \( \beta_2 \) must belong to the admissible set
\[
C_2 = \{(c_2, \beta_2) \in \mathbb{R}^2 \mid c_2 < 0 \} \quad \text{and} \quad C_{2,0} = \{(c_2, \beta_2) \in \mathbb{R}^2 \mid c_2 > 0 \}.
\]

Posterior distribution. The posterior distribution of \( \gamma_2 \) can be computed from the Bayes rule:
\[
\pi(\gamma_2 | l_X) \propto f(l_X | \gamma_2) \pi(\gamma_2).
\]

Due to the non-trivial dependence of \( f(l_X | \gamma_2) \) upon the parameters \( c_2 \) and \( \beta_2 \), 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 \( \pi(\gamma_2 | l_X) \) (denoted as \( \{\hat{\gamma}_t\}, t = 1, \ldots, N_{\text{MC}} \)) 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 sampling \( \{c_2(t) = \hat{c}_2(t), \beta_2(t) = \hat{\beta}_2(t)\} \) 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 \( f(c_2, \beta_2 | l_X) \). The reader is invited to consult [26] for more details regarding MCMC methods.

Sampling according to \( f(c_2) \). To sample according to the conditional distribution \( f(c_2 | c_2, \beta_2) \), 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 \( \gamma^*(t) = [c_2(t), \beta_2(t)]^T \) the current state vector at iteration \( t \) of the sampler. A candidate \( c_2^*(t) \) is drawn according to a proposal distribution \( q(c_2^*(t) | c_2(t), \beta_2(t)) \) chosen as the Gaussian distribution \( N(c_2^*(t), \eta^2) \) where \( \eta^2 \) is a given variance (to ensure good mixing properties). Then the proposed state vector \( \gamma^*_t = [c_2^*(t), \beta_2^*(t)]^T \) is 

\[
\text{f}(\gamma | l_x) = (2\pi)^{-n/2} |\det \Sigma|^\frac{1}{2} \exp \left( -\frac{1}{2} (l_x - \mu)^T \Sigma^{-1} (l_x - \mu) \right).
\]
accepted with the probability \( p_{c2} = \min(1, \rho_{c2}) \) where \( \rho_{c2} \) is the Metropolis-Hasting acceptance rate

\[
\rho_{c2} = \frac{f(\gamma_{2}^{(t)}|\mathbf{X}) p(\gamma_{2}^{(t)}|\mathbf{r}_{2}^{(t)})}{f(\gamma_{2}^{(t)}|\mathbf{X}) p(\gamma_{2}^{(t+1)}|\mathbf{r}_{2}^{(t+1)})} = \frac{\det \mathbf{R}_{(\gamma_{2}^{(t)})}^{1/2}}{\det \mathbf{R}_{(\gamma_{2}^{(t+1)})}^{1/2}} 1_{c2}(\gamma_{2}^{(t+1)}) \]  

(7)

\[ \times \exp \left[ -\frac{1}{2} \mathbf{X}^{T} \left( \Sigma_{2}^{-1}(\gamma_{2}^{(t)}) - \Sigma_{2}^{-1}(\gamma_{2}^{(t+1)}) \right) \mathbf{X} \right] \]

Finally, the current vector \( \gamma_{2}^{(t)} \) is updated as \( \gamma_{2}^{(t+1)} = \gamma_{2}^{(t)} \) or as \( \gamma_{2}^{(t+1)} = \gamma_{2}^{(t)} \) with probability \( p_{c2} \) or as \( \gamma_{2}^{(t+1)} = \gamma_{2}^{(t+1)} \) with probability \( 1 - p_{c2} \), where \( p_{c2} = \min(1, \rho_{c2}) \) and \( \rho_{c2} \) is computed as in (7).

3.3. Approximating the Bayesian estimators

The proposed Gibbs sampler enables us to generate \( N_{mc} \) samples \( \{\gamma_{2}^{(t)}\}_{t=0}^{N_{mc}} \) 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.,

\[ \gamma_{2}^{\text{MMSE}} \approx \frac{1}{N_{mc}} \sum_{t=N_{bi}+1}^{N_{mc}} \gamma_{2}^{(t)}; \gamma_{2}^{\text{MAP}} \approx \arg \max_{t=1,...,N_{mc}} f(\gamma_{2}^{(t)}|\mathbf{X}) \]

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 Dubecchies’ wavelet with \( N_w = 2 \) 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 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-Bm 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.

| \( N = 2^n \) | \( c_2 \) | -0.01 | -0.02 | -0.03 | -0.04 | -0.06 | -0.08 |
|---|---|
| LF | -0.019 | -0.023 | -0.037 | -0.044 | -0.072 | -0.094 |
| MMSE | -0.023 | -0.031 | -0.045 | -0.058 | -0.073 |
| MAP | -0.019 | -0.029 | -0.045 | -0.063 | -0.076 |
| LF | 0.055 | 0.064 | 0.074 | 0.097 | 0.106 |
| MMSE | 0.015 | 0.018 | 0.020 | 0.021 | 0.024 |
| MAP | 0.016 | 0.020 | 0.022 | 0.022 | 0.025 |
| LF | 0.056 | 0.064 | 0.073 | 0.097 | 0.107 |
| MMSE | 0.020 | 0.021 | 0.022 | 0.021 | 0.025 |
| MAP | 0.018 | 0.021 | 0.023 | 0.022 | 0.025 |

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


