BOOTSTRAP FOR LOG WAVELET LEADERS CUMULANT BASED MULTIFRACTAL ANALYSIS

Herwig Wendt, Stephane G. Roux, Patrice Abry

Physics Lab., CNRS UMR 5672, Ecole Normale Supérieure de Lyon, France
Tel: 33 47272 8493, Herwig.Wendt@ens-lyon.fr, Stephane.Roux@ens-lyon.fr, Patrice.Abry@ens-lyon.fr

ABSTRACT

Multifractal analysis, which mostly consists of estimating scaling exponents related to the power law behaviors of the moments of wavelet coefficients, is becoming a popular tool for empirical data analysis. However, little is known about the statistical performance of such procedures. Notably, despite their being of major practical importance, no confidence intervals are available. Here, we choose to replace wavelet coefficients with wavelet Leaders and to use a log-cumulant based multifractal analysis. We investigate the potential use of bootstrap to derive confidence intervals for wavelet Leaders log-cumulant multifractal estimation procedures. From numerical simulations involving well-known and well-controlled synthetic multifractal processes, we obtain two results of major importance for practical multifractal analysis : we demonstrate that the use of Leaders instead of wavelet coefficients brings significant improvements in log-cumulant based multifractal estimation, we show that accurate bootstrap designed confidence intervals can be obtained for a single finite length time series.

1. MOTIVATION

Scaling or Multifractal analysis [1, 2, 3] is becoming a standard analysis procedure commonly available in empirical data analysis toolboxes. Scaling, or scale invariance, is indeed a property that has been extensively observed in empirical data produced from numerous applications of very different nature. Multifractal analysis mostly consists of measuring scaling exponents, whose values are then commonly involved in various detection, identification or classification tasks. Despite becoming increasingly popular in data analysis, multifractal estimation procedures remain poorly studied. Questions, that may appear natural or simple, such as *should one prefer increments or wavelet coefficients?*, *should one perform weighted or non weighted regressions?* or *what are the typical sizes of the confidence intervals?* still remain insufficiently addressed. However, for practical uses and purposes, elements of answers to such issues are crucial. Indeed, in many real life applications, the sizes of confidence intervals are as important as the values of the scaling exponents themselves, as no classification, discrimination or hypothesis testing are possible without them.

In the present contribution, we elaborate on multifractal analysis in combining together three key improvements: **wavelet Leaders, log-cumulants and bootstrap**.

First, it is now considered as classical and powerful to chose wavelet coefficients as the key multiresolution quantities multifractal analysis should be based on [2, 3]. Very recent findings reported in [4, 5] indicate that an accurate multifractal analysis should be based on wavelet Leaders rather than on wavelet coefficients. Indeed, the former enable to estimate exactly the entire multifractal spectrum and to analyze accurately processes containing oscillating singularities when the later do not. For further details, the reader is referred to [4, 5]. Wavelet Leaders are defined in Section 4.

Second, multifractal estimation procedures are commonly based on structure functions (i.e., power law behaviors of the moments of multiresolution quantities) as in Eqs. (1) or (2) below. However, it has been proposed to use instead the cumulants of the logarithm of the multiresolution quantities. This was originally introduced in the early nineties in [6] and largely developed in [7]. We follow here these promising developments. Cumulant based estimation procedures are described in Section 2.

Third, we investigate on potential benefits resulting from the use of non parametric bootstrap for multifractal estimation. In bootstrapping, the distribution of an estimator is approximated through repeated resampling with replacement from the available data. The technique was introduced in the eighties [8] and has recently regained interest due to continuously growing computer facilities [9, 10, 11]. Bootstrapping has been used in the wavelet domain after the pioneering work reported in [12]. It has also been considered for the estimation of the Hurst parameter of self-similar processes [13, 14]. In the present work, we intend to explore the use of bootstrap in two respects: estimation procedure enhancement and confidence intervals derivation. Bootstrap procedures are detailed in Section 3.

Therefore, the aims of the present article are to contribute to the answers of the two following questions: Does the use of (log-cumulant) wavelet Leaders improve multifractal estimation procedures ? Can bootstrap provide us with reliable confidence intervals ? To address these questions, we first compare the statistical performance of estimation procedures based on wavelet coefficients and on wavelet Leaders. Second, we use a coverage procedure to compare confidence intervals obtained from a simple bootstrap approach. Results are derived by applying our procedures to a large number of realizations of synthetic scaling processes with a priori known and controlled multifractal properties (cf. Section 5). In Section 6, we show that the use of wavelet Leaders instead of wavelet coefficients brings substantial improvements in multifractal estimation performance. We also clearly demonstrate that the use of bootstrap procedures enables us to obtain highly reliable confidence intervals. We end up with a practical procedure that provides us with both accurate multifractal estimates and confidence intervals and that can actually be used for analyzing a single run of empirical data with finite observation duration.

2. MULTIFRACTAL, CUMULANTS AND WAVELETS

2.1. Definitions

Wavelet Coefficients. Let $X(t),\ t \in [0,n)$ denote the process under analysis and n its observation duration. $\psi_0(t)$ is a reference pattern with fast exponential decay, called the *mother-wavelet*. It is characterized by its *number of vanishing moments*, a strictly positive integer $N_\psi \geq 1$ defined as: $\forall k = 0, 1, \ldots, N_\psi - 1$, $\int_{\mathcal{R}} t^k \psi_0(t)dt \equiv 0$ and $\int_{\mathcal{R}} t^{N_\psi} \psi_0(t)dt \neq 0$. Let us further denote by $\{\psi_{j,k}(t) = 2^{-j}\psi_0(2^{-j}t - k), j \in \mathcal{Z}, k \in \mathcal{Z}\}$ templates of ψ_0 dilated to scales 2^j, and translated to time positions $2^j k$. The wavelet coefficients of X are defined as $d_X(j,k) = \langle \psi_{j,k} | X \rangle$.

Scaling and Multifractal. A process X is said to possess *scale invariance* or *scaling* properties if, for some statistical orders $q \in [q_*^-, q_*^+]$ (cf. [15]), the time averages of $|d_X(j,k)|^q$ taken at fixed scales display power law behaviors with respect to scales $a = 2^j$,

$$\langle |d_X(j,\cdot)|^q \rangle = F_q |a|^{\zeta(q)}, \tag{1}$$

over a wide range of scales $a \in [a_m, a_M]$, $a_M/a_m \gg 1$. The $\zeta(q)$ are referred to as the scaling exponents of X and are closely related to its multifractal spectrum [5].

When $\zeta(q)$ is linear in q, i.e. $\zeta(q) = qH$, the process X is said to be *monofractal*. This is, for instance, the case for finite variance self similar processes such as fractional Brownian motion. When $\zeta(q) \neq qH$, X is said to be *multifractal*. This is clearly only a poor and *operational* definition of multifractality. However, for the purposes of this article it is sufficient. We refer the reader to, e.g., [1], for a thorough introduction to multifractal analysis.

Cumulants. [6, 7] For some processes, Eq. (1) is equivalent to

$$\mathbb{E}|d_X(j,\cdot)|^q = F_q |2^j|^{\zeta(q)}. \tag{2}$$

Using a second characteristic function type expansion, one can rewrite Eq. (2) as :

$$\ln \mathbb{E} e^{q \ln |d_X(j,\cdot)|} = \sum_{p=1}^{\infty} C_p^j \frac{q^p}{p!} = \ln F_q + \zeta(q) \ln 2^j, \tag{3}$$

where C_p^j stand for the cumulant of order $p \geq 1$ of the random variable $\ln |d_X(j,\cdot)|$. Combining Eqs (2) and (3) yields that the cumulants of $\ln |d_X(j,\cdot)|$ have to be of the form:

$$\forall p \geq 1: \quad C_p^j = c_p^0 + c_p \ln 2^j \tag{4}$$

and therefore that

$$\ln \mathbb{E} e^{q \ln |d_X(j,\cdot)|} = \underbrace{\sum_{p=1}^{\infty} c_p^0 \frac{q^p}{p!}}_{\ln F_q} + \underbrace{\sum_{p=1}^{\infty} c_p \frac{q^p}{p!}}_{\zeta(q)} \ln 2^j, \tag{5}$$

where c_p^0 and c_p do not depend on the scale 2^j.

Thus, the measurements of the scaling exponents $\zeta(q)$ can be interestingly replaced by those of the *log-cumulants* c_p. This is mainly motivated by the fact that it emphasizes the difference between monofractal ($\forall p \geq 2: \ c_p \equiv 0$) and multifractal processes [6, 7]. The next section describes estimation procedures for the c_ps.

2.2. Estimation Procedures

Commonly, the scaling exponents $\zeta(q)$ are estimated by linear fits performed in $\log_2 \langle |d_X(j,\cdot)|^q \rangle$ vs. $\log_2 |2^j|$ plots (see e.g., [2]). In the present work, we explore the alternative estimation of the equivalent quantities c_p.

Cumulant estimations. Given n_j coefficients $d(j,k)$ and thus samples $Y_j(k) = \ln |d_X(j,k)|$, the asymptotically unbiased and consistent standard estimators (see e.g., [16]) are employed to obtain estimates $\hat{C}_p^j$ for the cumulants of $\ln |d_X(j,\cdot)|$.

Linear regressions. From these $\hat{C}_p^j$s, the c_p can then be estimated by linear regression (cf. equation (4)),

$$\hat{c}_p = \log_2 e \sum_{j=j_1}^{j_2} w_j \hat{C}_p^j. \tag{6}$$

Theoretical performance. Since the $\hat{C}_p^j$s are asymptotically unbiased and consistent, the $\hat{c}_p$s are asymptotically unbiased. As detailed in, e.g., [2], the $d_X(j,k)$ of scaling processes are weakly correlated. Hence, one can approximate the variance of $\hat{c}_p$ as: $\mathrm{Var}\, \hat{c}_p \simeq (\log_2 e)^2 \sum_{j=j_1}^{j_2} w_j^2 \mathrm{Var}\, \hat{C}_p^j$. Thus, the $\hat{c}_p$s are as well consistent.

Weights. The weights w_j have to satisfy the constraints $\sum_{j_1}^{j_2} j w_j \equiv 1$ and $\sum_{j_1}^{j_2} w_j \equiv 0$ and can be expressed as $w_j = \frac{1}{b_j} \frac{S_0 j - S_1}{S_0 S_2 - S_1^2}$, with $S_i = \sum_{j_1}^{j_2} j^i / b_j$, $i = 0, 1, 2$. The positive numbers b_j are freely selectable and reflect the confidence granted to each $\hat{C}_p^j$. We have chosen to compare three cases, corresponding respectively to i) non-weighted regression, ii) the $d_X(j,k)$ can be idealized to independent random variables (cf. [2]), and iii) the confidence level is set proportional to the inverse of the estimated variances $\hat{\sigma}_p^2(j)$ of $\hat{C}_p^j$ (in the present case, these variances will be estimated by bootstrap):

$w_{0,j}$: $b_j = 1$. (non-weighted regression)
$w_{1,j}$: $b_j = 1/n_j$ (assuming $\hat{C}_p^{j_m}$, $\hat{C}_p^{j_n}$ uncorrelated)
$w_{2,j}$: $b_j = \hat{\sigma}_p(j)^2$ ($(\hat{\sigma}_p(j)^2$: estimate of variance of $\hat{\kappa}_{p,(\mathcal{Y}_a)}$).

3. BOOTSTRAP

We use bootstrap generated nonparametric empirical distributions (see e.g. [8, 9, 10, 11]) for first estimating the variance $\hat{\sigma}_p^2(j)$ of $\hat{C}_p^j$, and second for constructing confidence intervals for $\hat{c}_p$. As the wavelet coefficients of scaling processes at a given scale are weakly correlated, we adopt a moving blocks bootstrap with overlapping blocks of length L. At each scale $a = 2^j$, the R bootstrap resamples $\mathcal{D}_j^{*(1)}, \cdots, \mathcal{D}_j^{*(R)}$ are generated from the original sample $\mathcal{D}_j = \{d_X(j,1), \cdots, d_X(j,n_j)\}$. Each resample $\mathcal{D}_j^* = \left\{ d_X^{*(j,1)}(\cdot), \cdots, d_X^{*(j,n_j)}(\cdot) \right\}$ represents an unsorted collection of n_j sample points, drawn blockwise and with replacement from the original sample. These collections $\mathcal{D}_j^{*(r)}$ are used to compute R bootstrap cumulant estimates $r = 1, \ldots, R$, $\hat{C}_p^{j\,*(r)}$. In turns, these $C_p^{j\,*(r)}$ are used for obtaining i) variance estimates for $\hat{C}_p^j$, and ii) R bootstrap $\hat{c}_p$:

$$\hat{\sigma}_p^2(j) = \widehat{\mathrm{Var}}\, C_p^{j\,*(r)}, \tag{7}$$

$$\hat{c}_p^{*(r)} = \log_2 e \sum_{j_1}^{j_2} w_j \hat{C}_p^{j\,*(r)}. \tag{8}$$

The $\hat{\sigma}_p^2(j)$ are involved in the calculation of the weights $w_{2,j}$. The $\hat{c}_p^{*(r)}$ are used to construct $100(1-\alpha)\%$ confidence intervals for the $\hat{c}_p$s, according to:

$$\mathrm{CI}_p = \left(Q_p\left(\frac{\alpha}{2}\right), Q_p\left(1 - \frac{\alpha}{2}\right) \right) = \left(\hat{c}_p^{*(r_1)}, \hat{c}_p^{*(r_2)} \right) \tag{9}$$

Here $Q_p(\alpha)$ is the α-th empirical quantile of the empirical distribution of the R estimates of $\hat{c}_p$, i.e., $r_1 = \lfloor \frac{R\alpha}{2} \rfloor$ and $r_2 = R - r_1 + 1$. Alternatively to this *simple* bootstrap procedure, the use of pivotal statistics or variance stabilizing transformations can be considered (see e.g., [9]). This is currently being investigated.

4. WAVELET LEADERS

As indicated in Section 1, *Wavelet Leaders* consists of multiresolution quantities that present significant theoretical and practical qualities to perform Multifractal analysis. Notably, they enable the use of positive and negative q in Eq. 1 as well as a relevant analysis of chirp-type oscillating singularities and hence the correct analysis of the entire spectrum of multifractal properties of X. This has recently been proven in [4, 5]. Therefore, in the estimation procedures described in Section 2, wavelet Leaders are used instead of

wavelet coefficients.
Let us introduce the indexing $\lambda_{j,k} = [k2^j, (k+1)2^j)$ and the union $3\lambda_{j,k} = \lambda_{j,k-1} \cup \lambda_{j,k} \cup \lambda_{j,k+1}$. The wavelet Leaders $L_X(j,k)$ are defined as

$$L_X(j,k) = \sup_{\lambda' \subset 3\lambda_{j,k}} |d_{\lambda'}|, \qquad (10)$$

where the supremum is taken on the discrete wavelet coefficients $d_X(\cdot,\cdot)$ in the time neighborhood $3\lambda_{j,k}$ over all finer scales $2^{j'} < 2^j$. All relations in Subsection 2.1, in particular Eqs. (1-5) can be rewritten replacing the $d_X(j,k)$ by the $L_X(j,k)$. Thus, the estimation procedures detailed in Subsection 2.2 and in Section 3 can be rewritten, mutatis mutandis. For convenience, we introduce the superscript d and L to distinguish between estimates involving wavelet coefficients and Leaders, i.e., $\hat{C}_p^{j,\,d}$, $\hat{c}_p^d$, CI_p^d involve the $d_X(j,k)$ while $\hat{C}_p^{j,\,L}$, $\hat{c}_p^L$, CI_p^L involve the $L_X(j,k)$.

5. NUMERICAL SIMULATIONS

Monte Carlo Simulation. We evaluate the performance of the proposed estimation procedures by applying them to a large number N_{MC} of realizations of synthetic stochastic multifractal processes with known and controlled multifractal properties.
From averages $\langle\cdot\rangle$ over Monte Carlo realizations, we compute the standard deviations $\hat{s}_p = \sqrt{\langle \hat{c}_p^2(i)\rangle - \langle \hat{c}_p(i)\rangle^2}$, the biases $\hat{\beta}_p = \langle c_p(i) - \hat{c}_p(i)\rangle$ and mean-square errors $\mathrm{MSE}_p = \sqrt{\hat{s}_p^2 + \hat{\beta}_p^2}$ of the proposed estimators.
To evaluate the reliability of the confidence intervals obtained from bootstrap, we investigate the coverages produced by *re-centered confidence intervals*,

$$\mathrm{CI}_{p,R}(i) = \mathrm{CI}_p(i) - \hat{\beta}_p, \quad i = 1, \cdots, N_{MC}$$

i.e., confidence intervals that are corrected by the Monte Carlo estimates of the bias of the estimators. This allows us to determine the quality of the confidence intervals independently of the influence of a possible bias of the estimators. The *empirical coverages* of re-centered confidence intervals are then calculated as:

$$\mathcal{C}_p^{\mathrm{emp}} = \langle \varepsilon\left(c_p, \mathrm{CI}_{p,R}(i)\right)\rangle \,.$$

Here, $\varepsilon\left(c_p, \mathrm{CI}_{p,R}(i)\right) = 1$ if $c_p \in \mathrm{CI}_{p,R}(i)$ and 0 otherwise: i.e., the empirical coverages $\mathcal{C}_p^{\mathrm{emp}}$ equals the percentage of MC realizations for which the true c_p fall within the corresponding re-centered confidence intervals.
Scaling Processes. We use two well known scaling processes, Fractional Brownian motion (FBM) and Multifractal random walk (MRW), chosen because they provide us with simple yet representative examples of Gaussian monofractal processes and non Gaussian multifractal processes respectively. **FBM** is defined as the only Gaussian exactly self-similar process with stationary increments. Its full definition as well as that of self-similarity can be found in e.g., [17]. The statistical properties of FBM are entirely determined by the parameter H. FBM possesses scaling properties as in Eq. (2), with $\zeta(q) = qH$, for $q \in (-1, \infty)$. Thus, $c_1 = H$ and $c_p \equiv 0$ for all $p \geq 2$. **MRW** has been introduced in [18] as a simple multifractal (hence non Gaussian) process with stationary increments: $X(k) = \sum_{k=1}^{n} G_H(k) e^{\omega(k)}$, 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: $\mathrm{cov}(\omega(k_1), \omega(k_2)) = \lambda \ln\left(\frac{L}{|k_1-k_2|+1}\right)$ when $|k_1 - k_2| < L$ and 0 otherwise. MRW has interesting scaling properties as in Eqs. (1) or (2) for $q \in \left[-\sqrt{2/\lambda}, \sqrt{2/\lambda}\right]$ (cf. [15]), with $\zeta(q) = (H+\lambda)q - \lambda^2 q^2/2$, hence $c_1 = H + \lambda$, $c_2 = -\lambda^2$ and $c_p \equiv 0$ for all $p \geq 3$. One sees that the departure from a linear behavior in q is fully controlled by λ (or c_2).
Simulation Setup. The results presented here are obtained using Daubechies wavelets with $N_\psi = 3$. Parameters were set to $N_{MC} = 1000$, $n = 2^{15}$, $R = 200$, $L = 6$, $H = c_1 = 0.8$ for FBM and $(H,\lambda) = (0.72, \sqrt{0.08})$, i.e. $c_1 = 0.8$ and $c_2 = -0.08$, for MRW.

6. RESULTS

6.1. Statistical Performance.

Tables 1 and 2 compare the biases and MSEs (respectively) of $\hat{c}_p^d$ and $\hat{c}_p^L$ for $p = 1-5$, obtained for 1000 realizations of FBM and MRW.
Bias. Table 1 shows that, while $\hat{c}_1^d$ and $\hat{c}_1^L$ have comparable biases, for $p \geq 2$, $\hat{c}_p^L$ systematically exhibits smaller biases. Note that this discrepancy increases with p and that for $p = 5$, the difference counts at least 3 orders of magnitude ! Clearly, $\hat{c}_p^d$ become useless in practise for $p = 4, 5$, whereas $\hat{c}_p^L$ continue to give estimates of high accuracy, similar to those produced for $p = 1, 2$. Also, $\hat{c}_1^d$ has a bias smaller than that of $\hat{c}_1^L$ for (monofractal) FBM, but a larger bias for (multifractal) MRW. Together, these arguments clearly indicate that when the deviations from linearity of $\zeta(q)$ are of interest, Leaders must be preferred to wavelet coefficients. From the weight choice point of view, biases are equivalent.
Mean Square Error. Standard deviations are an order of magnitude larger than biases, so that they mostly contribute to MSEs. Hence MSEs only are reported. Table 2 shows that the MSEs of the $\hat{c}_p^L$ are systematically much smaller than those of the $\hat{c}_p^d$ for both processes, all weights and all orders. Again, this difference grows with p (roughly as 10^{p-1}), showing that estimates for the cumulant of order 3 can no longer be obtained from wavelet coefficients and require the use of Leaders. For the weight choice issue, we note that the MSEs of $\hat{c}_p^L$ based on $w_{1,j}$ and $w_{2,j}$ are of comparable order of magnitude in all cases and much smaller than those obtained with $w_{0,j}$. Four conclusions can be drawn: i) linear regressions for the cumulants must be weighted, ii) the independence assumption of the $d_X(j,k)$ underlying the choice $w_{1,j}$ remains valid for Leaders, iii) bootstrap does a good job in estimating the variances of the $\hat{C}_p^j$s, iv) as the choice $w_{2,j}$ involves a much higher computational cost for results equivalent to those obtained with the choice $w_{1,j}$, this latter is preferred. Fig. 1 displays histograms of $\hat{c}_p^d$ and $\hat{c}_p^L$ for $p = 1, 2, 3$ (weights $w_{1,j}$) obtained from 1000 realizations of MRW. Whereas the empirical distributions of $\hat{c}_p^d$ and $\hat{c}_p^L$ have similar shape and center, those of $\hat{c}_p^L$ have smaller spread, in particular for $p = 2, 3$. Again, this suggests the use of $\hat{c}_p^L$ rather than $\hat{c}_p^d$.
Conclusion. These results lead us to conclude that weighted wavelet Leader based estimators produce the most accurate estimates for log-cumulants. In particular, highly relevant estimates of log-cumulants of higher order are obtained. Hence, we recommend the use of $\hat{c}_p^L$ with weights $w_{1,j}$ for practical log-cumulant based multifractal analysis.
Block Bootstrap. From the definition of the Leaders $L_X(j,k)$ in Eq. (10), it is clear that they may display complicated intra- and inter-scale (i.e., joint time and scale) correlation structures. Therefore, one may expect poor results for leader-based simple bootstrap estimates. Our results show that this is not the case. Moreover, in Fig. 2, histograms for $\hat{c}_p^L$, based on 200 realizations of MRW (cf. below), and for 200 corresponding bootstrap estimates $\hat{c}_p^{L,*}$, based on a single realization, are shown. The closeness of Monte Carlo and bootstrap empirical distributions for $\hat{c}_p^L$ provides us with a clear

BIAS	FBM			MRW		
Estimator	$w_{0,j}$	$w_{1,j}$	$w_{2,j}$	$w_{0,j}$	$w_{1,j}$	$w_{2,j}$
$\hat{c}_1^d$	**0.4**	0.8	-3.3	32.8	31.2	28.6
$\hat{c}_1^L$	2.2	5.8	6.2	**28.3**	30.2	28.4
$\hat{c}_2^d$	10.7	1.3	48.3	**-0.1**	-10.6	33.1
$\hat{c}_2^L$	1.5	-2.8	**0.6**	-1.8	-10.0	1.0
$\hat{c}_3^d$	-51.3	-3.5	-267.9	-36.7	-6.7	-258.9
$\hat{c}_3^L$	**0.9**	1.2	1.4	2.2	**-0.9**	2.0
$\hat{c}_4^d$	394.9	54.3	1059.6	352.8	86.0	1078.0
$\hat{c}_4^L$	0.8	0.6	**0.2**	5.9	3.2	**1.7**
$\hat{c}_5^d$	-3513.8	-783.3	-3235.6	-3344.8	-1036.9	-3232.7
$\hat{c}_5^L$	**0.0**	-0.2	-0.4	7.8	4.2	**3.0**

Table 1. Bias ($\times 10^3$) of estimators $\hat{c}_p$ for FBM (left) and MRW (right) and $p = 1 - 5$. Best results are marked in bold.

MSE	FBM			MRW		
Estimator	$w_{0,j}$	$w_{1,j}$	$w_{2,j}$	$w_{0,j}$	$w_{1,j}$	$w_{2,j}$
$\hat{c}_1^d$	32.8	15.5	16.6	47.0	35.3	34.3
$\hat{c}_1^L$	19.8	**10.8**	12.0	38.1	**32.8**	32.9
$\hat{c}_2^d$	77.6	37.3	65.6	87.8	42.8	60.1
$\hat{c}_2^L$	6.5	**4.1**	4.6	24.1	**17.5**	20.6
$\hat{c}_3^d$	340.0	187.7	302.4	399.8	200.7	300.7
$\hat{c}_3^L$	3.1	**1.8**	1.9	23.4	**18.4**	19.9
$\hat{c}_4^d$	1822.4	1251.2	1125.4	2505.8	1366.4	1164.0
$\hat{c}_5^L$	2.0	1.0	**0.8**	31.0	30.0	**25.0**
$\hat{c}_5^d$	9622.5	9803.0	3621.3	15741.9	11068.3	3673.2
$\hat{c}_5^L$	2.6	0.7	**0.6**	72.9	**50.1**	52.3

Table 2. MSE ($\times 10^3$) of estimators $\hat{c}_p$ for FBM (left) and MRW (right) and $p = 1 - 5$. Best results are marked in bold.

indication in favor of a relevant use of leader-based bootstrap estimation procedures, as described in Section 3. Joint time-scale block bootstrap is however under current investigation.

6.2. Confidence Intervals.

Table 3 summarizes the empirical coverages of the re-centered confidence intervals (9) for the estimators $\hat{c}_p^d$ and $\hat{c}_p^L$ of order $p = 1-5$. The targeted coverage is 95%.
We observe that the bootstrap based procedure produces satisfactory confidence intervals in all cases. Moreover, we note that the coverages obtained with the choice $w_{1,j}$ are highly relevant and quasi systematically the best (or very close to the best). This may be because it allies relevant weighting without having recourse to estimated quantities. This yields two main conclusions. First, bootstrap approaches yield highly relevant confidence intervals for multifractal estimation. Therefore, we highly recommend their use in practical multifractal analyses. Second, the choice $w_{1,j}$ is to be favored. Improvements resulting from the use of pivotal statistics (see e.g., [9]) are under current investigation.

6.3. Practical procedure and regression range.

From these analyzes, we have designed a MATLAB routine that implement wavelet coefficient and wavelet Leader based log-cumulant multifractal analysis together with bootstrap based confidence intervals. Therefore, it enables us to obtain from a **single** observed times series with **finite length**, both estimates for the c_ps and error bars.

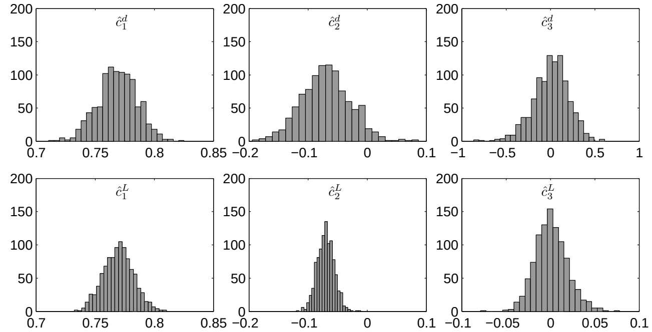

Fig. 1. Histograms of $\hat{c}_1$ (left), $\hat{c}_2$ (center) and $\hat{c}_3$ (right) for wavelet coefficients (top) and wavelet Leaders (bottom) for 1000 realizations of MRW and weights $w_{1,j}$ (note the different scale of the x-axis for $\hat{c}_3^d$ and $\hat{c}_3^L$.

Estimate	FBM			MRW		
	$w_{0,j}$	$w_{1,j}$	$w_{2,j}$	$w_{0,j}$	$w_{1,j}$	$w_{2,j}$
$\mathrm{CI}_{1,R}^d$	85.7	**92.1**	89.5	90.1	98.6	**95.5**
$\mathrm{CI}_{2,R}^d$	79.4	**92.3**	75.3	79.1	**95.5**	79.6
$\mathrm{CI}_{3,R}^d$	72.3	**91.2**	66.9	71.2	**92.6**	72.9
$\mathrm{CI}_{4,R}^d$	64.3	**87.8**	85.0	62.2	**90.3**	86.8
$\mathrm{CI}_{5,R}^d$	85.7	**87.8**	76.6	83.8	**90.8**	82.5
$\mathrm{CI}_{1,R}^L$	71.8	**83.4**	76.6	84.5	**98.8**	89.0
$\mathrm{CI}_{2,R}^L$	76.1	**90.3**	73.3	75.7	**97.0**	71.3
$\mathrm{CI}_{3,R}^L$	92.6	**94.7**	**94.7**	80.4	**96.4**	75.6
$\mathrm{CI}_{4,R}^L$	**96.3**	97.2	89.7	85.1	**95.4**	79.2
$\mathrm{CI}_{5,R}^L$	**96.1**	98.0	98.2	85.0	**96.6**	82.4

Table 3. Empirical coverage of re-centered 95% confidence interval for log-cumulant estimates from wavelet coefficients (top) and wavelet Leaders (bottom) for FBM and MRW and $p = 1 - 5$. Results closest to target coverage are marked in bold.

This significantly improves already available practical multifractal estimation procedures as for most applications error bars are as important as estimates themselves. We see this as a major result of the present contribution.
Fig. 3 illustrates this procedure at work and shows logscale diagrams (for $p = 1, 2, 3$): $\hat{C}_p^{j,\,L}$ (or $\hat{C}_p^{j,\,d}$) as a function of j for a single realization of MRW, together with corresponding regression lines and $\pm 1.96\hat{\sigma}_p(j)$ error bar estimates obtained from bootstrap. Whereas all $\hat{C}_p^{j,\,L}$s display a highly linear behaviour over a large range of scales j, even for large ps, a zone of linearity is more difficult to find for $\hat{C}_2^{j,\,d}$ and $\hat{C}_3^{j,\,d}$, suggesting estimates of poorer quality in these cases. Clearly, the selection of the regression range of octaves is a key practical issue that adds an extra complication to the analyzes reported here. Again, one sees that for Leader based multifractal, selecting the regression range should be easier.

7. CONCLUSION AND PERSPECTIVES

In the present work, we compared various log-cumulants multifractal estimation procedures. First, we demonstrated that the use of wavelet Leaders instead of wavelet coefficients brings substantial improvement in estimation performance. In particular, highly accurate estimates for log-cumulants of order $p \geq 2$ can be obtained. To the best of our knowledge, this had never been illustrated clearly before. Second, we showed that the simple bootstrap approach provides us with highly relevant confidence intervals for the estimates $\hat{c}_p^d$ and $\hat{c}_p^L$. This is another major improvement as, to the best of our

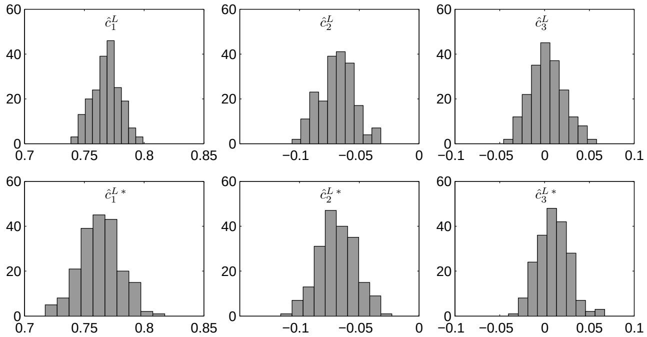

Fig. 2. Histograms of $\hat{c}_p^L$ obtained from 200 realizations of MRW (top row) and bootstrap resamples $\hat{c}_p^{L\,*}$ obtained from one single realization chosen at random (bottom row), for $p = 1, 2, 3$ (from left to right).

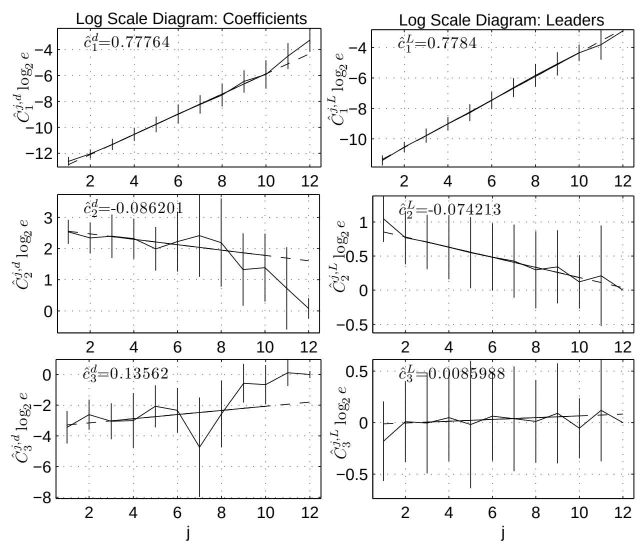

Fig. 3. Logscale diagram of $\hat{C}_p^{j,\,d}$ (left) and $\hat{C}_p^{j,\,L}$ (right) for a single realization of MRW, with weighted regression line with weights $w_{1,j}$ over scales $j_1 = 3$ to $j_2 = 10$. The error bars $\pm 1.96\hat{\sigma}_p(j)$ are estimated by bootstrap.

knowledge, this is the first time that a non parametric confidence interval estimation procedure, with excellent performance, is obtained for multifractal analysis. Wavelet Leaders and bootstrap confidence intervals together lead to the design of a multifractal analysis procedure (available in MATLAB upon request) of primary interest for the exploration of empirical data with possibly multifractal properties. A key feature of the results obtained in this work lies in the fact that they hold for both Gaussian monofractal and non Gaussian multifractal processes. This is very promising as it opens the track for the design of hypothesis tests aiming at discriminating between mono- and multi-fractal processes, and between different multi-fractal processes, two major practical issues. These ideas are currently under investigation.

8. REFERENCES

[1] S. Jaffard, “Multifractal formalism for functions, part 2: Self-similar functions,” *SIAM J. of Math. Anal.*, vol. 28, no. 4, pp. 971–998, 1997.

[2] P. Abry, P. Flandrin, M. Taqqu, and D. Veitch, “Wavelets for the analysis, estimation and synthesis of scaling data,” in *K. Park and W. Willinger Eds, Self-similar Network Traffic and Performance Evaluation*. spring 2000, Wiley.

[3] E. Bacry, J. Muzy, and A. Arneodo, “Singularity spectrum of fractal signals from wavelet analysis: Exact results,” *J. Stat. Phys.*, vol. 70, pp. 635–674, 1993.

[4] S. Jaffard, “Wavelet techniques in multifractal analysis,” in *Fractal Geometry and Applications: A Jubilee of Benoît Mandelbrot, M. Lapidus et M. van Frankenhuijsen Eds., Proceedings of Symposia in Pure Mathematics*. 2004, vol. 72(2), pp. 91–152, AMS.

[5] S. Jaffard, B. Lashermes, and P. Abry, “Wavelet leaders in multifractal analysis,” in *Wavelet Analysis and Applications*. 2005, University of Macau, China.

[6] B. Castaing, Y. Gagne, and M. Marchand, “Log-similarity for turbulent flows,” *Physica D*, vol. 68, pp. 387–400, 1993.

[7] J. Delour, J.F. Muzy, and A. Arneodo, “Intermittency of 1d velocity spatial profiles in turbulence: A magnitude cumulant analysis,” *The Euro. Phys. Jour. B*, vol. 23, pp. 243–248, 2001.

[8] B. Efron, *The Jackknife, the Bootstrap, and Other Resampling Plans*, Society for Industrial and Applied Mathematics, Philadelphia, 1982.

[9] A.M. Zoubir and D.R. Iskander, *Bootstrap Techniques for Signal Processing*, Cambridge University Press, ISBN 0-521-83127-X, 2004.

[10] A.M. Zoubir, “On confidence intervals for the coherence function,” in *Proceedings of the 30th IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Philadelphia, 2005.

[11] A.C. Davison and D.V. Hinkley, *Bootstrap methods and their application*, Cambridge series in statistical and probabilistic mathematics. Cambridge University Press, Cambridge, 1997.

[12] D.B. Percival, S. Sardy, and A.C. Davison, “Wavestrapping time series: Adaptive wavelet-based bootstrapping,” in *Nonlinear and Nonstationary Signal Processing,* W.J. Fitzgerald, R.L. Smith, A.T. Walden and P.C. Young, Eds., Cambridge, 2000, pp. 442–471, Cambridge University Press.

[13] A. Jach and P. Kokoszka, “Wavelet based confidence intervals for the self-similarity parameter,” *Journal of Statistical Computation and Simulation*, to appear.

[14] A.M. Sabatini, “Wavelet-based estimation of $1/f$-type signal parameters: Confidence intervals using the bootstrap,” *IEEE Transactions on Signal Processing*, vol. 47, no. 12, pp. 3406–3409, 1999.

[15] B. Lashermes, P. Abry, and P. Chainais, “New insight in the estimation of scaling exponents,” *Int. J. on Wavelets, Multiresolution and Information Processing*, vol. 2, no. 4, 2004.

[16] M. Kendall and A. Stuart, *The Advanced Theory of Statistics, Distribution Theory*, vol. 1, C. Griffin, 1977.

[17] G. Samorodnitsky and M. Taqqu, *Stable non-Gaussian random processes*, Chapman and Hall, New York ISBN 0-412-05171-0, 1994.

[18] E. Bacry, J. Delour, and J.F. Muzy, “Multifractal random walk,” *Phys. Rev. E*, vol. 64: 026103, 2001.