

## Bootstrap for Multifractal Analysis

Patrice Abry, Herwig Wendt

► **To cite this version:**

Patrice Abry, Herwig Wendt. Bootstrap for Multifractal Analysis. International Conference on Acoustics, Speech and Signal Processing (ICASSP), May 2006, Toulouse, France. IEEE-SPS, 2006. <ensl-00080285>

**HAL Id: ensl-00080285**

**<https://hal-ens-lyon.archives-ouvertes.fr/ensl-00080285>**

Submitted on 15 Jun 2006

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# BOOTSTRAP FOR MULTIFRACTAL ANALYSIS

Herwig Wendt, Patrice Abry

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

## ABSTRACT

Multifractal analysis, which mainly consists in estimating scaling exponents, has become a popular tool for empirical data analysis. Although widely used in different applications, the statistical performance and the reliability of the estimation procedures are still poorly known. Notably, little is known about confidence intervals, though they are of first importance in applications. The present work investigates the potential uses of bootstrap for multifractal estimation: Can bootstrap improve current estimation procedures or be used to obtain reliable confidence intervals? Comparing the statistical performance of different estimators, our major result is to show that bootstrap based procedures provide us both with accurate estimates and reliable confidence intervals. We believe that this brings substantial improvements to practical empirical multifractal analyses.

## 1. MOTIVATION

Multifractal (or scaling) analysis [1, 2, 3] is now considered as a useful and major tool for empirical data analysis. It is widely used in numerous applications of very different nature. The analysis mainly consists of measuring scaling exponents. Typically, such measurements are used to perform detection, classification or interpretation of data. Although largely used in practise, the statistical performance and reliability of the common estimation procedures remain poorly studied. For instance, little is known on the size of confidence intervals. In many applications and in real life data analysis, however, one is as much interested in the confidence intervals as in the values of the scaling exponents themselves. Indeed, without confidence intervals no classification, discrimination or hypothesis testing is possible. A popular and powerful approach to empirical multifractal analysis is based on wavelet estimation procedures [2, 3].

The aim of the present contribution is the investigation of potential benefits of the non parametric bootstrap for wavelet-based multifractal analysis. 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 [4] and has recently regained interest due to continuously growing computer facilities [5, 6, 7, 8]. Bootstrapping has been used in the wavelet domain after the pioneering work reported in [9]. It has also been considered for the estimation of the Hurst parameter of self-similar processes [10, 11].

In the present work, we extend this line of research by addressing two issues: Can the bootstrap be used to improve current estimation procedures? Can the bootstrap provide us with reliable confidence intervals? To answer these questions, we compare the statistical performance of six different estimation procedures as well as the confidence intervals designed either from asymptotic expansions and Gaussian expansions or from bootstrap approaches. Results are obtained by applying our procedures to a large number of realizations

of synthetic scaling processes with a priori known and controlled multifractal properties and show that bootstrapping enables us to obtain highly reliable confidence intervals. We end up with a practical procedure that provides us with both scaling exponent estimates and accurate confidence intervals, and that can actually be used for analyzing a single and finite length observation of empirical data.

## 2. SCALING AND MULTIFRACTAL

### 2.1. Definitions

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

The so-called structure functions  $S(j, q)$  are defined as time averages of the  $d_X(j, k)$  at fixed scale  $a = 2^j$  ( $n_j$  stands for the number of  $d_X(j, k)$  available at scale  $2^j$ , roughly  $n_j = n/2^j$ ):

$$S(j, q) = \frac{1}{n_j} \sum_{k=1}^{n_j} |d_X(j, k)|^q.$$

Processes are said to possess *scale invariance* or *scaling* properties if the  $S(j, q)$  display power law behaviors with respect to scales,

$$S(j, q) = c_q |a|^{\zeta(q)}, \quad (1)$$

over a wide range of scales,  $a \in [a_m, a_M]$ ,  $a_M/a_m \gg 1$ , and for some statistical orders  $q \in [q_*^-, q_*^+]$  (cf. [12]). The  $\zeta(q)$  are referred to as the scaling exponents. When  $\zeta(q) = qH$ ,  $X$  is said to be monofractal. For instance, this is 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 obviously a *poor* or *operational* definition of multifractal processes. However, this is sufficient for the purposes of this article. For a thorough introduction to multifractal analysis, the reader is referred to, e.g., [1].

### 2.2. Estimation Procedures

**Definition.** The scaling exponents  $\zeta(q)$  are practically estimated by performing weighted linear regressions in  $\log_2 S(j, q)$  versus  $\log_2 a = j$  diagrams:

$$\hat{\zeta}(q) = \sum_{j=j_1}^{j_2} w_{j,q} (\log_2 S(j, q) - g(j, q)), \quad (2)$$

where the  $g(j, q)$  are deterministic constants defined below. The finite range of scales within which regressions are performed is defined as  $a_m = 2^{j_1}$  and  $a_M = 2^{j_2}$ .

**Statistical performance.** A straightforward computation of the bias of  $\hat{\zeta}(q)$  implies that:

$$g(j, q) = \mathbb{E} \log_2 S(j, q) - \log_2 \mathbb{E} S(j, q). \quad (3)$$

Furthermore, as wavelet coefficients at different scales are weakly correlated, one can approximate the variance of  $\hat{\zeta}(q)$  as:  $\text{Var} \hat{\zeta}(q) \simeq \sum_{j=j_1}^{j_2} w_{j,q}^2 \text{Var} \log_2 S(j, q)$ . Hence, let us define

$$V(q) \equiv \sum_{j=j_1}^{j_2} w_{j,q}^2 \sigma^2(j, q) \text{ and } \sigma^2(j, q) = \text{Var} \log_2 S(j, q). \quad (4)$$

**Weights.** The weights have to satisfy the usual constraints  $\sum_{j_1}^{j_2} j w_{j,q} \equiv 1$  and  $\sum_{j_1}^{j_2} w_{j,q} \equiv 0$ . They are written as  $w_{j,q} = \frac{1}{b_{j,q}} \frac{S_{0j} - S_1}{S_0 S_2 - S_1^2}$ , where  $S_p = \sum_{j_1}^{j_2} j^p / b_{j,q}$ ,  $p = 0, 1, 2$ . The  $b_{j,q}$ s are arbitrary positive numbers reflecting the confidence granted to  $\log_2 S(j, q)$ . As thoroughly detailed in, e.g., [2], the discrete wavelet coefficients  $d_X(j, k)$  of scaling processes are weakly correlated. Therefore, the variance of  $S(j, q)$  and hence of  $\log_2 S(j, q)$  decreases approximately as  $1/n_j$ . This naturally leads to the choice  $b_{j,q} = 1/n_j = 2^j/n$ . Note that with this choice, valid for the remainder of the present work, the  $b_{j,q}$ s and  $w_{j,q}$ s no longer depend on  $q$ .

### 3. ASYMPTOTIC EXPANSIONS VERSUS BOOTSTRAP

Our goal is to use the procedure described above to obtain, from a single finite length observation of  $X$ , reliable estimates of  $\zeta(q)$ , as well as reliable confidence intervals for these estimates. This involves calculating or estimating the bias (3) and the distribution of the estimator (2).

The bias  $g(j, q)$  can be estimated by using either asymptotic or Gaussian expansions or the bootstrap. Confidence intervals can be obtained either by assuming  $\hat{\zeta}(q)$  to be Gaussian and using the approximation (4), with estimates for  $\sigma^2(j, q)$ , or by bootstrapping  $\hat{\zeta}(q)$ .

#### 3.1. Asymptotic Expansions

Standard approximation formulae for a change of variable  $Y = f(X)$  give  $\mathbb{E} f(X) \simeq f(\mathbb{E} X) + f''(\mathbb{E} X) \frac{\sigma_X^2}{2}$  and  $\sigma_Y^2 \simeq |f'(\mathbb{E} X)|^2 \sigma_X^2$  where  $\sigma_X^2$  and  $\sigma_Y^2$  denote the variances of  $X$  and  $Y$  respectively. As  $\text{Var} S(j, q)$  decreases as  $1/n_j$  in the limit of large  $n_j$ , we can apply those formulae to  $\log_2 S(j, q)$ . Tedious calculations [2] lead to:

$$\begin{aligned} g(j, q) &\simeq -\frac{\log_2 e}{2} \frac{\text{Var} S(j, q)}{(\mathbb{E} S(j, q))^2} \simeq -\frac{(\log_2 e)}{2n_j} \frac{\text{Var} |d_X(j, \cdot)|^q}{(\mathbb{E} |d_X(j, \cdot)|^q)^2} \\ \sigma^2(j, q) &\simeq (\log_2 e)^2 \frac{\text{Var} S(j, q)}{(\mathbb{E} S(j, q))^2} \simeq \frac{(\log_2 e)^2}{n_j} \frac{\text{Var} |d_X(j, \cdot)|^q}{(\mathbb{E} |d_X(j, \cdot)|^q)^2}. \end{aligned}$$

The second approximation involves the assumption that wavelet coefficients are quasi-uncorrelated, hence  $\text{Var} S(j, q) \simeq \frac{\text{Var} |d_X(j, \cdot)|^q}{n_j}$ . This asymptotic expansion leads to defining the estimates:

$$\begin{aligned} \hat{g}_A(j, q) &= -\frac{\log_2 e}{2n_j} \frac{\widehat{\text{Var}} |d_X(j, \cdot)|^q}{(\widehat{\mathbb{E}} |d_X(j, \cdot)|^q)^2}, \\ \hat{\sigma}_A^2(j, q) &= \frac{(\log_2 e)^2}{n_j} \frac{\widehat{\text{Var}} |d_X(j, \cdot)|^q}{(\widehat{\mathbb{E}} |d_X(j, \cdot)|^q)^2}, \end{aligned} \quad (5)$$

where  $\widehat{\mathbb{E}} \cdot$  and  $\widehat{\text{Var}} \cdot$  stand for the standard sample mean and sample variance estimators, respectively.

**Gaussian expansions.** For the cases where the wavelet coefficients are Gaussian, the ratio  $\text{Var} |d_X(j, \cdot)|^q / (\mathbb{E} |d_X(j, \cdot)|^q)^2$  can be computed analytically, yielding (note that no quantity needs to be estimated in this case):

$$\begin{aligned} \hat{g}_G(j, q) &= -\frac{\log_2 e}{2n_j} \left( \sqrt{\pi} \Gamma \left( q + \frac{1}{2} \right) / \Gamma \left( \frac{q+1}{2} \right)^2 - 1 \right) \\ \hat{\sigma}_G^2(j, q) &= \frac{(\log_2 e)^2}{n_j} \left( \sqrt{\pi} \Gamma \left( q + \frac{1}{2} \right) / \Gamma \left( \frac{q+1}{2} \right)^2 - 1 \right). \end{aligned} \quad (6)$$

#### 3.2. The Non Parametric Bootstrap Approach

Rather than using asymptotic or Gaussian expansions, the desired characteristics of  $S(j, q)$  and  $\hat{\zeta}(q)$  can be obtained through bootstrap generated approximate empirical distributions [4, 5, 6, 7, 8]. Since wavelet coefficients at each given scale are weakly correlated, we use a block bootstrap with overlapping blocks of length  $L$ . At each scale  $j$ ,  $P$  bootstrap resamples  $\mathcal{D}_j^{*(1)}, \dots, \mathcal{D}_j^{*(P)}$  are generated, where each resample

$$\mathcal{D}_j^* = \left\{ d_X^{*(1)}(j, \cdot), \dots, d_X^{*(n_j)}(j, \cdot) \right\}$$

is an unsorted collection of  $n_j$  sample points that are drawn block-wise with replacement from the original sample  $\mathcal{D}_j = \{d_X(j, 1), \dots, d_X(j, n_j)\}$ . From these collections  $\mathcal{D}_j^{*(p)}$ , the  $P$  bootstrap quantities

$$p = 1, \dots, P, \quad S^{*(p)}(j, q) = \frac{1}{n_j} \sum_{k=1}^{n_j} |d_X^{*(p)}(j, k)|^q$$

are calculated, yielding the estimates

$$\begin{aligned} \hat{g}_B(j, q) &= \widehat{\mathbb{E}} \log_2 S^*(j, q) - \log_2 \widehat{\mathbb{E}} S^*(j, q) \\ \hat{\sigma}_B^2(j, q) &= \widehat{\text{Var}} \log_2 S^*(j, q). \end{aligned} \quad (7)$$

#### 3.3. Estimators

At this stage, we gathered material to define 2\*3 different estimators, (2 different  $S(j, q)$ , standard or bootstrapped, times 3 different bias corrections, asymptotic (5), Gaussian (6) or bootstrap (7)):

$$\begin{aligned} \hat{\zeta}_X(q) &= \sum_{j_1}^{j_2} w_j (\log_2 S(j, q) - \hat{g}_X(j, q)) \\ \hat{\zeta}_X^B(q) &= \widehat{\mathbb{E}} \hat{\zeta}_X^*(q) \end{aligned} \quad (8)$$

with  $X = \{G, B, A\}$ . Here,  $\hat{\zeta}_X^B(q)$  is seen as a modification of (2), and is defined as the sample mean of the bootstrap quantities

$$\hat{\zeta}_X^{*(p)}(q) = \sum_{j_1}^{j_2} w_j \left( \log_2 S^{*(p)}(j, q) - \hat{g}_X(j, q) \right).$$

#### 3.4. Confidence Intervals

Plugging the estimates (5), (6) or (7) for  $\hat{\sigma}_X^2(j, q)$  in the approximate variance expression (4), the  $100(1 - \alpha)\%$  asymptotic confidence intervals for each estimator  $\hat{\zeta}(q)$  in (8) are defined as

$$\text{CI}_{V_X}(q) = \left( \hat{\zeta}(q) + p \left( \frac{\alpha}{2} \right) \sqrt{V_X(q)}, \hat{\zeta}(q) + p \left( 1 - \frac{\alpha}{2} \right) \sqrt{V_X(q)} \right) \quad (9)$$

with  $X = \{G, B, A\}$ . Here,  $p(\alpha)$  is the  $\alpha$ -th quantile of the standard normal distribution.

The  $100(1 - \alpha)\%$  bootstrap confidence interval for  $\hat{\zeta}(q)$  is

$$\text{CI}_B(q) = \left( p\left(\frac{\alpha}{2}\right), p\left(1 - \frac{\alpha}{2}\right) \right) = \left( \hat{\zeta}^{*(p_1)}(q), \hat{\zeta}^{*(p_2)}(q) \right), \quad (10)$$

where  $p(\alpha)$  is the  $\alpha$ -th empirical quantile of the empirical distribution of the  $P$  estimators  $\hat{\zeta}^*(q)$ , i.e.,  $p_1 = \lfloor \frac{P\alpha}{2} \rfloor$  and  $p_2 = P - p_1 + 1$ . Alternative bootstrap confidence intervals, using pivotal statistics and variance stabilizing transformations, can be considered [6].

## 4. SIMULATION STUDY AND RESULTS

### 4.1. Monte Carlo Simulation

The performance of the proposed estimation procedures are assessed by applying them to a large number  $N_{MC}$  of realizations of synthetic stochastic multifractal processes with known and controlled scaling properties. We compare the biases  $\hat{\beta}(q) = \frac{\mathbb{E}\hat{\zeta}(q) - \zeta(q)}{q}$ , standard deviations  $\hat{s}(q) = \sqrt{\text{Var}\hat{\zeta}(q)}/q$ , and mean-square errors  $\text{MSE}(q) = \sqrt{\hat{s}(q)^2 + \hat{\beta}(q)^2}$  of the estimators (8).

The reliability of the confidence intervals (9) and (10) is studied via their empirical coverages:

$$C_{\text{emp}}(q) = \frac{\sum_{i=1}^{N_{MC}} \varepsilon(\zeta(q), \text{CI}_i(q))}{N_{MC}},$$

where  $\varepsilon(\zeta(q), \text{CI}(q)) = 1$  if  $\zeta(q) \in \text{CI}(q)$  and 0 otherwise: i.e., the empirical coverage corresponds to the percentage of MC realizations for which the true  $\zeta(q)$  falls within the computed confidence interval.

### 4.2. Scaling Processes

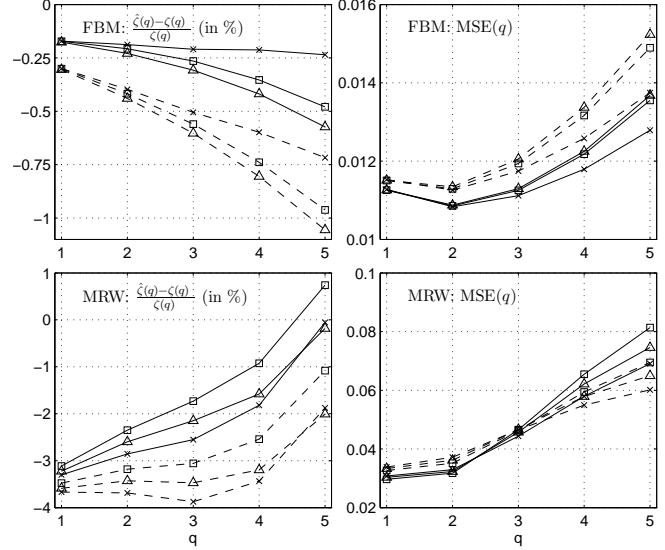
We use two scaling processes that are chosen as simple yet representative examples of Gaussian monofractal processes and non Gaussian multifractal processes respectively.

**Fractional Brownian motion (FBM)** is defined as the only Gaussian self-similar process with stationary increments [13]. Its finite dimensional distributions satisfy  $X(t) \stackrel{\text{fd}}{=} a^H X(t/a)$  for all  $a > 0$ . Its properties are entirely determined by the parameter  $H$ , and it possesses scaling properties as in Eq. (1) for  $q \in [0, \infty)$ , with  $\zeta(q) = qH$ .

**Multifractal random walk (MRW)** has been introduced in [14] as a non Gaussian processes 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  $\omega$  is independent of  $G_H$ , Gaussian, with non trivial covariance:  $\text{cov}(\omega(k_1), \omega(k_2)) = c_2 \ln\left(\frac{L}{|k_1 - k_2| + 1}\right)$  when  $|k_1 - k_2| < L$  and 0 otherwise. MRW has scaling properties as in Eq. (1) for  $q \in [0, \sqrt{2/c_2}]$ , with  $\zeta(q) = (H + c_2)q - c_2 q^2/2$ . One sees that the departure from a linear behavior in  $q$  is fully controlled by  $c_2$ .

### 4.3. Results

**Simulation setup.** The results presented here were obtained using Daubechies wavelets with  $N = 3$ . Parameters were set to  $N_{MC} = 3000$ ,  $n = 2^{15}$ ,  $P = 200$ ,  $L = 6$ ,  $H = 0.8$  for FBM and  $(H, c_2) = (0.75, 0.08)$  for MRW.



**Fig. 1.** Relative bias (left column) and MSE (right column) of estimators  $\hat{\zeta}(q)$  for FBM (top row) and MRW (bottom row) for  $q \in \{1, 2, 3, 4, 5\}$ . Solid lines correspond to  $\hat{\zeta}_X(q)$ , dashed lines to  $\hat{\zeta}_X^B(q)$ , the symbols ( $\times, \square, \triangle$ ) to  $X = (G, B, A)$ .

**Statistical performance.** Fig. 1 (left column) shows relative biases. We first note that, for all  $X \in \{G, B, A\}$ ,  $\hat{\zeta}_X(q) \geq \hat{\zeta}_X^B(q)$ , and also that the biases are of comparable orders of magnitude for all estimators. Still, we observe that, for FBM, the bias of  $\hat{\zeta}_G(q)$  is the closest to 0, while for MRW,  $\hat{\zeta}_G^B(q)$  exhibits the smallest bias. Therefore, for both Gaussian and non Gaussian processes, the Gaussian based bias correction (cf. Eq. (6)) works best.

Fig. 1 (right column) shows MSEs. Because biases are small compared to standard deviations (roughly 25%), MSEs mostly reflect standard deviations (which are hence not shown here). Let us first note that, for most cases, the MSEs of the  $\hat{\zeta}_X(q)$  are smaller than those of the  $\hat{\zeta}_G^B(q)$  for FBM, while this is the converse for MRW. More precisely, we observe that the MSEs of  $\hat{\zeta}_G(q)$  and  $\hat{\zeta}_G^B(q)$  are systematically the two lowest ones, with a minimum for  $\hat{\zeta}_G(q)$  in the case of FBM and for  $\hat{\zeta}_G^B(q)$  in the case of MRW.

These analyses lead us to recommend the use of estimators with Gaussian based bias correction,  $\hat{\zeta}_G(q)$  and  $\hat{\zeta}_G^B(q)$ , for both Gaussian and non Gaussian processes. The reasons why  $\hat{\zeta}_G^B(q)$  performs slightly better for non Gaussian processes are being further inquired.

**Confidence intervals.** Table 1 summarizes the empirical coverages obtained from the confidence intervals (9) and (10) for all estimators in (8) (the targeted coverage is 95%). Here, we report only results for  $q = 2$ , (as it can be classically related to standard spectral estimation (see e.g. [2])) and  $q = 5$  (as it constitutes a difficult estimation task). In the case of FBM, for  $q = 2$ ,  $\text{CI}_{V_B}$  has best coverage while for  $q = 5$ ,  $\text{CI}_{V_G}$  performs slightly better ( $\text{CI}_{V_B}$  remains comparable however). For both cases, the bootstrap confidence intervals  $\text{CI}_B$  produce slightly lower coverages, but stay nonetheless highly relevant.

In the case of MRW, the  $\text{CI}_B$ s produce the best (and excellent) coverages, even for  $q = 5$  (cf. Table 1). The  $\text{CI}_{V_B}$ s display similar accuracy, while, in contrast, the empirical coverages of the Gaussian based  $\text{CI}_{V_G}$ s diminish dramatically (below 45%).

| q | X | $\hat{\zeta}_X(q)$ |                             |                             |                             | $\hat{\zeta}_X^B(q)$        |                             |                             |
|---|---|--------------------|-----------------------------|-----------------------------|-----------------------------|-----------------------------|-----------------------------|-----------------------------|
|   |   | CI <sub>B</sub>    | CI <sub>V<sub>G</sub></sub> | CI <sub>V<sub>B</sub></sub> | CI <sub>V<sub>A</sub></sub> | CI <sub>V<sub>G</sub></sub> | CI <sub>V<sub>B</sub></sub> | CI <sub>V<sub>A</sub></sub> |
| 2 | G | 88.1               | 86.8                        | <b>90.6</b>                 | 85.9                        | 86.3                        | 89.0                        | 85.4                        |
|   | B | 87.9               | 86.8                        | 90.3                        | 86.0                        | 86.2                        | 88.9                        | 85.5                        |
|   | A | 87.9               | 86.9                        | 90.4                        | 86.0                        | 85.9                        | 88.8                        | 85.3                        |
| 5 | G | 84.6               | <b>90.4</b>                 | 89.4                        | 85.8                        | 87.5                        | 86.6                        | 82.5                        |
|   | B | 79.9               | 88.3                        | 87.1                        | 83.6                        | 84.7                        | 82.3                        | 78.3                        |
|   | A | 79.1               | 88.0                        | 86.8                        | 82.8                        | 83.4                        | 81.5                        | 76.6                        |
| 2 | G | 82.2               | 35.2                        | 89.8                        | 74.1                        | 28.7                        | 84.4                        | 66.5                        |
|   | B | 84.9               | 38.7                        | <b>90.9</b>                 | 76.3                        | 32.1                        | 87.2                        | 69.8                        |
|   | A | 83.6               | 37.3                        | 90.4                        | 75.5                        | 30.9                        | 85.8                        | 68.5                        |
| 5 | G | <b>92.1</b>        | 26.2                        | 86.3                        | 73.9                        | 28.3                        | 91.0                        | 79.9                        |
|   | B | 88.9               | 24.2                        | 81.2                        | 67.1                        | 26.0                        | 87.6                        | 74.6                        |
|   | A | 90.9               | 24.3                        | 84.1                        | 70.9                        | 26.8                        | 89.7                        | 76.6                        |

**Table 1.** Empirical coverage (in %) of nominal 95% confidence intervals for FBM (top part) and MRW (bottom part).

We conclude that the bootstrap based confidence intervals, CI<sub>B</sub> and CI<sub>V<sub>B</sub></sub>, are highly relevant for both Gaussian and non Gaussian processes. Therefore, we strongly recommend their use as they involve no a priori on Gaussianity and hence on mono- versus multi-fractality.

**Additional results.** Complementary Tables and Figures, not reported here for space reasons but available at [perso.ens-lyon.fr/patrice.abry/MYWEB/ICASSP06\\_results.pdf](http://perso.ens-lyon.fr/patrice.abry/MYWEB/ICASSP06_results.pdf), show that the analyses and conclusions reported above for statistical performance and confidence intervals of the estimators hold for various values of  $q$  and various sample lengths  $n = \{2^9, 2^{12}, 2^{15}\}$ . We also investigated the impact of varying the number of bootstrap resamples  $P = \{100, 200, 1000\}$ . We found empirically that  $P = 200$  is sufficient with respect to the 95% confidence interval target, and that increasing  $P$  does not bring significant improvements while having much higher computational costs. Furthermore, we explored various block lengths  $L = \{1, 4, 6, 12\}$  and found empirically that the optimal choice is  $L = 6$ . A possible explanation for this fact can be found in the weak correlation of the wavelet coefficients at each given scale, which remains significant only over the time support of the wavelet, in our case  $2N = 6$ . This is being further investigated. Reports for numerical simulations related to varying  $P$  and  $L$  are available at the URL mentioned above.

## 5. CONCLUSIONS AND PERSPECTIVES

We compared various scaling exponent estimation procedures involving asymptotic and Gaussian expansions and the bootstrap. We showed that Gaussian based bias corrected  $\hat{\zeta}_G(q)$  and  $\hat{\zeta}_G^B(q)$  display the best statistical performance. The difference in performance being weak, we conclude that the bootstrap based modification  $\hat{\zeta}_G^B(q)$  of the standard  $\hat{\zeta}_G(q)$  does not bring significant improvements. However, we showed that bootstrap approaches are providing us with highly relevant confidence intervals, CI<sub>B</sub> and CI<sub>V<sub>B</sub></sub>. This is the main benefit of bootstrapping in scaling exponent estimation as, to the best of our knowledge, this is the first time that a non parametric confidence interval estimation procedure, with excellent performance, is obtained. This constitutes a major progress in applied multifractal and scaling analysis as it provides the user not only with (scaling exponent) estimates but also with accuracy measures. Further improvements such as variance stabilization or the use of pivotal statistics [6] are under current analysis. Preliminary attempts are promising.

Furthermore, a wavelet based scaling exponent estimation procedure  $\hat{\zeta}_G^B(q)$  with bootstrap based confidence intervals, that can be applied to a single finite duration observation of empirical data, has been implemented in MATLAB and is available upon request.

A key point in 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 analytically derived results can not be obtained for multifractal processes. This leads to the idea of extending this bootstrap based approach to other multifractal quantities such as wavelet log-cumulants or wavelet leaders. In addition, this opens the track for the design of a hypothesis test aiming at discriminating between mono- and multi-fractal processes, a major practical issue. These ideas are under current work.

## 6. 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 *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] B. Efron, *The Jackknife, the Bootstrap, and Other Resampling Plans*, Society for Industrial and Applied Mathematics, Philadelphia, 1982.
- [5] A.M. Zoubir, "Multiple bootstrap tests and their application," in *Proc. of the IEEE Int. Conf. on Acoustics, Speech and Sig. Proc.*, Adelaide, 1994.
- [6] A.M. Zoubir and D.R. Iskander, *Bootstrap Techniques for Signal Processing*, Cambridge University Press, 2004.
- [7] A.M. Zoubir, "On confidence intervals for the coherence function," in *Proc. of the 30th IEEE Int. Conf. on Acoustics, Speech and Sig. Proc.*, Philadelphia, 2005.
- [8] A.C. Davison and D.V. Hinkley, *Bootstrap methods and their application*, Cambridge series in statistical and probabilistic mathematics. Cambridge University Press, Cambridge, 1997.
- [9] D.B. Percival, S. Sardy, and A.C. Davison, "Wavestrapping time series: Adaptive wavelet-based bootstrapping," in *Non-linear 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.
- [10] A. Jach and P. Kokoszka, "Wavelet based confidence intervals for the self-similarity parameter," *Journal of Statistical Computation and Simulation*, to appear.
- [11] A.M. Sabatini, "Wavelet-based estimation of 1/f-type signal parameters: Confidence intervals using the bootstrap," in *IEEE Transactions on Signal Processing*, vol. 47, no. 12, pp. 3406–3409, 1999.
- [12] 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.
- [13] G. Samorodnitsky and M. Taqqu, *Stable non-Gaussian random processes*, Chapman and Hall, New York, 1994.
- [14] E. Bacry, J. Delour, and J.F. Muzy, "Multifractal random walk," *Phys. Rev. E*, vol. 64, pp. 026103, 2001.