Contributions of Wavelet Leaders and Bootstrap to Multifractal Analysis: Images, Estimation Performance, Dependence Structure and Vanishing Moments. Confidence Intervals and Hypothesis Tests.

Herwig Wendt

N° d'ordre : 474
N° attribué par la bibliothèque : 07ENSL0 474

THÈSE

en vue d'obtenir le grade de

Docteur de l'Université de Lyon - École Normale Supérieure de Lyon
Spécialité : Traitement du Signal - Physique

Laboratoire de Physique
École doctorale de Physique

Présentée et soutenue publiquement le 23/09/08 par

HERWIG WENDT

Contributions of Wavelet Leaders and Bootstrap to Multifractal Analysis: Images, Estimation Performance, Dependence Structure and Vanishing Moments. Confidence Intervals and Hypothesis Tests.

Devant la commission d'examen formée de :

PATRICE ABRY	Membre, Directeur de thèse
Directeur de Recherche, CNRS	
STÉPHANE JAFFARD	Membre
Professeur, Université de Paris 12	
JEAN-FRANÇOIS PINTON	Membre, Président
Directeur de Recherche, CNRS	
FRANÇOIS G. SCHMITT	Membre, Rapporteur
Directeur de Recherche, CNRS	
JEAN-YVES TOURNERET	Membre, Rapporteur
Professeur, Institut National Polytechnique de Toulouse	
ABDELHAK ZOUBIR	Membre, Rapporteur
Professeur, Technische Universität Darmstadt	

Acknowledgements

First and foremost, I would like to thank Patrice Abry, who supervised this thesis work during the last three years, and who made this exciting adventure possible in the first place. I highly appreciated his commitment, enthusiasm and generosity, and I profited from his scientific excellence and guidance. I count myself very lucky for had the opportunity of working with him, and even more so for the personal side of this encounter.

I would like to thank Stéphane Jaffard, Jean-François Pinton, François Schmitt, Jean-Yves Tourneret and Abdelhak Zoubir for having given me the honor of judging my thesis work. I owe particular gratitude to François Schmitt, Jean-Yves Tourneret and Abdelhak Zoubir for having accepted the burden of evaluating this thesis manuscript, and for their thorough work. Also, I would like to thank Jean-Yves Tourneret for valuable comments and suggestions for the final version of this thesis.

This thesis would not be what it is without Stéphane Jaffard, to whom I owe a lot for the fruitful collaboration and for the stimulating scientific exchange, without which I would have a different view of applied functional analysis.

What made these three years such a good time was the fact that I had the chance of being part of an exceptional research team. I would like to thank all the members of SiSyPhe for their openness and availability that made my stay so enjoyable, both scientifically and personally. Thank you Patrice, Pierre, Bernard, Laurent, Patrick, Nicolas, Edmundo, Gabriel, Stéphane, Antoine, Béatrice and Jun for this wonderful experience.

Thanks go also to all the other members of the Physics department who all contribute to this great scientific environment and its favorable ambiance, and to Laurence and Nadine for their kindness and aptitude to master administrative trouble with a smile. I am also grateful for the trust of the members of the teaching department, who offered me the opportunity of teaching within the prestigious French Agrégation de Physique.

My thesis would have been much less comprehensive without the ongoing and past collaborations. Thanks to Sophie Achard, Philippe Ciuciu, Vladas Pipiras and Béatrice Vedel for this enriching experience, both scientifically and personally. I would also like to express my gratitude towards Alfred Hero for his supervision during my stay at University of Michigan.

It would be absurd to try to express my gratitude towards Faustine for all that she did and is, and for all the support, patience and understanding during the last three years. Being far from home means that I did not see my family as much as I would have wished. I would like to thank my father, my mother and my brother for their support and understanding over all these years.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

In many domains, scientists and engineers are used to conducting the study of systems or signals with respect to some *characteristic scales* of time (or space, energy, ...). Such scales serve as a reference for characterizing the properties of the signal or system. For instance, a physicist is often naturally confronted with a characteristic scale of time (a period, for example) or space (the size of a structure), and a chemist usually has natural reference scales of time (reaction time, equilibrium time) and space (length of chemical bonds, crystalline patterns or polymer chains). Similarly, in signal analysis, one commonly aims at identifying a correlation length and works with samples several correlation lengths apart, which are considered to be independent.

Scale invariance implies to adopt exactly the opposite perspective. The notion of scale invariance, or scaling, refers to signals or systems for which no scale can be singled out or identified to play a characteristic role. Equivalently, this amounts to saying that all scales are of equal importance, or that all scales play an equivalent role. From a signal analysis point of view, this implies a major change of paradigm: Rather than identifying scales playing a characteristic role, one has to identify and *characterize the mechanisms relating different scales*.

Scale invariant signals have been involved in a large variety of applications and in the analysis of data of very different nature, ranging from natural phenomena — physics (hydrodynamic turbulence [75, 121, 128], statistical physics [28, 98], roughness of surfaces [147]), biology (human heart beat rhythms [114, 116], physiological signals or images), geology (fault repartition [71]) — to human activities — computer network traffic [135, 143], texture image analysis [53, 96, 147], population geographical repartition [72], social behaviors or financial markets [124], to name but a few.

Scale invariance is intimately tied to *power laws*, and practical scaling analysis mostly amounts to detecting such power laws and measuring the exponents that characterize them. Historically, it has been associated with invariance to dilatation of the expected average power $\Gamma(\nu)$ of a signal, which implies the absence of a characteristic frequency ν in the signal. The power spectral density then has to verify $\frac{\Gamma(\nu_2)}{\Gamma(\nu_1)} = f\left(\frac{\nu_2}{\nu_1}\right), \quad \forall \nu_1, \nu_2,$ and therefore, $f\left(\frac{\nu_3}{\nu_1}\right) = f\left(\frac{\nu_3}{\nu_2}\right) \cdot f\left(\frac{\nu_2}{\nu_1}\right)$. It follows for 1d signals that f and consequently also Γ are power laws $\Gamma(\nu) \sim C|\nu|^{-\gamma}$. Therefore, practically, scale invariance has been defined and analyzed through the power law behavior of standard spectrum estimates

$\hat{\Gamma}(\nu)$.

The first major *model* for scale invariance is *fractional Brownian motion* (fBm) $B_H(t)$. This process is Gaussian and *self-similar*, which means that its (finite dimensional) distributions are exactly identical to those of a dilated and rescaled version $a^{-H}B_H(at)$ for all $a > 0$. As a result, the process and its dilated versions are statistically indistinguishable. These scale invariance properties are fully controlled by the single parameter H, called the Hurst parameter or self-similarity parameter. Due to its simplicity – Gaussian and entirely controlled by one single parameter – fBm became an attractive and commonly used model for stochastic scale invariance.
By definition, fBm is a nonstationary process and is long range dependent when $H > 0.5$. This poses major theoretical and practical difficulties for (spectrum based) interpretation and analysis. In their celebrated and seminal works, Flandrin [70] and Tewfik [166] showed that the coefficients of the dyadic wavelet transform of fBm form stationary sequences and are practically uncorrelated as long as the analyzing wavelet has sufficient vanishing moments $N_\psi \geq H + \frac{1}{2}$, while exactly reproducing the scale invariance properties of the process through power law behaviors across scales of their moments that are entirely controlled by H. Due to these key properties, the dyadic wavelet transform became the predominant scaling analysis tool and has since been used massively in applications for the estimation of the self-similarity parameter H.

Fractional Brownian motion constitutes a very satisfactory model for scale invariance, since it connects *all* statistical moments of dilated versions for *any* positive dilation factor a by one single parameter H. This has, however, soon been noticed to be too restricting in many situations, since it does often not match scale invariance observed in data, hence the need for richer models. Therefore, in practice, one often prefers the second major model, constituted by the more general and useful class of *multiplicative cascade processes*. These processes are non Gaussian and feature more flexible scale invariance properties since, for a restricted range of dilation factors and a finite range of statistical orders, the moments of dilated versions of the process are related through power laws with different exponents. This deeply connects to their *multifractal properties*. They are characterized by a dense and rich variety of fractal distributions of local singularities in their trajectories[1], which constitutes a strong link between scaling analysis and *multifractal analysis*.
The change in model induces two major variations in analysis: First, accounting for the extra flexibility provided by multiplicative cascade processes, the single parameter H must be replaced with a *whole collection of parameters*, called the multifractal attributes. Second, recent theoretical results show that for a relevant and accurate analysis of the scaling and multifractal properties of data, wavelet coefficients need to be replaced with new multiresolution quantities, called *wavelet Leaders* [89, 92]. At the outset of this work, the first and only practical implementation of wavelet Leaders was obtained for 1d signals by Lashermes in [8, 107, 110]. He demonstrated that practical 1d wavelet Leaders preserve key theoretical functional analysis properties and considered them for the analysis of hydrodynamic turbulence data. Yet, a number of key questions remained unanswered.

Therefore, the first main goal of the present work is to study the properties of wavelet Leader based multifractal analysis. This constitutes the core material of Part I of this manuscript. To this end, we introduce in Chapter 2 the state of the art of scaling and

[1] In contrast, the trajectories of fBm have identical locally singular behavior everywhere.

multifractal analysis, the relevant mathematical definitions and analysis tools, estimation procedures and common model processes. Then, we investigate in Chapters 3 to 5 a number of central statistical issues that have never been addressed so far: *How do wavelet Leaders actually work in practice, and what are the practical implications of finite size effects with respect to their (infinite resolution) theoretical definition? What is the form of their marginal or bivariate distribution?* These questions concern primary properties of wavelet Leaders and are accounted for in Chapter 3. The main conclusions are that meaningful Leaders can be practically obtained, but their precise analytical statistical characterization remains limited. Therefore, numerical simulations constitute the core methodology for the study of estimation procedures: — *What statistical performance can be achieved in practice, and how does it compare to those of wavelet coefficients?* This central question is investigated in Chapter 4.1. It leads to the main conclusion that wavelet Leaders do have significantly superior performance than wavelet coefficients and are hence to be preferred in practice. Moreover, we study two effects of major importance in applications: — *What are the origins of the so-called linearization effect in practical estimation?* (Chapter 5.2) — *Are estimations robust with respect to quantization of data?* (Chapter 5.3). The characterization of the properties of wavelet Leaders and wavelet Leader based analysis in Chapters 3 to 5 constitutes one of the contributions of the present work.

The second main goal of the present work is the extension of wavelet Leader based multifractal analysis to higher dimensional signals. Nowadays, in a large number of applications of very different natures, the data collected by sensors for analysis consist of images, i.e., are naturally bi-dimensional signals. This is mostly due to the recent and significant progresses achieved in digital sensor, fast rate and high resolution camera and video camera design. For a number of these applications the corresponding statistical analysis of the images amounts to performing texture characterization. This is the case notably for clouds or rainfalls analyses in geophysics [147, 155], bio-medical diagnosis for human body rhythms or structure (bones, tissues, mammography,...) [93, 116, 158], universe or galaxy structures in astronomy [162], growth phenomena in physics [57, 142] or texture classification in computer vision [136, 185], to name but a few examples. There is an increasing number of research articles which suggest that texture characterization should be conducted within the mathematical framework of scaling or multifractal analysis. However, practical tools with reasonable calculation cost and tight mathematical support are lacking, and important theoretical questions related with the multifractal analysis of textures have never been clearly stated before: *Can wavelet Leaders be practically extended to the analysis of higher dimensional signals, and what are the statistical performance?* (Chapter 2.5 and Chapter 4.2) — *Can the analysis of texture images (measures) be conducted in the same framework as the analysis of processes? How can wavelet Leader based analysis be modified to address this difficulty? What are the theoretical and practical implications of such a modification?* (Chapter 5.1). The operational definition and the implementation and validation of a 2d wavelet Leader analysis, the general theoretical and practical investigation of related difficulties and the proposed practical solutions and tools to characterize and overcome them represent another contribution of this work. It leads to the overall conclusion that wavelet Leader based multifractal analysis of images is practically feasible and exhibits satisfactory performance. The study of these issues and the aforementioned ones for 1d signals occupies Part I of the present manuscript.

An important consequence of the change of model from fBm to multifractal multiplicative cascades is that little can be said *theoretically* about the statistical performance of the analysis procedures. Whereas for fBm, asymptotical results (in the limit of large observation durations) can be established (see e.g., [4, 26, 52, 84, 169]), this is in general not possible for cascade processes. The lack of results is mainly due to the fact that their construction consists of multiplicative martingales and is extremely involved. What is more, such mathematical models result in stochastic processes with strong dependence and heavy tailed, strongly non Gaussian marginal distributions. As a result, analytical derivation of the statistical performance of analysis procedures is not feasible, even asymptotically, and the formulation of relevant definitions for estimators, confidence intervals or tests is in itself an issue. This implies that practitioners are lacking tools to assess the confidence they should grant to the obtained estimates. Along the same line, no hypothesis tests validating the precise multifractal nature and properties of the data under analysis are available, while this issue turns out to be mentioned as crucial in most contributions where multifractal analysis is used. For instance, there is so far no statistical procedure available in the literature that enables to decide whether real-life data are better described by finite variance self-similar processes (such as fractional Brownian motion) or by truly multifractal multiplicative cascade processes. Answering such questions is of major both theoretical and practical importance. The inferred understanding of the (physical, biological, ...) phenomena producing the data under analysis may be dramatically changed: Self-similar processes are indeed deeply related to random walks and additive phenomena, while most multifractal processes are historically tied to multiplicative structures. Also, both the number of parameters that need to be matched and the computational complexities that need to be handled for these different classes of models are radically different. However, despite the huge collection of research articles describing the practical use of multifractal analysis on real-life data, the state-of-the-art tools to assess confidence in estimates and decision making remained, at the outset of this work, the *experience* and *eyes* of the practitioners. This makes up for considerable practical limitations of the analysis tool.

The third main goal of this manuscript is to overcome such limitations. To this end, we propose in Part II of the present work the use of *nonparametric bootstrap* techniques [56, 64, 104, 195]. The principles of these techniques are stated in Chapter 6. The proposed procedures are inspired by wavelet domain bootstrap, pioneered in [137, 151]. The contribution lies in the adaptation of bootstrap resampling and estimation to match the specificities of multifractal analysis and to account for the difficult statistical context. This is addressed by an original blocks of wavelet Leaders in the time- (or space-) scale plane bootstrap (Chapter 7). The construction is based on commonly accepted intuitions deduced from the key properties of the dyadic wavelet transform for fBm [4, 6, 70, 77, 137, 166, 169]. This bootstrap procedure is in turn used for the design of confidence intervals (Chapter 8), the construction of statistical tests for multifractal models (Chapter 9) and for stationarity of multifractal attributes (Chapter 10). Their relevance is validated by means of numerical simulations. The use of bootstrap for multifractal analysis has never been reported before and these procedures represent, to the best of our knowledge, the only practically available confidence interval and statistical test procedures with satisfactory performance for scaling and multifractal analysis, readily applicable to single finite length observations of real-world data.

The attempt to study the *theoretical* validity of bootstrap for multifractal analysis collides with the limited amount of theory on the statistics of estimation procedures that is avail-

able in this context and precisely motivates its use. This is further amplified for wavelet Leader based analysis, since their statistical description remains difficult even for the model case of Gaussian and independent coefficients. The key issue turns out to be the theoretical characterization of the *dependence* of wavelet coefficients and Leaders for non Gaussian self-similar and multiplicative cascade processes. Therefore, we revisit in Chapter 11 the common heuristic of assuming that the vanishing moments N_ψ — responsible for weak correlation of wavelet coefficients for fBm — play a similar key role for more general processes. We establish analytical results on the dependence structure of multiresolution quantities, for both self-similar and multiplicative cascades processes, and validate them by numerical simulation. The implications for bootstrap are investigated numerically in Chapter 11.4. The results indicate that common intuitions need to be partly revised: The number of vanishing moments of the analyzing wavelet does indeed in general reduce the correlation among wavelet coefficients, yet it does not affect or control their dependence. These results have, to our knowledge, never been clearly reported elsewhere and have strong implications for practical analysis and applications. They represent another original contribution of the present work.

Part III of the present work is dedicated to applications based on the bootstrap and wavelet Leader based analysis procedures. In Chapter 12, we show them at work on real-life data produced by hydrodynamic turbulence experiments. Hydrodynamic turbulence is the scientific domain that gave birth to the concept of multifractal [75]. The seminal works by Yaglom and Mandelbrot in the 60s and 70s indeed proposed to describe the celebrated Richardson energy cascade of turbulence flows by means of multiplicative cascade split and multiply iterative constructions. It has since been recognized that velocity or dissipation turbulence fields possess scale invariance properties and are better described with multifractal models than with self-similar ones. It remains, however, to decide which precise multiplicative model better fits the data, an open and controversial issue. Answering such a question is of theoretical importance as it may help to better understand the physical mechanisms at work in the development of turbulence flows. The bootstrap based confidence intervals and hypothesis tests enable us to revisit this old question.
In Chapter 13, we investigate the application of wavelet Leader based multifractal analysis to texture image classification. We propose to replace commonly used features for characterizing their regularity, which are often obtained by heuristic arguments and analysis methods, with 2d wavelet Leader based multifractal attributes, and employ them for the classification of a large texture image database. Results indicate the relevance of multifractal attributes for texture image regularity characterization.
Finally, in Chapter 14, we conclude on the present work and state further issues and perspectives.

Part I

Multifractal Analysis

Chapter 2

Scale Invariance and Multifractal Analysis

Contents

In this chapter, we introduce the main concepts of scale invariance, self-similarity and scaling analysis, and the mathematical definitions and notions of multifractal analysis. We define and characterize the related analysis tools and practical estimation procedures, and we introduce self-similar and multiplicative cascade model processes and discuss their properties. The goal of this chapter is to give an overview on the main theoretical and practical notions and consequences of these mathematical concepts. Therefore, the chapter consists mostly of a synthetic overview presentation of the state of the art on the topic.
The main contribution of this chapter lies in the synthetic overview exposition itself. It also contains a couple of contributions on precise technical points. These are clearly stated in the text, and their detailed proofs are annexed in Appendices A and B.
Without loss of generality, most of the definitions will be stated in 1d for convenience of notation.

2.1 Scale Invariance: Intuitions

2.1.1 $1/f$ processes

Historically, scale invariance had been tied to $1/f$ stochastic 2nd order stationary processes, sometimes as well called $1/f$ noise since used for modeling thermal noise in electronics: Let X denote the signal under analysis, and $\hat{\Gamma}_X(\nu, t_k)$ any standard spectrum estimation procedure, such as the average over sliding time windows (centered around times t_k) of smoothed periodograms or Welch estimator. Scale invariance is related to a power law behavior of the spectrum estimate with respect to frequency over a wide range of frequencies, $\nu \in [\nu_m, \nu_M], \nu_M/\nu_m \gg 1$:

$$\frac{1}{n}\sum_{k=1}^{n} \hat{\Gamma}_X(\nu, t_k) \simeq C|\nu|^{-\gamma},\ \gamma > 0, \tag{2.1}$$

where C is a positive constant. Therefore, no frequency in (ν_m, ν_M) plays a privileged role, and the spectrum is covariant with respect to (w.r.t.) dilatation: $\Gamma_X(a\nu) = a^{-\gamma}\Gamma_X(\nu)$. Equivalently, no characteristic scale is present in the auto-covariance function[1] of the process: $c_X(\tau) = |a|^{1-\gamma}c_X(a\tau)$.

There exist two ways of reading Eq. (2.1), depending on the actual range $[\nu_m, \nu_M]$, namely *long range dependence* and *local self-similarity*. The first is obtained by fixing the lower characteristic frequency to $\nu_m = 0$ and has strong links to self-similarity. The latter is obtained by letting the high characteristic frequency go to infinity, $\nu_M = +\infty$. It has a strong link with the local regularity of the trajectories of $X(t)$ and deeply relates to multifractal analysis (cf. Section 2.4).

2.1.1-a) Long range dependence (LRD)

Let us first explore the power law behavior in the limit of $|\nu| \to 0$ or, equivalently, $|\tau| \to \infty$:

[1] From now on, and throughout this manuscript, we refer to the auto-covariance function of a process as its covariance function. When the term covariance function is used for the cross-covariance, this is clearly and explicitly stated. Similarly, we refer to the auto-correlation function as the correlation function.

Definition 2.1 (Long range dependence) *A second-order stationary process with finite variance $X(t)$ is called long range dependent (LRD) if its power spectral density $\Gamma_X(\nu)$ behaves, in the limit of $|\nu| \to 0$, as:*

$$\Gamma_X(\nu) \sim C|\nu|^{-\gamma}, \tag{2.2}$$

with $0 < \gamma < 1$ and C a positive constant [33, 153].

Equivalently, this means that the covariance function,

$$c_X(\tau) \sim \tilde{C}|\tau|^{-\alpha}, \quad |\tau| \to \infty, \quad 0 < \alpha < 1, \quad \alpha = 1 - \gamma, \tag{2.3}$$

decays very slowly in the limit $|\tau| \to \infty$ ($\tilde{C}$ being a positive constant), such that:

$$\int_A^{+\infty} c_X(\tau) d\tau = +\infty, \tag{2.4}$$

for any $0 < A < +\infty$.
A direct and major practical consequence of LRD is that estimation becomes very difficult. For instance, the variance of the sample mean $\hat{\mu}_N = \frac{1}{N}\sum_{n=1} X_n$ decays as Var$\hat{\mu}_N \sim CN^{-\alpha}$ with $0 < \alpha < 1$ as $N \to \infty$, hence much slower than the common N^{-1}. In general, for higher order sample moment estimation, such as variance, the estimations are strongly biased and have very slowly decreasing variance [33, 153].

2.1.1-b) Scale invariance, fractal trajectories and local regularity

Let us now explore the other limit, when the covariance function (equivalently, power spectral density) displays power law behavior for small time lags $|\tau| \to 0$ (equivalently, large frequencies $|\nu| \to \infty$).
Suppose that the covariance function takes the form

$$c_X(\tau) = \mathbb{E}X(t+\tau)X(t) \sim C(1 - |\tau|^{2h}), \quad |\tau| \to 0, \tag{2.5}$$

which implies $\Gamma_X(\nu) \sim C|\nu|^{-(2h+1)}$, $|\nu| \to +\infty$. Notice that then, h describes the local regularity of the trajectories of $X(t)$: The closer h to 0, the more irregular the trajectories, the larger h, the smoother they are. In the case of a Gaussian process, we can be more precise: The trajectories have Hausdorff dimension $2 - h$, and are $\tilde{h} < h$ Hölder (for precise definitions of Hölder regularity and Hausdorff dimension, see Section 2.4). For a covariance function as in Eq. (2.5), the increments $X(t+\tau) - X(t)$ of the process are, for small time lags τ, characterized by [84]:

$$\mathbb{E}|X(t+\tau) - X(t)|^2 \sim C|\tau|^{2h}, \quad |\tau| \to 0. \tag{2.6}$$

This locally singular behavior of X intimately relates scale invariance to multifractal analysis (cf. Section 2.4).

2.1.2 Practical definition of scale invariance and scaling analysis

A number of authors have proposed to enlarge or enrich the notion of scale invariance [2, 4, 5, 6]: Scale invariance is now commonly and operationally defined as the power law behaviors of (the time average of the $q-$th power of) multiresolution quantities, labeled

$T_X(a,t)$, with respect to the analysis scale a, for a given (large) range of scales $a \in (a_m, a_M), a_M/a_m \gg 1$:

$$\frac{1}{n_a}\sum_{k=1}^{n_a} |T_X(a,k)|^q \simeq c_q a^{\zeta(q)}. \tag{2.7}$$

With respect to Eq. (2.1), the definition in Eq. (2.7) implies two major changes in paradigms:

1. Standard spectral estimates are replaced with *multiresolution quantities* $T_X(a,t)$. These are quantities describing the content of X around a time position t, and a scale a. Standard examples for the $T_X(a,t)$ are given by wavelet, increment or box-aggregated coefficients. Qualitatively, the analysis scale a acts as the inverse of the frequency: $a \sim \nu_0/\nu$ (ν_0 being an arbitrary constant). The multiresolution quantities $T_X(a,t)$ can therefore be seen as some sort of spectral estimates. For instance, for wavelet coefficients, it can be shown that:

$$\mathbb{E}T_X(a,t)^2 = \int \Gamma_X(\nu) a |\tilde{\psi}_0(a\nu)|^2 d\nu, \tag{2.8}$$

   where $\tilde{\psi}_0$ is the Fourier transform of the mother wavelet [1, 6, 7, 12].

2. The second statistical order $q = 2$ is replaced with a *whole range* of (positive and negative) statistical orders q.

It has been shown that scale invariance as in Eq. (2.7) can fruitfully be modeled with self-similar [153] and/or multifractal processes [145].

Definition 2.2 (Scaling analysis) *The aim of scaling analysis is to validate the existence of power law behaviors as in Eq. (2.7), and to measure the scaling exponents $\zeta(q)$ that characterize them.*

Essentially, analysis and estimation procedures consist in tracking straight lines and estimating slopes in log-log plots, as suggested by Eq. (2.7) above. The estimated exponents can in turn be used for the physical understanding of the data or of the systems producing them or are involved in standard signal processing tasks such as detection, hypothesis testing, identification or classification.
The link of scaling analysis and multifractal analysis is made explicit in Section 2.5.3.

2.1.3 Scale invariance and self-similarity

The central theoretical notion of scale invariance is that of self-similarity: The whole resembles the part, and the part resembles the whole. Therefore, the (statistical) information that can be obtained from an object is independent of the scale of observation. The paradigm and most popular examples of self-similar objects are probably given by fractals – objects that show the same geometric features at a (typically discrete) set of scales. Fig. 2.1 shows the famous von Koch curve as an example of a fractal object: It is clear that the essential information on the nature of such objects can not be gathered by measurements on a fixed scale (where we essentially observe that the object is constructed from equilateral triangles) but only by identifying the mechanisms relating different scales (from which one can, for example, deduce that the curve has infinite length and a fractal box-counting dimension of $\log 4/\log 3$).
In the next section, we are interested in stochastic versions of this deterministic scheme.

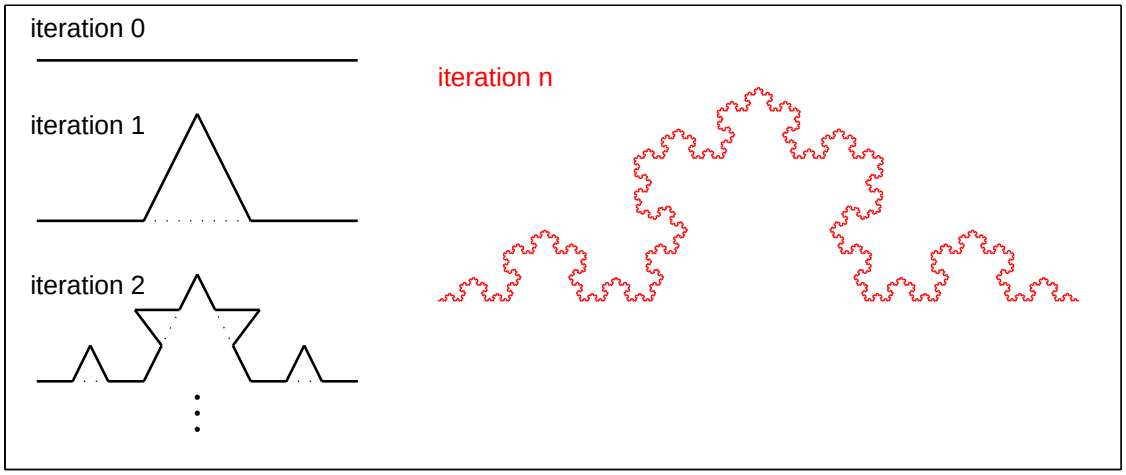


Figure 2.1: **Von Koch curve.** Iterative construction principle of the von Koch curve (left), and the curve after $n = 6$ iterations. In the limit $n \to \infty$, the curve has infinite length and a box-counting dimension of $\log 4/\log 3$. This essential information on the geometrical nature of the object has to be deduced from the mechanisms relating different scales. It can not be obtained by measurements at a fixed observation scale.

2.2 Statistical Self-Similarity and H-sssi Processes

The most important and commonly used models for random self-similarity are *self-similar processes*.

2.2.1 Self-similar process

Definition 2.3 (Self-similar process) *A process $\{X(t), t \in \mathbb{R}^+\}$ is said to be self-similar (H-ss) if and only if for all $a > 0$ [69, 153]:*

$$X(t) \stackrel{fdd}{=} a^{-H} X(at). \tag{2.9}$$

The symbol $\stackrel{fdd}{=}$ means that the process $X(t)$ and the process $a^{-H}X(at)$ have the same finite dimensional distributions, that is, for all $a > 0$ and $n \in \mathbb{N}$ and $t_1 < t_2 < \cdots < t_n$,

$$[X(t_1), X(t_2), \cdots, X(t_n)] \stackrel{d}{=} a^{-H}[X(at_1), X(at_2), \cdots, X(at_n)], \tag{2.10}$$

where $\stackrel{d}{=}$ stands for equality in distribution. The parameter $H > 0$ is called the *self-similarity parameter* or *Hurst parameter*. The definition implies that one can not statistically distinguish the process $X(t)$ from the dilated process $a^{-H}X(at)$, for any dilation factor $a > 0$.
A first direct consequence of self-similarity is that all finite moments of $X(t)$ display power law behaviour[2]:

$$\mathbb{E}|X(t)|^q = |t|^{qH}\mathbb{E}|X(1)|^q, \quad \forall t > 0, \quad \forall q : \mathbb{E}|X(t)|^q < +\infty, \tag{2.11}$$

[2]By convention, and without loss of generality, $X(0) \equiv 0$.

which will play a central role in parameter estimation procedures. It follows that therefore, any self-similar process $X(t)$ is by definition non stationary. Let us note that the theorem of Lamperti provides an interesting theoretical connection between self-similar and stationary processes: If $\{X(t), t \in \mathbb{R}^+\}$ is a self-similar process, then $\{Z(t) = \exp(-Ht)X(\exp(t)), t \in \mathbb{R}\}$ is a stationary process, and if $\{Z(t), t \in \mathbb{R}\}$ is a stationary process, then $\{X(t) = t^H Z(\ln t), t \in \mathbb{R}^+\}$ is a self-similar process [39, 105].

2.2.2 Self-similar process with stationary increments: H-sssi process

In practice, the non stationarity of self-similar processes is problematic from both modeling and analysis points of view. Therefore, one usually considers the sub-class of *self-similar processes with stationary increments* (H-sssi processes).

2.2.2-a) Definition

Definition 2.4 (Stationary increments) *A process $X(t)$ is said to have stationary increments if and only if*

$$\delta_\tau X(t) \stackrel{fdd}{=} \delta_\tau X(0), \quad \forall \tau \geq 0. \tag{2.12}$$

where $\delta_\tau X(t)$ are the increments of $X(t)$:

$$\delta_\tau X(t) = X(t+\tau) - X(t). \tag{2.13}$$

The self-similarity of $X(t)$ translates to the $\delta_\tau X(t)$:

$$\delta_\tau X(t) \stackrel{fdd}{=} a^{-H} \delta_{a\tau} X(at), \quad \forall a > 0, \tag{2.14}$$

just as the power law behavior of the moments:

$$\mathbb{E}|\delta_\tau X(t)|^q = |\tau|^{qH} \mathbb{E}|\delta_1 X(0)|^q, \quad \forall t, \tau > 0. \tag{2.15}$$

H-sssi processes, as well as their increments, have mean zero. If the process $X(t)$ has finite variance (fv), its covariance is given by:

$$\mathbb{E}X(t)X(s) = \frac{\mathbb{E}|X(1)|^2}{2} \left(|t|^{2H} + |s|^{2H} - |t-s|^{2H} \right), \tag{2.16}$$

from which follows that:

$$0 < H < 1, \tag{2.17}$$

since the covariance function is positive definite. It is also easy to show that then, the covariance function of the increments is given by [153]:

$$\mathbb{E}\delta_\tau X(t+s)\delta_\tau X(t) = \frac{\mathbb{E}|X(1)|^2}{2} \left(|s+\tau|^{2H} + |s-\tau|^{2H} - 2|s|^{2H} \right). \tag{2.18}$$

2.2.2-b) H-sssi process and long range dependence

For $s \gg \tau$, the auto-covariance function of the increments of a H-sssi process (cf. Eq. (2.18)) behaves as:

$$\mathbb{E}\delta_\tau X(t+s)\delta_\tau X(t) \simeq Cs^{2(H-1)}\tau^2\mathbb{E}|X(1)|^2 H(2H-1), \quad H \neq \frac{1}{2},\ s \gg \tau. \tag{2.19}$$

Consequently, the increments of a H-sssi process form a long range dependent process when the self-similarity parameter is in the range:

$$\frac{1}{2} < H < 1 \tag{2.20}$$

which corresponds to Definition 2.1 and Eq. (2.3) with $\gamma = 2H-1$ and $\alpha = 2(1-H)$, respectively. This establishes a strong link between the asymptotic property of long range dependence and self-similarity.
Although H-sssi processes are non stationary, one commonly says that they are long range dependent if $\frac{1}{2} < H < 1$, and heuristically associates a spectrum $\sim |\nu|^{-(2H+1)}$ with them.

2.2.2-c) H-sssi process, fractal trajectories and local regularity

Eq. (2.15) implies that H-sssi processes display local power law behavior as in Eq. (2.6):

$$\mathbb{E}|X(t+\tau) - X(t)|^2 = \mathbb{E}X(1)^2|\tau|^{2H}. \tag{2.21}$$

Hence, H characterizes the local regularity of $X(t)$ everywhere, and the local regularity is the same all along the trajectory. These properties strongly connect H-sssi and multifractal analysis. This will be explained in Section 2.4.

2.2.3 H-sssi processes: Examples

In the present work, we exclusively consider processes with *finite variance* (fv). In particular, empirical studies will concentrate on the (Gaussian) fractional Brownian motion, and the (non Gaussian) Rosenblatt process.

2.2.3-a) Fractional Brownian motion (fBm)

Fractional Brownian motion (fBm) $B_H(t)$ constitutes the archetype of fv H-sssi processes. It is the only Gaussian finite variance self-similar process with stationary increments [153], and it is fully defined by the self-similarity parameter $H \in (0,1)$, up to a multiplicative factor. Therefore, it is a LRD process when $H \in (0.5, 1)$. Reintroduced and made popular much later by Mandelbrot [121], it has first been studied by Kolmogorov in the context of turbulent flows in the early 40s [100]. Fractional Brownian motion can equivalently be defined as:

$$B_H(t) = k_H \int_{-\infty}^{\infty} \left((t-x)_+^{\frac{H}{2}-1} - (-x)_+^{\frac{H}{2}-1} ds \right) dB(x), \tag{2.22}$$

where k_H is a normalizing constant (e.g. such that $\mathbb{E}B_H(1)^2 = 1$), $x_+ = \max\{x, 0\}$ for $x \in \mathbb{R}$ and $B(t)$, $t \in \mathbb{R}$, is a standard Brownian motion[3] [153]. The increment process of

[3] Note that for $H = 0.5$, this is to be interpreted as an integral representation of standard Brownian motion.

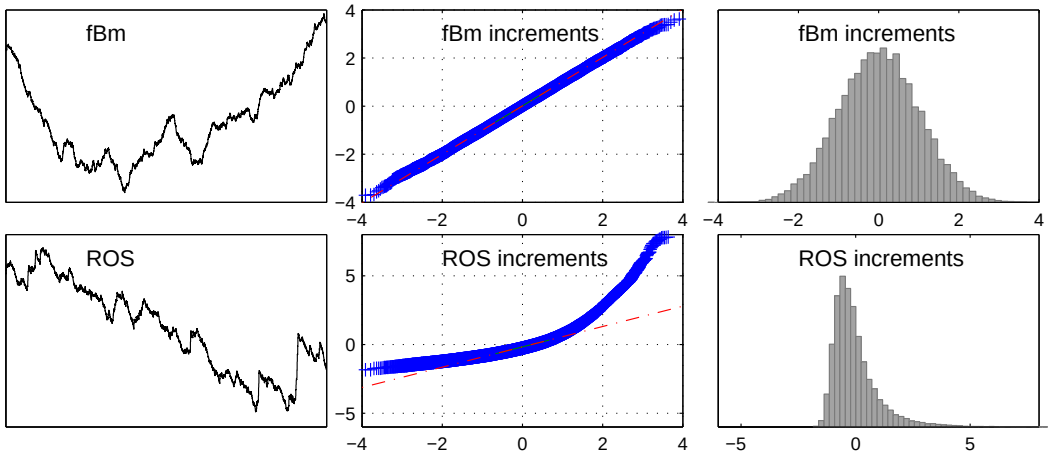


Figure 2.2: **Trajectories and increments of H-sssi processes.** Trajectories (left), quantile-quantile plots versus standard normal (quantiles given by abscissa) of normalized empirical distributions of increments (center), and normalized empirical distributions of increments (right) for fBm (top, $H = 0.7$) and ROS (bottom, $H = 0.7$): Whereas the increments of fBm are Gaussian, they are strongly non Gaussian and skewed for ROS.

fractional Brownian motion is commonly referred to as *fractional Gaussian noise* (fGn). The reader is referred to [153] and references therein for more details. For numerical synthesis procedures in 1d and 2d, see e.g. [27, 153] and [163], respectively.

2.2.3-b) Rosenblatt Process (ROS)

There also exist fv non Gaussian $H-$sssi processes, the most well known being the *Rosenblatt process* (ROS) [146]. The Rosenblatt process, with $H \in (1/2, 1)$, is defined as

$$Z_H(t) = k_H \int_{\mathbb{R}^2}{}' \left\{ \int_0^t (s-u)_+^{\frac{H}{2}-1} (s-v)_+^{\frac{H}{2}-1} ds \right\} dB(u) dB(v), \tag{2.23}$$

where k_H is a normalizing constant (e.g. such that $\mathbb{E} Z_H(1)^2 = 1$), $\int_{\mathbb{R}^2}{}'$ denotes the double Wiener-Itô integral and $'$ denotes that the integral is not taken over the diagonal (i.e., $u \neq v$). $B(t)$, $t \in \mathbb{R}$, stands for the standard Brownian motion and $x_+ = \max\{x, 0\}$ for $x \in \mathbb{R}$. For further technical details on the definition, properties and numerical simulation procedures, the reader can consult e.g., [10] and references therein. The two major properties of the Rosenblatt process of interest in this manuscript are as follows: ROS has exactly the same covariance function as fBm but more complex higher order dependence. Also, ROS has a non Gaussian highly skewed marginal distribution (as illustrated from numerical simulations for its increments in Fig. 2.2). Finally, since it is defined only for $H \in (1/2, 1)$, ROS is always an LRD process.

2.2.3-c) α-stable motion.

For sake of completeness, we mention another class of self-similar stationary increment processes, which is given by the α-stable processes. They will not be considered any further in this thesis, and the interested reader is referred to e.g. [153].

2.3 Wavelets and H-sssi Processes

Many different practical methods have been proposed for scaling analysis and for the estimation of the self-similarity parameter H. A prominent place is occupied by the *discrete wavelet transform* (DWT). It will also play a key role in the definition of practical procedures for performing a multifractal analysis (see Section 2.5). Therefore, we concentrate on this method, and the reader is referred to e.g. [33] for an overview of methods.

2.3.1 Wavelet transform in 1d

The original idea of the wavelet transform is to provide a decomposition of the signal $X(t)$ in the time-scale plane [43, 55, 117]. It compares the signal by means of inner products to a set of analyzing functions or reference patterns that are obtained, by time shift and dilation (change of scale) operations, from an elementary function $\psi_0(t)$, referred to as the *mother wavelet*. The mother wavelet has to fulfill the admissibility condition

$$\int_{\mathbb{R}} \psi_0(t)dt \equiv 0, \tag{2.24}$$

and is chosen such that its energy remains mostly concentrated in a narrow support both in the time and frequency domains. It is characterized by its *number of vanishing moments*, a strictly positive integer $N_\psi \geq 1$ defined as:

$$\forall l = 0, 1, \ldots, N_\psi - 1 : \int_{\mathbb{R}} t^l \psi_0(t)dt \equiv 0, \text{ and } \int_{\mathbb{R}} t^{N_\psi} \psi_0(t)dt \neq 0. \tag{2.25}$$

The mother-wavelet, and its derivatives up to order N_ψ, have fast exponential decays in the time domain.

Continuous wavelet transform. Let $\psi_{a,t}(u)$ denote the template of $\psi_0(t)$ dilated to scale a and translated to position u:

$$\psi_{a,t}(u) = \frac{1}{\sqrt{a}} \psi_0 \left(\frac{u-t}{a} \right). \tag{2.26}$$

The *continuous wavelet transform* (CWT) is defined trough its coefficients:

$$C_X(a,t) = \langle \psi_{a,t} | X \rangle = \int_{\mathbb{R}} X(u)\ \psi_{a,t}(u)\, du. \tag{2.27}$$

The CWT is a very rich (redundant) representation of $X(t)$, since it maps a time signal – hence a signal whose information can be entirely represented in the time domain – in the time-scale domain. Therefore, it is natural to ask whether one can restrict the number of coefficients while keeping all the information of the signal in the wavelet representation.

Discrete wavelet transform. At a first reading, the discrete wavelet transform (DWT) can be seen as such a sampling of the time-scale plane by keeping only wavelet coefficients for dilations to scales and translations to positions:

$$a = 2^j, \quad t = k2^j. \tag{2.28}$$

The DWT is, however, much more than that: It is constructed in such a way that the collection of templates of ψ_0:

$$\{\psi_{j,k}(t) = 2^{-j/2} \psi_0(2^{-j}t - k), j \in \mathbb{Z}, k \in \mathbb{N}\} \tag{2.29}$$

forms an orthonormal basis of $L^2(\mathbb{R})$, i.e.:

$$X(t) = \sum_{j,k} \langle X|\psi_{j,k}\rangle \psi_{j,k}(t) = \sum_{j,k} D_X(j,k)\psi_{j,k}(t) \tag{2.30}$$

where the $D_X(j,k)$ are the discrete wavelet transform coefficients:

$$D_X(j,k) = \langle \psi_{j,k}|X\rangle = \int_{\mathbb{R}} X(t)\ 2^{-j/2}\psi_0(2^{-j}t-k)\,dt. \tag{2.31}$$

The solution to finding mother wavelets $\psi_0(t)$ that satisfy these constraints is embedded in the so-called *multiresolution theory* [55, 117]. It relies on the design of a *scaling function* $\phi_0(t)$, from which the mother wavelet $\psi_0(t)$ yielding the orthonormal basis can be constructed. Moreover, discrete time quadrature mirror filters $H_0(k)$ and $G_0(k)$ can be associated with the scaling function $\phi_0(t)$ and the mother wavelet $\psi_0(t)$, respectively, except at the first scale. This gives rise to fast algorithms. The absence of discrete time filters for the first scale makes it theoretically necessary to perform a *projection step*, which is often omitted in practice [3, 172].
For the definition of the DWT in 2d in the next section, we prefer, for convenience, a notation that translates to the 1d case as: $\phi_0 \leftrightarrow \psi_0^{(0)}$ and $\psi_0 \leftrightarrow \psi_0^{(1)}$.

2.3.2 Wavelet transform in 2d

Continuous wavelet transform. The continuous wavelet transform in 2d is constructed in a similar way as the 1d CWT [14, 117]. It is defined by its coefficients:

$$C_X(a,\mathbf{t}) = \langle \psi_{a,\mathbf{t}}|X\rangle = \int_{\mathbb{R}^2} X(\mathbf{u})\ \psi_{a,\mathbf{t}}(\mathbf{u})\,d\mathbf{u}, \quad \mathbf{t},\mathbf{u}\in\mathbb{R}^2. \tag{2.32}$$

Discrete wavelet transform. The construction of the discrete wavelet transform in 2d is, however, slightly different from the 1d case. Its lower approximation and the detail coefficients are given by:

$$D_X^{(m)}(j,\mathbf{k}) = \langle \psi_{j,\mathbf{k}}^{(m)}|X\rangle, \quad m=0,1,2,3, \quad \mathbf{k}\in\mathbb{N}^2 \tag{2.33}$$

where the collection $\{\psi_{j,\mathbf{k}}^{(m)}, j\in\mathbb{Z}, \mathbf{k}\in\mathbb{N}^2, m=0,1,2,3\}$ forms a basis of $L^2(\mathbb{R}^2)$. For the general theory on the design of such functions $\psi^{(m)}$, the reader is referred to [14]. Here, we will restrict ourselves to the following definition: A 2d orthonormal DWT can be practically defined via the use of 4 bi-dimensional filters $G^{(m)}(k_1,k_2)$, $m=0,1,2,3$ obtained as tensor products of the quadrature mirror filters H_0 and G_0 (low-pass and high-pass, respectively) defining a 1d orthonormal DWT. By convention, $G^{(0)}(k_1,k_2) = H_0(k_1)H_0(k_2)$ corresponds to the 2d low pass filter providing a lower approximation, while $G^{(m)}$, $m=1,2,3$, correspond to the high pass filters yielding the wavelet coefficients: $G^{(1)}(k_1,k_2) = G_0(k_1)H_0(k_2)$, $G^{(2)}(k_1,k_2) = H_0(k_1)G_0(k_2)$ and $G^{(3)}(k_1,k_2) = G_0(k_1)G_0(k_2)$.

2.3.3 Normalization of wavelet coefficients and practical implementation

Normalization of wavelet coefficients. For scaling analysis, it is more suitable (cf. [2, 24]) to re-normalize the standard L^2-norm wavelet coefficients according to a L^1-norm:

$$d_X(j,k) = 2^{-j/2} D_X(j,k), \tag{2.34}$$

$$d_X^{(m)}(j,\mathbf{k}) = 2^{-j} D_X^{(m)}(j,\mathbf{k}), \quad \mathbf{k} \in \mathbb{Z}^2, \quad m = 1,2,3. \tag{2.35}$$

We will refer to d_X as the (discrete) wavelet coefficients of X.

Practical implementation. There exist a certain number of families of mother wavelets for the definition of a DWT [14, 117]. Throughout this manuscript, we will make use of the family of *Daubechies wavelets* [54]. These wavelets have compact support, and the associated discrete time filters have minimum length.
The codes for the 1d DWT that are used for obtaining the results presented in this manuscript have been implemented in MATLAB© by ourselves. The 2d DWT is performed using (a corrected version of) the Rice Wavelet Toolbox (`www.dsp.rice.edu/software/rwt.shtml`).

2.3.4 Properties for finite variance H-sssi processes

The wavelet coefficients of a fv H-sssi process $X(t)$ have the following properties (e.g. [4]):

P1 If $N_\psi \geq 1$, the coefficients $\{d_X(j,k),\, k \in \mathbb{N}\}$ form a *stationary* process.

P2 The coefficients $\{d_X(j,k),\, k \in \mathbb{N}\}$ reproduce the *scaling* property Eq. (2.9):

$$d_X(j,k) \overset{fdd}{=} 2^{jH} d_X(0,k). \tag{2.36}$$

P3 The correlation function of the coefficients $d_X(j,k)$ and $d_X(j,k')$ behaves as[4]:

$$\mathbb{E} d_X(j,k) d_X(j,k') \sim C|k-k'|^{2(H-N_\psi)},\ |k-k'| \to \infty, \tag{2.37}$$

for large time lags $|k-k'|$. This is a direct consequence of the specific covariance structure of fv H-sssi processes Eq. (2.16). Therefore, the coefficients $\{d_X(j,k),\, k \in \mathbb{Z}\}$ are *not LRD* if:

$$N_\psi \geq H + \frac{1}{2}. \tag{2.38}$$

Hence, the wavelet coefficients exactly reproduce the scaling properties of $X(t)$ while getting rid of the statistically difficult context of $X(t)$ (LRD and non-stationary): They form *stationary* and only *weakly correlated* sequences. The first is a consequence of the fact that the family of analyzing wavelets itself exhibits a scale invariance feature, since obtained by a dilation operation, the latter of the (sufficient) number of vanishing moments of the mother wavelet. Combining (**P1**) and (**P2**) gives, for all finite moments $\forall q > 0 : \mathbb{E}|d_X(0,\cdot)|^q < \infty$:

$$\mathbb{E}|d_X(j,\cdot)|^q = \underbrace{\mathbb{E}|d_X(0,\cdot)|^q}_{c_q} \cdot 2^{jqH}. \tag{2.39}$$

[4] This can be written more generally for coefficients at different scales j and j' as in Eq. (11.2), cf. [70].

Finally, for the specific case of fBm:

P4 The marginal distributions of the coefficient of fBm are Gaussian with zero mean and variances $\sigma^2(j) = \text{Var} d_X(j,\cdot)$:

$$d_X(j,\cdot) \sim \mathcal{N}(0, \sigma(j)).$$

This explains and underlines the predominant role that the discrete wavelet coefficients play in both theoretical and practical analysis of scale invariance, notably for $q = 2$.

2.4 Multifractal Analysis: Theory

2.4.1 Scale invariance, higher order statistics and multifractal analysis

Self-similarity is a very demanding property, as it implies exact invariance to dilatation for *all scale factors* $a > 0$ of all finite dimensional distributions, therefore involving *all statistical orders* of the process (cf. Eq. (2.9)). This is in practice often too restrictive a model, be it because of practical estimation problems for large statistical orders, or due to the nature of the process under analysis. Therefore, investigation of scale invariance is sometimes restricted to the statistical order 2, and to a certain range of scale factors only: This is precisely the intuition and definition of $1/f$ processes in Eq. (2.1), which is in some sense practically convenient, since it allows to model and analyze a wider range of scale invariance. It is, however, as well a weak model in the sense that it gives only a coarse description of the scale invariance properties of a process in terms of the second statistical order and is by definition blind to any aspects involving higher order statistics.
One definition of statistical scale invariance – which is at the same time stronger and more flexible – is obtained by relaxing the self-similarity property in a similar way as in the practical definition in Section 2.1.2: First, scale invariance has to hold only for a restricted range of scaling factors $a \in (a_m, a_M), a_M/a_m \gg 1$, rather than for all $a > 0$. Second, the (single) self-similarity parameter H is replaced with the function $\zeta(q)$, called the scaling function or the *scaling exponents* of $X(t)$:

Definition 2.5 (Scale invariant stationary increment process) *Suppose $X(t)$ is a process with stationary increments. Then it is scale invariant if*

$$\mathbb{E}|X(at)|^q = |a|^{\zeta(q)} \mathbb{E}|X(t)|^q \tag{2.40}$$

for a range of scales $a \in (a_m, a_M), a_M/a_m \gg 1$ and some range of statistical orders q.

This implies for the increment process:

$$\mathbb{E}|\delta_\tau X(t)|^q = |\tau|^{\zeta(q)} \mathbb{E}|\delta_1 X(0)|^q,\ 0 < \tau_m < \tau < \tau_M < \infty. \tag{2.41}$$

Comparing this definition with that of H-sssi (Def. 2.3 plus Def. 2.4), we see that Eq. (2.41) is highly reminiscent of Eq. (2.15), with $\zeta(q) = qH$. A central difference lies in the fact that in Eq. (2.41) the single parameter H is replaced with a function $\zeta(q)$, which can in general be non linear, $\zeta(q) \neq qH$, and hence represents a whole collection of parameters for the characterization of the process.

The analysis of processes that satisfy Eqs. (2.40) and (2.41) is often conducted in the

framework of *multifractal analysis* (MFA). Multifractal analysis studies how the (pointwise) *local regularity* of X fluctuates in time (or space). It therefore starts from a point of view that is a priori different from that of scaling analysis. We will, however, see that although the starting points look very different at first, scaling and multifractal analysis have a strong link, and that multifractal analysis will bring us back to Eqs. (2.40) and (2.41) in the limit of small scales $a_m, \tau_m \to 0$. Let us now introduce the notions and concepts of multifractal analysis. For general and exhaustive theoretical accounts for multifractal analysis, the reader is referred to e.g. [86, 87, 145].

2.4.2 Local regularity of functions

Definition 2.6 (Pointwise Regularity) *A function $f(t)$, $f : \mathbb{R}^d \to \mathbb{R}$ is $C^\alpha(t_0)$ with $\alpha > 0$, denoted as $f \in C^\alpha(t_0)$, if there exist $C > 0$, $\varepsilon > 0$ and a polynomial $P_{t_0}(\tau)$ of order strictly smaller than α such that:*

$$\text{if } |\tau| \leq \varepsilon, \quad |f(t_0 + \tau) - P_{t_0}(\tau)| \leq C|\tau|^\alpha. \tag{2.42}$$

If such a polynomial $P_{t_0}(\tau)$ exists, it is unique. Its constant part is always given by $P_{t_0}(0) = f(t_0)$.

Definition 2.7 (Hölder exponent) *The Hölder exponent $h_f(t_0)$ of f at t_0 is*

$$h_f(t_0) = \sup\{\alpha \,:\, f \,\in\, C^\alpha(t_0)\}. \tag{2.43}$$

Heuristically, if f has Hölder exponent $h(t_0)$ at t_0, one can write

$$\left|f(t_0 + \tau) - (f(t_0) + a_1(t_0)\tau + a_2(t_0)\tau^2 + \cdots + a_N(t_0)\tau^N)\right| \leq C(t)|\tau|^{h_f(t_0)}, \tag{2.44}$$

with $N < \alpha$. Intuitively, the Hölder exponent therefore gives a much finer account to the local regularity of a function, in between the notions "differentiable" and "continuous". Most interestingly, if $P_{t_0}(\tau) = f(t_0)$ reduces to a constant[5], the Hölder exponent characterizes the power law behavior of the increments at t_0:

$$|f(t_0 + \tau) - f(t_0)| \leq C(t_0)|\tau|^{h_f(t_0)}. \tag{2.45}$$

Hölder exponent and singularities. The Hölder exponent generalizes the heuristic definition of "singularity" that we introduced in Section 2.1.1-b): If f has Hölder exponent $h(t_0) = h_f(t_0) < 1$ at t_0, then f has at t_0 either a *cusp-type singularity* (cf. Fig. 2.5, top left):

$$|f(t_0 + \tau) - f(t_0)| \sim C|\tau|^{h(t_0)}, \tag{2.46}$$

or an *oscillating singularity* (or chirp-type singularity, cf. Fig. 2.5, top right):

$$|f(t_0 + \tau) - f(t_0)| \sim C|\tau|^{h(t_0)} \sin\left(\frac{1}{|\tau|^\beta}\right), \tag{2.47}$$

with oscillation exponent $\beta > 0$. Conversely, if f has either a cusp or an oscillating singularity at t_0 and $1 > h(t_0) > 0$, then $h(t_0) = h_f(t_0)$ is the Hölder exponent of f at t_0. Note that the converse can only hold for $h(t_0) > 0$, since in Def. 2.7 Hölder exponents are *positive*. Technically, the definition of negative Hölder exponents is feasible but much more involved. This is beyond the scope of this work [91, 178]. Nonetheless, note that the heuristic $|f(t_0 + \tau) - f(t_0)| \sim C|\tau|^{h(t_0)}$, $\tau \to 0$ stays valid with negative exponents $h(t_0) < 0$.

[5] This case is most interesting, since it can always be reached by differentiating enough times.

2.4.3 Multifractal spectrum

The aim of multifractal analysis is to provide a description of the collection of Hölder exponents h of the function f. Since the Hölder exponent may jump from one point to another, describing them for each time instance in form of a function $h_f(t)$ is rather meaningless. Therefore, multifractal analysis provides a global description of the regularity of the function of f in form of a *multifractal spectrum* (also called the *singularity spectrum*). It describes the "size" of the set of points for which the Hölder exponent takes a certain value h. The measure for "size" most commonly taken is the Hausdorff dimension and gives rise to the Hausdorff spectrum (which we will associate with and call the multifractal spectrum). It describes the collection of Hölder exponents $h(t)$ by mapping to each value of h the Hausdorff dimension $D(h)$ of the collection of points t_i at which $h_f(t_i) = h$:

Definition 2.8 (Iso-Hölder sets) *The iso-Hölder set $I_f(h)$ is the collection of points t_i for which the Hölder exponent takes a certain value h.*

$$I_f(h) = \{t_i | h_f(t_i) = h\}. \tag{2.48}$$

For defining the Hausdorff dimension, we first have to recall the definition of a Hausdorff measure:

Definition 2.9 (Hausdorff measure) *Let $S \subset \mathbb{R}^d$, $\varepsilon > 0$, and let $\gamma_\varepsilon(S)$ be ε-coverings of S, that is, bounded sets $\{c_n\}_{n\in\mathbb{N}}$ of radius $|c_n| \leq \varepsilon$ (maximal distance between two elements of c_n) that cover S: $S \subset \gamma_\varepsilon(S)$. Let $\mathcal{C}_\varepsilon(S)$ be the collection of all ε-coverings $\gamma_\varepsilon(S)$ of S. The δ-dimensional Hausdorff measure of S is:*

$$m_\delta(S) = \lim_{\varepsilon\to 0} \inf_{\mathcal{C}_\varepsilon(S)} \sum_{\gamma_\varepsilon(S)} |c_n|^\delta. \tag{2.49}$$

It can be shown that either $m_\delta(S) = 0$ if $\delta < \delta_c$, or $m_\delta(S) = +\infty$ if $\delta > \delta_c$. The Hausdorff dimension of S is defined as the critical value δ_c:

Definition 2.10 (Hausdorff dimension) *The Hausdorff dimension $\dim_H(S)$ of $S \subset \mathbb{R}^d$ is given by:*

$$\dim_H(S) = \inf_\delta \{m_\delta(S) = +\infty\} = \sup_\delta \{m_\delta(S) = 0\}. \tag{2.50}$$

The multifractal spectrum assigns now to each Hölder exponent – as a measure of its geometric importance – the Hausdorff dimension of the set of points that share the same exponent h:

Definition 2.11 (Multifractal spectrum) *The multifractal spectrum of a function f is defined as the Hausdorff dimension of the iso-Hölder sets $I_f(h)$:*

$$\mathcal{D}_f(h) = \dim_H(I_f(h)). \tag{2.51}$$

By convention, if h' is not a Hölder exponent of f, $\mathcal{D}_f(h') = -\infty$. Following [145], we consider the multifractal spectrum of a *process* to be the multifractal spectrum of each of his realizations. We will not distinguish any more between functions and processes in what follows, and we will simply speak of multifractal spectrum, for both functions and processes. For more technical details, the reader is referred to [145].

Therefore, the goal of multifractal analysis is to determine the multifractal spectrum. It describes a local property, the point-wise regularity of a function, globally through the geometrical importance of different Hölder exponents, disregarding any information on their precise geometric repartition.

2.4.4 Multifractal classification of processes

Functions and processes are classified by the properties of their multifractal spectra.

2.4.4-a) Homogenous processes

In this manuscript, we will only consider homogenous processes, that is, processes $X(t)$ for which the Hölder exponents $h_X(t)$ are homogeneously distributed on the support of $X(t)$:

Definition 2.12 (Homogenous function or process) *A function or process $X(t)$ is called homogenous if it has the same multifractal spectrum on all nonempty open sets of its support [106].*

For such processes, it can be shown that $h_X(t)$ is either a constant, or discontinuous everywhere. This definition excludes processes which are smoother in some regions than in others: For instance, it excludes *multifractional* processes, for which the Hölder exponent is a regular, smooth function which takes a different value for each point of the support [31, 32]. Hence, their multifractal spectrum is $\mathcal{D}_X(h) = 0$ if $\exists t : \ h_X(t) = h$, i.e. if h is a Hölder exponent of the process, and $-\infty$ elsewhere. Therefore, multifractal analysis is of no interest for their description.

2.4.4-b) Monofractal processes

Definition 2.13 (Monofractal function or process) *A function or processes $X(t)$ in $\mathbb{R}^d$ for which $h_X(t)$ is a constant, $\forall t : \ h_X(t) = H$, is called* monofractal*.*

The multifractal spectrum of monofractal processes reduces to:

$$\mathcal{D}_X(h) = \begin{cases} d & h = H \\ -\infty & h \neq H. \end{cases}$$

The fv H-sssi processes fBm and ROS (cf. Subsection 2.2.3) are monofractal processes and have Hölder exponent $h_X(t) = H$ everywhere.

2.4.4-c) Multifractal processes

Multifractal functions or processes are commonly defined as follows:

Definition 2.14 (Multifractal function or process) *A function or process $X(t)$ is called multifractal if it contains more than one Hölder exponent h that is living on a support with non-zero Hausdorff dimension.*

Technically, however, it is more precise to distinguish between functions or processes $X(t)$ that contain more than one Hölder exponent h, termed *multi-Hölder*, and functions or processes that do in addition contain more than one Hölder exponent that lives on a support with non-zero Hausdorff dimension [85].
All homogenous processes X whose Hölder exponents $h_X(t)$ are not constant fall into this category. Their Hölder exponents $h_X(t)$ are discontinuous everywhere and hence are highly variable and change widely from point to point, and from sample path to sample

path. In contrast to *multifractional* processes, multifractal processes have many different Hölder exponents on any subinterval, and are hence interesting for multifractal analysis. In this manuscript, we will only consider multifractal *multiplicative cascade* (MMC) processes [95, 121, 145], which represent the multifractal model processes that are quasi-exclusively used in practical applications. They are defined in Section 2.7.

2.5 Multifractal Formalism

A key practical issue consists in estimating the multifractal spectrum $\mathcal{D}_X(h)$ from a single finite length observation of data X. This can not be undertaken by direct application of Def. 2.11, since this would involve measuring $h_X(t)$ at each time instant t, which is practically intractable: The Hölder exponent $h_X(t)$ of multifractal processes $X(t)$ is discontinuous everywhere and hence practically inaccessible to direct local numerical determination. Because of the finite resolution of data, each sample of the observation represents a subinterval of t, which contains many different Hölder exponents of X (theoretically, all of them are present on any subinterval). One therefore has to resort to mathematical formulas which allow, under certain mathematical hypotheses, to obtain the spectrum $\mathcal{D}_X(h)$ from quantities that can be numerically calculated. Such formulas are called *multifractal formalisms*.
Recently, it has been shown that this can be achieved using the so-called *wavelet Leader multifractal formalism* [8, 89, 92, 110]. For it to be well defined, the measure or function $X(t)$ has to be *bounded*, which implies that the wavelet Leaders (defined below) are bounded. A sufficient condition for $X(t)$ to be bounded is that it is *uniform Hölder* (cf. e.g. [90, 178, 182]).

Definition 2.15 (Uniform Regularity) *A function or process $X(t)$ is said to belong to C^ε, $\varepsilon \in \mathbb{R}$, when its wavelet coefficients satisfy:*

$$\exists C > 0 : \ \ \forall j, k \ \ |d_X(j,k)| \leq C 2^{j\varepsilon}. \tag{2.52}$$

A *uniform regularity exponent* h_{min} can hence be defined:

$$h_{min} = \sup\{\varepsilon : \ X \in C^\varepsilon\}. \tag{2.53}$$

Definition 2.16 (Uniform Hölder) *A function or process $X(t)$ is said to be uniformly Hölder if $h_{min} > 0$.*

In turn, $h_{min} > 0$ implies:

$$\forall t_0, \ \ h(t_0) \geq h_{min},$$

and yields that $X(\mathbf{t})$ is uniform Hölder, hence is continuous, hence is bounded, hence possesses finite wavelet Leaders in the limit of fine scales. These implications are strict. Therefore, we assume throughout this manuscript that $X(t)$ is uniform Hölder.

The uniform Hölder condition may seem rather restrictive in practice. This is in particular the case for the multifractal analysis of images, since they consist of an intensity local average that can naturally be seen as the approximation, at a given resolution level, of a positive measure: There is hence a priori no guarantee that they are in the class of bounded functions. This practical limitation will be addressed in Section 5.1.

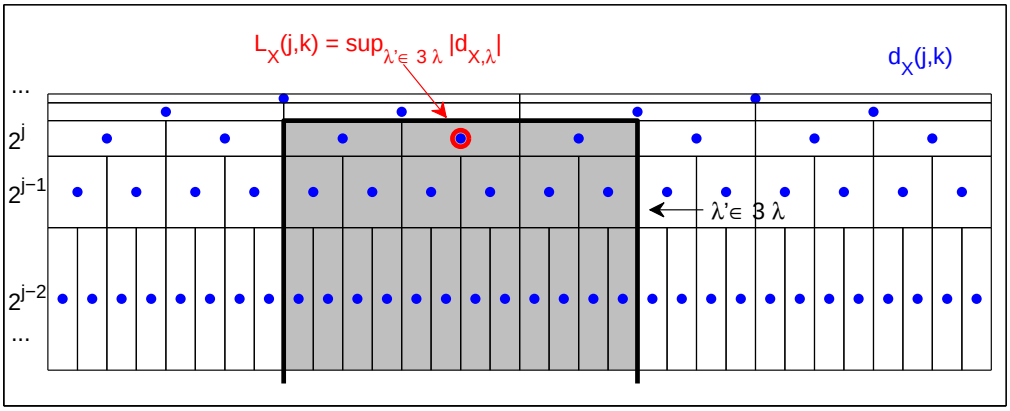


Figure 2.3: **Definition of Wavelet Leaders: 1d.** The wavelet Leader $L_X(j,k)$ ('○') is defined as the largest coefficient $|d_X(\cdot,\cdot)|$ ('•') within the time neighborhood $3\lambda_{j,k}$ over all finer scales $2^{j'} < 2^j$ (area in gray, truncated at fine scales).

2.5.1 Wavelet Leaders

Wavelet Leaders consist of multiresolution quantities that advantageously replace wavelet coefficients in multifractal analysis since they possess significant theoretical and practical qualities for the construction of a multifractal formalism.

In this manuscript, we only consider the analysis of 1d and 2d signals. Therefore, we only provide the theoretical definition of 1d and 2d wavelet Leaders. The general nd definition can be found in [89, 90, 92].

We assume that the mother wavelet $\psi_0(t)$ has a compact time support, and that the quadrature mirror filters $H_0(k), G_0(k)$ have finite impulse responses. This condition is fulfilled for the Daubechies bases used in this work [55].

Definition 1d. Let us define dyadic intervals as:

$$\lambda_{j,k} = \left[k2^j, (k+1)2^j\right).$$

Also, let 3λ denote the union of the interval λ with its 2 adjacent dyadic intervals:

$$3\lambda_{j,k} = \lambda_{j,k-1} \cup \lambda_{j,k} \cup \lambda_{j,k+1}.$$

Following [89], we define wavelet Leaders as:

$$L_X(j,k) = \sup_{\lambda' \subset 3\lambda} |d_{X,\lambda'}|. \tag{2.54}$$

This definition means that the wavelet Leader $L_X(j,k)$ consists of the largest wavelet coefficient $|d_X(j',k')|$ computed at all finer scales $2^{j'} \leq 2^j$ within a narrow time neighborhood, $(k-1)\cdot 2^j \leq 2^{j'}k' < (k+2)\cdot 2^j$. The definition is illustrated in Fig. 2.3.

Definition 2d. Let us introduce a dyadic indexing of squares as:

$$\lambda_{j,k_1,k_2} = \left\{[k_1 2^j, (k_1+1)2^j), [k_2 2^j, (k_2+1)2^j)\right\}.$$

The union of 9 such neighbor intervals is denoted as:

$$3\lambda_{j,k_1,k_2} = \bigcup_{n_1,n_2=\{-1,0,1\}} \lambda_{j,k_1+n_1,k_2+n_2}.$$

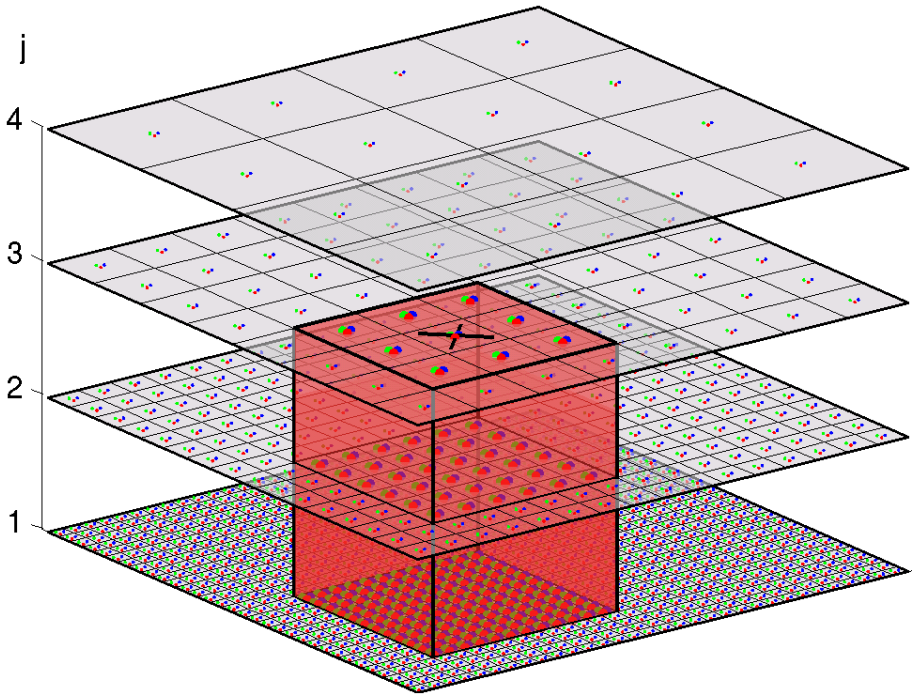


Figure 2.4: **Definition of Wavelet Leaders: 2d.** The wavelet Leader $L_X(j,k_1,k_2)$ at scale 2^j and position (k_1,k_2) (black cross) is defined as the largest of the wavelet coefficients $|d_X^{(m)}(j',k_1',k_2')|$, $m=1,\cdots,3$ ('•', '•' and '•') within a spatial neighborhood of (k_1,k_2) and within all finer scale $2^{j'}\leq 2^j$ (red volume, truncated at fine scales). The wavelet coefficients over which the supremum is taken are marked by fat dots.

2d wavelet Leaders are defined as [89]:

$$L_X(j,k_1,k_2)=\sup_{m,\,\lambda'\subset 3\lambda_{j,k_1,k_2}}|d_X^{(m)}(\lambda')|, \tag{2.55}$$

This definition is illustrated in Fig. 2.4 and means that the Leader $L_X(j,k_1,k_2)$ is obtained as the largest amongst the wavelet coefficients $|d_X^{(m)}(j',k_1',k_2')|$, $m=1,2,3$ existing in a (narrow) spatial neighborhood of (k_1,k_2), at any finer scale $2^{j'}\leq 2^j$.

Key properties. Under the uniform Hölder regularity condition, wavelet Leaders are *hierarchical quantities*, i.e., (by construction) monotonously increasing with analysis scale 2^j, and they accurately measure local Hölder exponents of the function f or sample path X. These are the central properties underlying their use as multiresolution quantities for multifractal analysis [89, 92]: If X is uniform Hölder and has Hölder exponent $h(t_0)\geq 0$ at t_0, then, on condition that $N_\psi > h$ and when $2^jk=t_0$:

$$L_X(j,k)\sim_{2^j\to 0}2^{jh(t_0)}, \tag{2.56}$$

where $X_a\sim_{a\to 0}Y_a$ means that $\lim_{a\to 0}X_a=\lim_{a\to 0}Y_a$. Therefore, the wavelet Leader exactly reproduces the Hölder exponent of X at t_0 by a local power law behavior in the limit of fine scales ($2^j\to 0$). The general validity of this local power law behavior, in all cases and for all processes, is the key property ensuring the validity and relevance of the multifractal formalism developed below. For the theoretical proof, see [89, 92].

2.5.2 Wavelet Leader Multifractal Formalism (WLMF)

Let us, for fixed analysis scales $a = 2^j$, form the time (space) averages of (the q-th powers of) the $L_X(j,k)$, referred to as the *structure functions*:

$$S^L(j,q) \quad = \quad \frac{1}{n_j}\sum_{k=1}^{n_j} L_X(j,k)^q. \tag{2.57}$$

Here, n_j denotes the number of $L_X(j,k)$ available at scale 2^j. Following intuitions originally developed in [76] for an increment based multifractal formalism, Eq. (2.56) above suggests that the wavelet Leader structure functions $S^L(j,q)$ possess power law behavior with respect to scales in the limit of small scales:

$$S^L(j,q) \simeq F_q 2^{j\zeta^L(q)},\ 2^j \to 0. \tag{2.58}$$

This relation establishes a clear and deep connection between the property of scale invariance and the analysis tool constituted by multifractal analysis: Indeed, Eq. (2.58) is highly reminiscent of Eq. (2.7), postulated as the (practical) definition of scale invariance, with wavelet Leaders as multiresolution quantities. Therefore, we call the function $\zeta^L(q)$ the scaling exponents[6] of X.

Legendre spectrum. The function $q \to \zeta^L(q)$ is necessarily concave (cf. [92]). Therefore, it can be replaced with its *Legendre transform* without loss of information, since this transformation is bijective for concave functions:

$$D^L(h) = \min_{q\neq 0}(d + qh - \zeta^L(q)). \tag{2.59}$$

The function $D^L(h)$ is called the *Legendre spectrum* of X.
Under the uniform Hölder regularity condition for X that is assumed to hold, it can be shown that $D^L(h)$ provides a tight upper bound for the multifractal spectrum $\mathcal{D}(h)$:

$$\mathcal{D}(h) \leq D^L(h) = \min_{q\neq 0}(d + qh - \zeta^L(q)). \tag{2.60}$$

Multifractal formalism. The *wavelet Leader multifractal formalism* (WLMF) asserts that this inequality turns into an equality:

$$\forall h, \quad \mathcal{D}(h) = D^L(h) = \min_{q\neq 0}(d + qh - \zeta^L(q)), \tag{2.61}$$

and therefore that the Legendre spectrum of X can be interpreted in terms of the Hölder singularities of X.
Eq. (2.61) constitutes the relation that is used in practice for the numerical calculation of the multifractal spectrum $\mathcal{D}(h)$: In applications, one tries to estimate the Legendre spectrum $D(h)$, which is the only quantity that is numerically accessible. Estimation procedures are defined in Section 2.6.

[6] Note that in the context of theoretical multifractal analysis, $\zeta^L(q)$ is generally referred to as the *scaling function*.

2.5.3 Multifractal analysis versus scaling analysis

Strictly speaking, multifractal analysis aims at measuring the multifractal spectrum $\mathcal{D}(h)$, while scaling analysis (cf. Def. 2.2) rather concentrates on scaling exponents $\zeta(q)$. Because of the general validity of the WLMF Eq. (2.61) for most commonly used scaling processes (fv H-sssi and MMC processes), the scaling exponents and the multifractal spectrum are closely related one to the other via a Legendre transform. Hence, the power law behaviors in Eqs. (2.7) and (2.58) together with Eq. (2.61) constitute the fundamental relations establishing the connection between scale invariance and multifractality. The multifractal formalism explains why the terms multifractal analysis and scaling analysis are often used one for the other in practical situations.

2.5.4 Validity of the WLMF

Obviously, the WLMF relation Eq. (2.61) is not valid *in general*, for all functions and all realizations of stochastic processes: A first necessary condition is that the multifractal spectrum $\mathcal{D}(h)$ of $X(t)$ is a concave function, and a second necessary condition is that $X(t)$ is a bounded function, $X \in L^0$, such that wavelet Leaders are finite.
The conditions of validity of Eq. (2.61) are, in fact, not very well known. Nevertheless, it has been proven that the WLMF Eq. (2.61) is strictly valid for most classes of functions and processes $X(t)$ practically used in applications. In particular, it is exact for finite variance H*-sssi processes* such as fBm or ROS (cf. Section 2.2.3) [92]. For advanced discussions of this question, the reader is referred to [89, 90, 92, 178]

Legendre spectrum in applications. In applications, one tries to estimate the Legendre spectrum, without being in general able to determine the relevance and validity of its interpretation in terms of a multifractal spectrum. In what follows, we will not distinguish any more between the multifractal spectrum $\mathcal{D}(h)$ of $X(t)$ and its Legendre spectrum $D^L(h)$ for estimation procedures and in applications, and refer to the Legendre spectrum $D^L(h)$ as the multifractal spectrum of $X(t)$.
Let us note that regardless of its interpretation, the Legendre spectrum constitutes an important quantity in applications. It can be shown that for $X(t)$ uniform Hölder, the Legendre spectrum $D^L(h)$ is *independent* of the precise choice of mother wavelet $\psi_0(t)$ (on condition that N_ψ is larger than the largest singularity exponent encountered in $X(t)$, cf. [92] for a detailed account). This shows that it constitutes a quantity that is intimately related to properties of $X(t)$ and can hence be used in applications for detection, classification and identification tasks, without explicit reference to the validity of the WLMF.

2.5.5 Other multifractal formalisms

Wavelet coefficient based multifractal formalism (WCMF). Previous multifractal formalisms were based on increments or wavelet coefficients instead of wavelet Leaders. They are obtained by replacing wavelet Leaders L_X in Eq. (2.57) (and subsequently

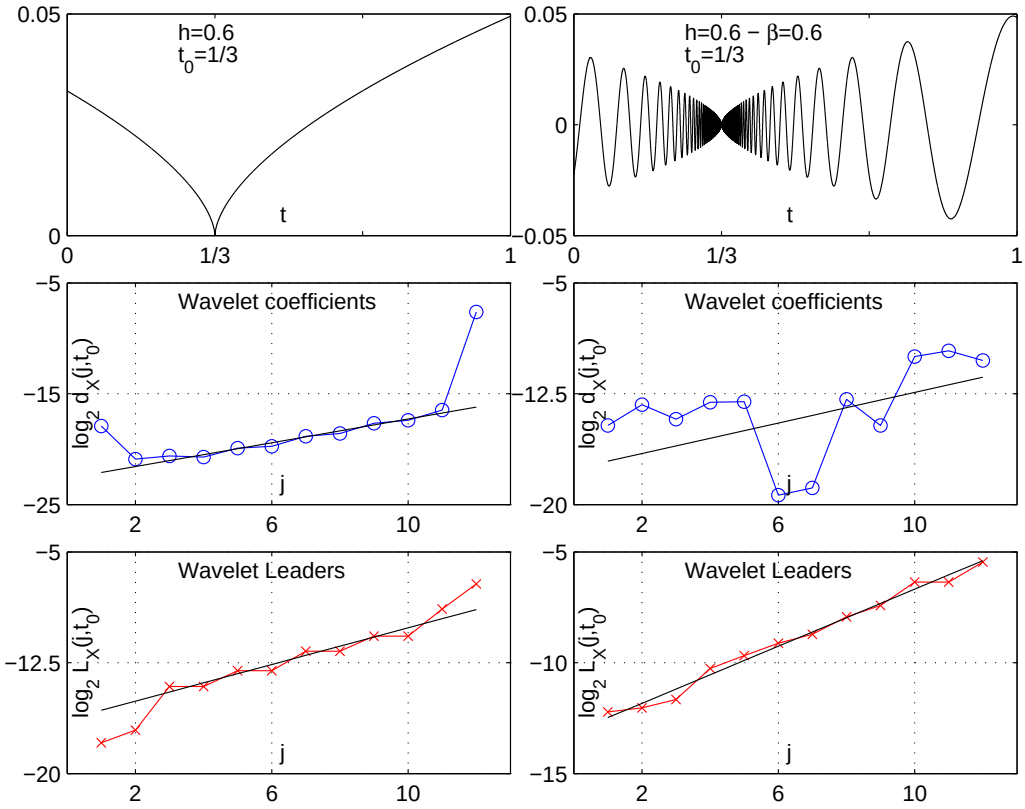


Figure 2.5: **Analysis of cusp-type and oscillating singularities.** Analysis of cusp-type (left) and chirp-type (right) singularities (top row) with wavelet coefficients (center row) and Leaders (bottom row): Whereas wavelet Leaders reproduce the Hölder exponent by a power law $L_X(j,t_0) \sim C2^{jh(t_0)}$ for both the cusp- and the chirp-type singularity, this is not the case for wavelet coefficients for the chirp-type singularity.

through Eqs. (2.58) to Eqs. (2.61)) with wavelet coefficients d_X:

$$S^d(j,q) = \frac{1}{n_j}\sum_{k=1}^{n_j}|d_X(j,k)|^q \simeq F_q 2^{j\zeta^d(q)},\ 2^j \to 0, \tag{2.62}$$

$$D^d(h) = \min_{q\neq 0}(d+qh-\zeta^d(q)), \tag{2.63}$$

$$\mathcal{D}(h) = D^d(h) \quad \text{(WCMF)}. \tag{2.64}$$

In contrast to the WLMF, the WCMF does not hold for all multiplicative martingale processes and fails to provide the practitioners with a correct analysis of the entire multifractal spectrum:

1. First, in contrast to wavelet Leaders, for which Eq. (2.56) is of general validity, wavelet coefficients are not appropriate for measuring Hölder exponents for oscillating singularities, because they are not hierarchical quantities. This is illustrated in Fig. 2.5 (see also [110])[7]. Therefore, the WCMF can not hold for processes containing such singularities.

[7] Lashermes [110] actually provides the first practical illustration for the case of general singularities. He

2. Second, since wavelet coefficients are sparse quantities and mainly centered around the zero value, the WCMF suffers from a major practical drawback since structure functions Eq. (2.62) are numerically unstable for negative statistical orders q. Consequently, they can not be used for the estimation of the decreasing part of the multifractal spectrum, which is associated with negative orders q via the multifractal formalism. In contrast, wavelet Leaders are, for X uniform Hölder, practically always positive and do allow the exploration of positive *and negative* statistical orders q, hence of the entire multifractal spectrum (cf. [110] for illustrations, Chapter 3 for a statistical characterization of wavelet Leaders, and Chapter 4 for estimation performance comparisons between the WLMF and the WCMF).

Historically, the first multifractal formalism was proposed in [76], based on increments $T_X(a,t) = \delta_{a\tau_0}X(t)$ (cf. Eq. (2.13)) as multiresolution quantities. Increments can be seen as some sort of wavelets, referred to as poor man's wavelets, with specific mother wavelet $\psi_0(t) = \delta(t+\tau_0) - \delta(t)$. Higher order increments are obtained by taking increments of increments, and the corresponding mother wavelets by convolutions of ψ_0 with itself. Therefore, the increment based multifractal formalism suffers from the same drawbacks, stated in 1. and 2. above, as the WCMF. In addition, the specific mother wavelets have very poor frequency resolution. Despite these major disadvantages, the increment based multifractal formalism is still used in practice by physicists.

Wavelet coefficient modulus maxima based multifractal formalism (MMMF). Another multifractal formalism, based on the Modulus Maxima Wavelet Transform (MMWT), has been previously proposed to resolve the numerical instability issue for negative q of the WCMF, initially for the analysis of 1d signals [24, 117, 130]. It is also commonly used in the context of turbulence, for instance (see e.g., [19, 24, 129]). The solution relies on the use of the coefficients of a continuous wavelet transform, from which a skeleton, consisting of maxima along scales lines, is extracted. The wavelet coefficients living on this skeleton are then involved in the computation of the structure functions. This technique has also been extended to image analysis, at the price of significant computational (2d-CWT + 2d-skeleton) and conceptual (maxima lines become maxima manifolds) complexities [16, 17].
On top of the practical difficulties related to its implementation and its computational cost, the MMMF, be it 1d or 2d, despite its showing satisfactory experimental results, is still lacking a theoretical mathematical support, such as provided for the WLMF in [89, 92].
We will not study the MMMF in this manuscript. For more details, the reader is referred to [16, 17, 24, 117].

2.6 Multifractal Analysis: Estimation

2.6.1 Moments - scaling exponents

Eqs. (2.7) and (2.58) above suggest that the scaling exponents can be estimated by linear regressions in log-log coordinates, i.e. by linear regressions of $Y(j,q) = \log_2 S^L(j,q)$ vs.

considers the case $t = 1/2$, which is the only point that falls exactly on the dyadic grid on all scales. Fig. 2.5 demonstrates this for $t = 1/3$, which is not on the dyadic grid for finite length observations (and hence potentially more difficult). It confirms that Leaders reproduce Hölder exponents for both cusp and chirp-type oscillating singularities, while coefficients do not.

$j = \log_2 2^j$ over the range of scales $j \in [j_1, j_2]$ for which scale invariance is evidenced:

$$\hat{\zeta}^L(q) = \sum_{j=j_1}^{j_2} w_j \log_2 S^L(j,q), \tag{2.65}$$

Weights in linear regressions. 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 $w_j = b_j \frac{V_0 j - V_1}{V_0 V_2 - V_1^2}$ with $V_i = \sum_{j_1}^{j_2} j^i b_j$, $i = 0, 1, 2$. The positive numbers b_j are freely selectable and reflect the confidence granted to each $Y(j,q) = \log_2 S^L(j,q)$.
In the present work, we perform either non-weighted, ordinary linear fits ($b_j = 1$, denoted by w^0), or weighted fits, following [4]: $b_j = n_j$ (denoted by w^1), where n_j is the number of coefficients at scale j. These weighted w^1 regressions are based on Gaussian and independence assumptions for the coefficients. Alternative choices are reported in [173, 180]. The relative performance of w^0 and w^1 is studied in Section 4.1.4.

2.6.2 Cumulants

The structure functions Eq. (2.57) consist of time averages and can be read as sample mean estimators for the ensemble averages $\mathbb{E}L_X(j,\cdot)^q$. This heuristic analysis was first proposed using increments as multiresolution quantities in [44] and further developed for continuous wavelet coefficients in [60]. Here, we further extend this interpretation to wavelet Leaders. Hence, Eq. (2.58) is rewritten as:

$$\mathbb{E}L_X(j,\cdot)^q = F_q 2^{j\zeta^L(q)}. \tag{2.66}$$

For the range of qs (necessarily including $q = 0$) where $\mathbb{E}L_X(j,\cdot)^q < \infty$, a standard generating function expansion yields:

$$\ln \mathbb{E} e^{q \ln L_X(j,\cdot)} = \sum_{p=1}^{\infty} C^L(j,p) \frac{q^p}{p!}, \tag{2.67}$$

where the $C^L(j,p)$ stand for the cumulants of order $p \geq 1$ of the random variables $\ln L_X(j,\cdot)$. Combining Eqs. (2.66) and (2.67) compels the $C^L(j,p)$ to take the following scale dependence:

$$\forall p \geqslant 1, C^L(j,p) = c^L_{0,p} + c^L_p \ln 2^j, \tag{2.68}$$

and thus implies:

$$\ln \mathbb{E} e^{q \ln L_X(j,\cdot)} = \underbrace{\sum_{p=1}^{\infty} c^L_{0,p} \frac{q^p}{p!}}_{\ln F_q} + \underbrace{\sum_{p=1}^{\infty} c^L_p \frac{q^p}{p!}}_{\zeta^L(q)} \ln 2^j. \tag{2.69}$$

Hence,

$$\zeta^L(q) = \sum_{p=1}^{\infty} c^L_p \frac{q^p}{p!} = c^L_1 q + c^L_2 \frac{q^2}{2} + c^L_3 \frac{q^3}{6} + \cdots, \tag{2.70}$$

and the knowledge of $\zeta^L(q)$ (and therefore also of $D^L(h)$, cf. Section 2.6.3-c) below) can be rephrased in terms of the coefficients c^L_p, called the *log-cumulants*.

Eq. (2.68) suggests that estimations can be performed by means of linear regressions in $\ln 2^j$ vs. $C^L(j,q)$ coordinates:

$$\hat{c}_p^L = (\log_2 e) \cdot \sum_{j=j_1}^{j_2} w_j \hat{C}^L(j,p), \tag{2.71}$$

where the estimates $\hat{C}^L(j,p)$ for the cumulants of $\ln L_X(j,\cdot)$ are obtained from standard sample cumulant estimators (see, for instance, [97]):

$$\hat{C}^L(j,p) = \hat{m}^L(j,p) - \sum_{n=1}^{p-1} \binom{p-1}{n-1} \hat{C}^L(j,n)\hat{m}^L(j,n-k). \tag{2.72}$$

Here, the $\hat{m}^L(j,p) = \frac{1}{n_j}\sum_{k=1}^{n_j} \ln L_X(j,\cdot)^p$ are the sample raw moments of $\ln L_X(j,\cdot)$. The triplet (c_1^L, c_2^L, c_3^L) gathers most of the information, related to $\zeta^L(q)$, practically available from empirical data: c_1^L measures the linear part of $\zeta^L(q)$ ($c_1^L = H$ for H-sssi processes, and c_1^L plays the equivalent role of H for MMC processes), c_2^L the first departure from linearity, and c_3^L and higher terms a more complicated departure of $\zeta^L(q)$ from linearity. For applications, such a set of attributes is highly advantageous: For practical purposes such as detection or classification of data from multifractal features, manipulating and comparing the functions $\zeta^L(q)$ is not comfortable. For instance, detecting departure from linearity of $\zeta^L(q)$ would be intricate when based directly on $\zeta^L(q)$, since it would have to involve multiple estimates of $\zeta^L(q)$, which are likely to be strongly dependent. In contrast, this can in first approximation be meaningfully cast into the detection of departure of c_2 from the zero value. Approximation of $\zeta^L(q)$ and $D^L(h)$ with a limited number of c_p^L can hence significantly simplify the classification or detection tasks based on MF attributes.

We note, however, that the expansion in Eq. (2.70) is not of strict general validity for all processes. Notably, one condition is that the derivatives $d^p/dq^p\zeta^L(q)$ of $\zeta^L(q)$ exist in the neighborhood of $q = 0$, which is not fulfilled for all processes. One example for which the log-cumulant development of $\zeta^L(q)$ is not valid are the log-stable cascades, used as a model for turbulence intermittency, for which $\zeta^L(q) = \alpha\frac{1-q^\alpha}{1-q}$, $0 < \alpha < 2$ [155]. Another example is given by α-stable motion, for which $\zeta^L(q) = qH$, but only for the range of statistical orders $-1 < q < \alpha$, $0 < \alpha < 2$ (see e.g. [153]). We choose to exclude these types of models and to restrict ourselves, throughout this manuscript, only to processes for which Eq. (2.70) is valid up to at least $p = 3$.

2.6.3 Multifractal spectrum

2.6.3-a) Legendre transform

Eq. (2.61) indicates that estimations of the multifractal spectrum $D^L(h)$ can be obtained by numerical Legendre transform of the scaling exponent estimates $\hat{\zeta}^L(q)$:

$$\hat{D}^L(h) = \min_{q\neq 0}(d + qh - \hat{\zeta}^L(q)). \tag{2.73}$$

For some practical purposes, however, estimates Eq. (2.73) are not very useful. The numerical Legendre transform of the estimates $\hat{\zeta}^L(q)$ for a certain discrete range of orders

q does not enable to assign one specific position in the spectrum $D^L(h)$ to each value $\hat{\zeta}^L(q)$. This is, for instance, very limiting for the design of nonparametric confidence intervals (cf. Chapter 8). Therefore, it is in practice often advantageous to resort to parametric estimators for $D^L(h)$.

2.6.3-b) Direct determination of the multifractal spectrum

Chhabra et al. [49] proposed a method for direct determination of the multifractal spectrum that is obtained without explicit numerical Legendre transformation. The parametric form $(D^L(q), h^L(q))$ of the spectrum is estimated by linear regression of $U(j,q)$ and $V(j,q)$ versus j:

$$\hat{D}^L(q) = 1 + \sum_{j=j_1}^{j_2} w_j U(j,q), \tag{2.74}$$

$$\hat{h}^L(q) = \sum_{j=j_1}^{j_2} w_j V(j,q), \tag{2.75}$$

where

$$U(j,q) = \log_2 n_j + \sum_{k=1}^{n_j} R^q(j,k) \log_2 R^q(j,k), \tag{2.76}$$

$$V(j,q) = \sum_{k=1}^{n_j} R^q(j,k) \log_2 L(j,k), \tag{2.77}$$

$$R^q(j,k) = \frac{L(j,k)^q}{\sum_{k=1}^{n_j} L(j,k)^q}. \tag{2.78}$$

The derivation of Eqs. (2.74-2.78) is sketched in [49] and has been fully detailed in Appendix B. This parametric form for estimation of $D^L(h)$ is mainly used in this manuscript.

2.6.3-c) A log-cumulant expansion of the multifractal spectrum

The multifractal spectrum admits a polynomial expansion around its maximum, parametrized by the log-cumulants c_p^L:

$$D^L(h) = d + \frac{c_2^L}{2!}\left(\frac{h-c_1^L}{c_2^L}\right)^2 + \frac{-c_3^L}{3!}\left(\frac{h-c_1^L}{c_2^L}\right)^3 + \frac{-c_4^L + 3{c_3^L}^2/c_2^L}{4!}\left(\frac{h-c_1^L}{c_2^L}\right)^4 + \dots \tag{2.79}$$

This result constitutes an original contribution of this work. The derivation is detailed in Appendix A. The result shows that at first order, $p = 2$, $D^L(h)$ is approximated as a parabola: $D^L(h) = d - (h - c_1^L)^2/(2c_2^L)$. The log-cumulants c_p^L therefore admit a meaningful interpretation: c_1^L corresponds to the location of the maximum of $D^L(h)$, c_2^L to its width, while c_3^L is an asymmetry parameter and $c_4^L - 3{c_3^L}^2/c_2$ a flatness term.
The reduced set of parameters $\{c_1^L, c_2^L, c_3^L, c_4^L\}$ can therefore be regarded as a relevant approximation of $D^L(h)$ and hence as a meaningful and relevant summary of the multifractal properties of X.

2.6.4 Coefficient and increment based estimation

The estimations of $\zeta(q)$, $D(h)$ and c_p can also be based on increments or wavelet coefficients, mutatis mutandis, by replacing wavelet Leaders with increments or wavelet coefficients in the estimation procedures Eq. (2.65), Eqs. (2.71–2.72) and Eqs. (2.74–2.78). We use the superscripts ^{I}, ^{d} and ^{L} to distinguish between increment, wavelet coefficient and Leader based estimation, respectively. When it is clear from the context on which quantities estimation is based, or when a distinction is irrelevant, we drop these superscripts. The respective performance of these estimation procedures will be investigated in Chapter 4 and Section 5.3.

2.7 Multifractal Processes: Multiplicative Cascade Processes

In this section, we introduce some of the most commonly used multifractal processes and discuss their properties. The most famous constructions of multifractal processes are based on multiplicative martingales [95, 121, 145]. They are commonly referred to as multiplicative cascade processes. In this manuscript, we choose to call them *multifractal multiplicative cascade* (MMC) processes, to further emphasize their difference with respect to (monofractal) fv H-sssi processes. Indeed, the starting point for the construction of MMC processes can be seen in the need for a class of processes with richer scale invariance and multifractal properties than H-sssi processes. Whereas for H-sssi processes, everything is controlled by one single parameter H, this central role is taken by a function $\varphi(q)$ for MMC processes, which accounts for more flexibility. It depends on the precise definition of the cascade process and characterizes both the scale invariance and multifractal properties of the process.
For an overview of MMC processes, the reader is referred to, e.g., [45].

2.7.1 Cascades

2.7.1-a) Canonical Mandelbrot Cascades (CMC)

The multiplicative cascades of Mandelbrot (CMC) constitute the archetype of multifractal processes and have for long been the only example of multifractal processes practically available. They have been introduced by Yaglom [186], who tried to mimic the energy transfer in turbulence phenomena following the celebrated energy cascade based heuristic analysis of turbulence flows by Richardson [144]. Their construction is based on an iterative split-and-multiply random procedure on an interval, which matches the physical intuitions beyond the vorticity stretching mechanisms at work in turbulence flows. In the 1970s, Mandelbrot studied the properties of these models and fruitfully gathered them in the unified framework of multiplicative martingales [121, 145].
Definition. CMC are constructed by the following iterative procedure. The construction starts from a uniform unit mass on the interval $]0,1[$. First, the interval is cut into two intervals of equal size[8], $]0,\frac{1}{2}[$ and $]\frac{1}{2},1[$. Then a random mass is attributed to each of the intervals by multiplying the original masses with random multiplier $W_{11}, W_{12} \overset{i.i.d}{\sim} p_W$, respectively, i.e., the interval on the left has mass $\frac{1}{2}W_{11}$, and the interval on the right $\frac{1}{2}W_{12}$.

[8] We concentrate on the binary case, which we use as model processes in Chapters 4 and 8. More generally, CMC can be defined as a b-ary cascade, $b \geq 2$.

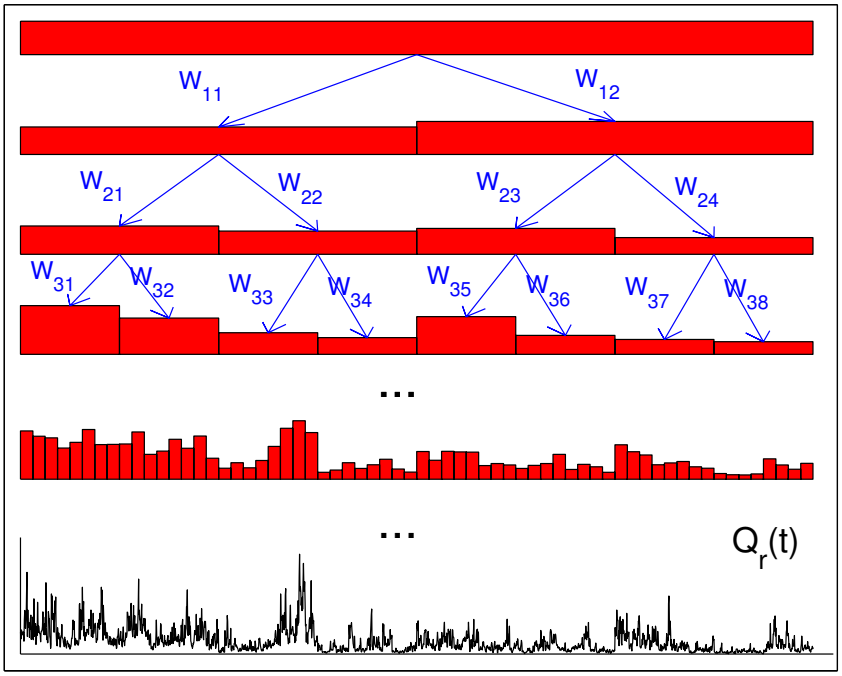


Figure 2.6: **Canonical Mandelbrot Cascade.** Iterative construction principle of CMC and the resulting $Q_r(t)$ (Eq. (2.80)) after 12 iterations.

Now this procedure is applied to both intervals: both are cut in two equal parts, and the mass of each of the resulting subintervals is multiplied with a multiplier, which is drawn i.i.d. from p_W for each subinterval. Then the procedure is iterated on the subintervals, and so on and so forth. The construction of CMC is illustrated in Fig. 2.6.
The multipliers have to be strictly positive random variables, $W > 0$, and they have to satisfy the constraint $\mathbb{E}W = 1$, ensuring that the cascade conserves mass in average.
Let the multipliers for the jth iteration be denoted by W_{jk}, $k = 1, \cdots, 2^j$. The measure $Q_r(t)$ at resolution $r = 2^{-J}$ (i.e., after J iterations) is defined as:

$$Q_r(t) = \prod_{j=1,\cdots,J,\,\{k:\,t\in[2^j k, 2^j(k+1)[\}} W_{jk}. \tag{2.80}$$

Extension to higher dimension. The definition of CMC, given above for an interval on the real line $\mathbb{R}$, can easily be extended to intervals in higher dimensions d, i.e. hypercubes in $\mathbb{R}^d$. The iterative construction principle remains the same: Starting from a uniform unity mass on $[0,1]^d$, the unit hypercube is divided into 2^d sub-hypercubes with side length $\frac{1}{2}$, 2^d multipliers $W \overset{i.i.d.}{\sim} p_W$ are drawn and determine the mass of each sub-hypercube. Then the operation is iterated for each of the sub-hypercubes, and so on and so forth.
Motion. Scaling and multifractal properties. Let us define the function:

$$\varphi(q) = -\log_2 \mathbb{E}W^q, \tag{2.81}$$

which has the following properties: It is concave in q, and it satisfies: $\varphi(0) = 0$, $\varphi(1) = 0$ (conservation of mass). Kahane and Peyrière [95] showed that on condition that:

$$\varphi'(1^-) > -d, \tag{2.82}$$

the limit:

$$A(t) = \lim_{r \to 0} \int_0^t Q_r(u) du. \tag{2.83}$$

is a well-defined process. It is referred to as the CMC *motion*. From a practical point of view, this limit can of course never be reached numerically, since the iteration process has to be stopped after finite time, hence for $r > 0$. One supposes then that the number of iterations is large enough to observe the same properties on the synthesized realizations as one would observe on the limit process $A(t)$.
$A(t)$ exhibits scaling properties of the form of Eqs. (2.7) and (2.58) for $q_c^- < q < q_c^+$ and for the discrete scaling factors $a = 2^j$:

$$\frac{1}{n_a} \sum_{k=1}^{n_a} L_A(a,k)^q \simeq c_q |a|^{\zeta^L(q)}, \tag{2.84}$$

$$\text{with } \zeta^L(q) = \lambda(q) \quad \text{for } q \in [q_*^-, q_*^+], \tag{2.85}$$

$$\text{where } \lambda(q) = dq + \varphi(q). \tag{2.86}$$

Outside the range $[q_*^-, q_*^+]$, $\zeta^L(q)$ behaves as a linear function of q. The critical orders q_*^- and q_*^+ are defined in Section 2.8 (Eqs. (2.129) and (2.130)). The range of finite moments $[q_c^-, q_c^+]$ is defined in Section 2.7.1-d) (Eqs. (2.114) and (2.115)).
The multifractal spectrum of $A(t)$ can be derived from the Legendre transform $D(h) = \min_{q \neq 0}(1 + qh - \lambda(q))$ of $\lambda(q) = dq + \varphi(q)$ as:

$$\mathcal{D}(h) = \begin{cases} D(h), & \text{if } D(h) \geq 0, \\ -\infty, & \text{otherwise.} \end{cases} \tag{2.87}$$

One commonly associates a multifractal spectrum $\mathcal{D}_Q(h)$ with the limit measure $Q_r(t)$, $r \to 0$:

$$\mathcal{D}_Q(h) = \mathcal{D}(h + d). \tag{2.88}$$

In this work, we will use binary CMC as (easy to generate and control) examples of multifractal images (i.e., processes in 2d). We will use two specific choices of multipliers W, namely log-Normal (CMC-LN) and log-Poisson (CMC-LP) multipliers, which are commonly used in applications, notably in hydrodynamic turbulence (see e.g. [45]).
Example 1: CMC-LN. For the construction of CMC-LN, log-Normal multipliers $W = 2^{-U}$ are used, i.e. $U \sim \mathcal{N}(m, \sigma)$ is Gaussian with mean m and variance σ^2, and therefore: $\varphi(q) = mq - \sigma^2 \ln(2)/2q^2$. Conservation of mass, $\mathbb{E}W = 1$, implies $\sigma^2 = 2m/\ln(2)$ and thus:

$$\varphi(q) = mq(1-q) \tag{2.89}$$

$$\zeta^L(q) = mq(1-q) + dq \quad \text{for } q \in [q_*^-, q_*^+] \tag{2.90}$$

$$D^L(h) = d - \frac{(h-d-m)^2}{4m} \tag{2.91}$$

$$c_1 = m + d, \quad c_2 = -2m \tag{2.92}$$

$$\forall p \geq 3 : c_p \equiv 0. \tag{2.93}$$

Hence, CMC-LN is seen as a simple multifractal process, since departure of its $\varphi(q)$ and $\zeta(q)$ from a linear behavior in q simplifies to a 2nd order polynomial form. Departure

from linearity in q is hence fully controlled by the single parameter c_2, and the multifractal spectrum consists of a simple parabola.

Example 2: CMC-LP. The construction of CMC-LP uses log-Poisson multipliers $W = 2^\gamma \exp\left(\ln(\beta)\pi_\lambda\right)$, where π_λ is a Poisson random variable with parameter $\lambda = -\frac{\gamma \ln(2)}{(\beta-1)}$. Its scaling and multifractal characteristics are therefore given by:

$$\varphi(q) = -\gamma q + \frac{\gamma(\beta^q - 1)}{\beta - 1} \tag{2.94}$$

$$\zeta^L(q) = (d-\gamma)q + \frac{\gamma(\beta^q - 1)}{\beta - 1} \quad \text{for } q \in [q_*^-, q_*^+] \tag{2.95}$$

$$D^L(h) = d + \frac{\gamma}{\beta - 1} + \frac{\gamma + h - d}{\ln \beta} \cdot \left[\ln\left(\frac{(\gamma + h - d)(\beta - 1)}{\gamma \ln \beta}\right) - 1\right] \tag{2.96}$$

$$c_1 = \gamma\left(\frac{\ln(\beta)}{\beta - 1} - 1\right) + d \tag{2.97}$$

$$c_p = -\frac{\gamma}{\beta - 1}\left(-\ln(\beta)\right)^p, \quad p \geq 2. \tag{2.98}$$

Further examples. In this manuscript, we choose to use the two specific CMC constructions CMC-LN and CMC-LP described above as representatives of simple ($c_3 \equiv 0$) and more elaborated ($c_3 \neq 0$) multifractal models. We note that other constructions have been proposed in the literature, such as the log-stable cascades [155], which have been widely used as model processes in the context of hydrodynamic turbulence. They will not be considered in this manuscript. For more details on other examples and possible choices of multipliers, we refer to e.g. [45] and references therein.

Limitations. The CMC construction suffers from important practical drawbacks that have been recognized since their first uses:

- The split/multiply construction results in discrete scale invariance, such that only specific dilation factors $a = 2^j$ can be used in Eqs. (2.7) and (2.58),
- increments are not stationary,
- $Q_r(t)$ is only positive valued, and $A(t)$ has only positive variation. Hence, they are inappropriate models for many physical signals.

CMC remained, however, the only known multifractal synthetic reference processes used in applications up to the mid 90s, when several approaches were proposed to overcome these limitations. One of these propositions consists in constructing continuous time multifractal cascades, involving infinitely divisible stochastic integrals. The reader is referred to [156] for the precise definition and further details. In this manuscript, we choose to concentrate on another approach, which was proposed in [30].

2.7.1-b) Compound Poisson Cascades (CPC)

Following intuitions by Barral and Mandelbrot [30] and theoretical developments by Barral [29], compound Poisson cascades (CPC) were proposed to overcome some of the drawbacks of CMC. In short, the construction of compound Poisson cascades replaces the deterministic geometric grid underlying the construction of Mandelbrot's multiplicative cascades – responsible for discrete scale invariance and non stationary increments –

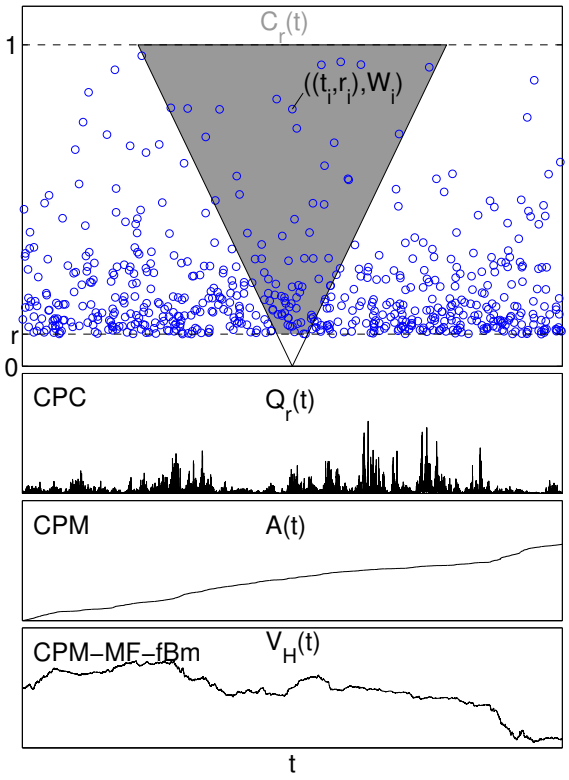


Figure 2.7: **Compound Poisson Cascade.** Illustration of the construction of a compound Poisson cascade (top line), one realization of $Q_r(t)$ (second line), and one realization of compound Poisson motion $A(t)$ (third line) and of fractional Brownian motion in multifractal time $V_H(t)$ (bottom).

with a random (Poisson point process) grid. Compound Poisson cascades are defined as:

$$Q_r(t) = C \prod_{(t_i,r_i)\in\mathcal{C}_r(t)} W_i, \quad r > 0, \tag{2.99}$$

where

$$\mathcal{C}_r(t) = \{(t', r') : r \le r' \le 1, t - r'/2 \le t' \le t + r'/2\} \tag{2.100}$$

is a cone, (t_i, r_i) consists of a 2d Poisson point process on the rectangle $I = \{(t', r') : r \le r' \le 1, -1/2 \le t' \le T + 1/2\}$ with intensity measure $dm(t, r)$, W_i are positive i.i.d. multipliers associated with points (t_i, r_i), and $C = C(r, t)$ is a normalizing constant such that $\mathbb{E}Q_r(t) = 1$. The construction of CPC is illustrated in Fig. 2.7.
It can be shown that CPC satisfy the following key resolution equation:

$$\mathbb{E}Q_r(t)^q = \exp\left(-\varphi(q)m(\mathcal{C}_r(t))\right), \tag{2.101}$$

with $m(\mathcal{C}_r(t)) = \int_{\mathcal{C}_r(t)} dm(t', r')$ and:

$$\varphi(q) = c((1 - \mathbb{E}W^q) - q(1 - \mathbb{E}W)), \tag{2.102}$$

where c is an arbitrary positive constant. If $dm(t, r) = g(r)drdt$, the process $Q_r(t)$ (and therefore the increments of the motion processes $A(t)$ and $V_H(t)$ defined below) are

stationary. In this manuscript, we only consider the choice[9] proposed in [22, 23]:

$$g(r)dr = c(dr/r^2 + \delta_{\{1\}}(dr)), \tag{2.103}$$

where $\delta_{\{1\}}(dr)$ denotes a point mass at $r = 1$. Then, Eq. (2.101) simplifies to the power law:

$$\mathbb{E}Q_r(t)^q = r^{\varphi(q)}, \tag{2.104}$$

and it can also be shown that the correlation is given by [46, 47]:

$$\mathbb{E}Q_r(t)Q_r(s) = |t-s|^{\varphi(2)},\ |t-s| > r. \tag{2.105}$$

Compound Poisson Motion (CPM). On condition that $\varphi(1^-) \geq -1$, *compound Poisson motion* (CPM) $A(t)$ is a well-defined process:

$$A(t) = \lim_{r\to 0} \int_0^t Q_r(s)ds. \tag{2.106}$$

With the choice Eq. (2.103), the increments of $A(t)$ are stationary, and CPM exhibits scaling properties of the form of Eqs. (2.7) and (2.58) [46]:

$$\mathbb{E}L_A(a,t)^q \simeq c_q|a|^{\lambda(q)}, \tag{2.107}$$

$$\text{with } \lambda(q) = q + \varphi(q), \tag{2.108}$$

for $q_c^- < q < q_c^+$ (cf. Section 2.7.1-d), Eqs. (2.114) and (2.115)). From the results proven in [30], we can infer that the multifractal spectrum $\mathcal{D}(h)$ of $A(t)$ can be derived from the Legendre transform $D(h) = \min_{q\neq 0}(1 + qh - \lambda(q))$ of $\lambda(q)$ as:

$$\mathcal{D}(h) = \begin{cases} D(h), & \text{if } D(h) \geq 0, \\ -\infty, & \text{otherwise.} \end{cases} \tag{2.109}$$

As for CMC, one commonly associates a multifractal spectrum $\mathcal{D}_Q(h)$ with the limit measure $Q_r(t)$, $r \to 0$:

$$\mathcal{D}_Q(h) = \mathcal{D}(h+1). \tag{2.110}$$

The scaling exponents are given by:

$$\zeta^L(q) = \lambda(q) \quad \text{for } q \in [q_*^-, q_*^+], \tag{2.111}$$

and behave as a linear function of q beyond the range $[q_*^-, q_*^+]$ (cf. Section 2.8, Eqs. (2.129) and (2.130)).

Example: Log-Normal multipliers. In this work, we will only consider CPC with log-Normal multipliers $W = \exp(Y)$, where each $Y \sim \mathcal{N}(\mu, \sigma)$ is an independent Gaussian random variable (r.v.) with mean μ and variance σ^2. Then, the form of $\varphi(q)$:

$$\varphi(q) = c\left[\left(1 - \exp\left(\mu q + \frac{\sigma^2}{2}q^2\right)\right) - q\left(1 - \exp\left(\mu + \frac{\sigma^2}{2}\right)\right)\right], \tag{2.112}$$

[9] Other choices are possible, leading to cascades with different properties, cf. e.g. [46].

is such that the cumulant expansion of $\zeta^L(q)$ yields a non zero c_3 [46, 47]. More precisely, the log-cumulants of CPM have the following explicit expressions:

$$\begin{aligned} c_1 &= 1 - c\left(\mu + 1 - \exp\left(\mu + \frac{\sigma^2}{2}\right)\right) \\ c_2 &= -c(\mu^2 + \sigma^2) \\ c_3 &= -c(\mu^3 + 3\mu\sigma^2). \end{aligned}$$

These expressions of the log-cumulants of order p of $A(t)$ have been obtained by ourselves from Eq. (2.112). Note that for higher orders than the first three c_p stated here, log-cumulants are also non-zero, $p \geq 4 : \ c_p \neq 0$.

2.7.1-c) Infinitely Divisible Cascades

Since compound Poisson distributions belong to the more general class of infinitely divisible distributions, the product Eq. (2.99) can be generalized to the exponential of a continuous random measure. This defines the *infinitely divisible cascades* (IDC):

$$Q_r(t) = C \exp\left(\int_{\mathcal{C}_r(t)} dM(t', r')\right).$$

Here, C is a normalizing constant such that $\mathbb{E}Q_r(t) = 1$, and $\mathcal{C}_r(t)$ and $dm(t, r)$ are often defined as for CPC to ensure power law behaviors and stationary increments[10]. In a similar manner as for CMC and CPC, the corresponding motion $A(t)$ can be defined:

$$A(t) = \lim_{r \to 0} \int_0^t Q_r(s) ds. \tag{2.113}$$

Despite theoretical interests, these processes remain rarely used in applications and difficult and time consuming to synthesize. We will not consider IDC any further in this manuscript. For more details, the reader is referred to e.g. [22, 46].

2.7.1-d) Finiteness of moments

The range of statistical orders $[q_c^-, q_c^+]$ for which the moments of the wavelet Leaders of the multiplicative cascades are defined is given by [95]:

$$q_c^+ = \sup\{q \geq 1 : q + \varphi(q) \geq d\}, \tag{2.114}$$

$$q_c^- = \inf\{q \in \mathbb{R} : \mathbb{E}W^q \leq \infty\}. \tag{2.115}$$

2.7.1-e) Limitations

Any of the processes $Q_r(t)$ and $A(t)$ defined above are in practice often not rich enough to model empirical data: $Q_r(t)$ is by construction only positive valued, and consequently, $A(t)$ has in addition only positive valued variations. For many physical signals, however, model processes need to take on both positive and negative values, and possess both positive and negative variations.

[10] Again, other choices for $dm(t, r)$ are possible and lead to different properties.

2.7.2 Processes

2.7.2-a) Fractional Brownian Motion in Multifractal Time (CPM-MF-fBm)

Subordinating fractional Brownian motion $B_H(t)$ to compound Poisson motion $A(t)$ yields the so-called *fractional Brownian motion in multifractal time* (CPM-MF-fBm) [124, 145]:

$$V_H(t) = B_H(A(t)). \tag{2.116}$$

CPM-MF-fBm is usually a more realistic model process than CPC or CPM, since it takes values in $\mathbb{R}$ and has both positive and negative valued variations. With the choice Eq. (2.103) for the intensity measure of the cascade, V_H has stationary increments and exhibits scaling properties of the form of Eqs. (2.7) and (2.58) for $q_c^- < q < q_c^+$ [46, 47]:

$$\mathbb{E}L_{V_H}(a,t)^q \simeq c_q|a|^{\lambda(q)}, \tag{2.117}$$

$$\text{with } \lambda(q) = qH + \varphi(qH), \tag{2.118}$$

with $\varphi(q)$ given by Eq. (2.112). Its moments are finite for $q_c^- < q < q_c^+$, the q_c being defined by Eqs. (2.114) and (2.115) by replacing q with qH, i.e., $q_c^{V_H} = q_c/H$. The multifractal spectrum $\mathcal{D}(h)$ of $V_H(t)$ can be derived from the Legendre transform $D(h)$ of $\lambda(q)$ as: $\mathcal{D}(h) = D(h)$ if $D(h) \geq 0$, and $\mathcal{D}(h) = -\infty$ otherwise [30]. The scaling exponents are given by: $\zeta^L(q) = \lambda(q)$ for $q \in [q_*^-, q_*^+]$, and behave as a linear function of q beyond the range $[q_*^-, q_*^+]$ (cf. Section 2.8, Eqs. (2.129) and (2.130)). The log-cumulants $c_p^{V_H}$ of the compound motion process $V_H(t)$ are related to the log-cumulants c_p^A of the motion process $A(t)$ by:

$$c_p^{V_H} = H^p c_p^A. \tag{2.119}$$

The CPM-MF-fBm construction therefore provides us with a theoretically controlled multifractal process with a priori prescribed non zero c_3 and stationary increments [46, 47].

2.7.2-b) Multifractal Random Walk (mrw)

The cascade measures $Q_r(t)$ can also be involved in the theoretical definition of a true continuous time *multifractal random walk* according to [22, 23, 138]:

$$Z_H(t) = \int_{\mathbb{R}} Q_r(s) dB_H(s). \tag{2.120}$$

The definition of such a continuous time process is, however, intricate and requires further developments. Hence, in practice, it is defined from the following discrete-time ad-hoc construction [21], which we simply refer to as *multifractal random walk* (mrw):

$$X(k) = \sum_{k=1}^{n} G_H(k) \exp(\omega(k)). \tag{2.121}$$

In this definition, the process $G_H(k)$ consists of the increments of fBm with parameter $H > 1/2$: $G_H(k) = B_H(k+1) - B_H(k)$. The process ω is independent of G_H, Gaussian, with the following specific covariance:

$$\text{cov}(\omega(k_1), \omega(k_2)) = \begin{cases} \beta \ln\left(\frac{L}{|k_1 - k_2| + 1}\right) & \text{if } |k_1 - k_2| < L \\ 0 & \text{otherwise,} \end{cases} \tag{2.122}$$

with $\beta > 0$. The form of this covariance, which is in no way intuitive, has been chosen to mimic the one observed for CMC. By construction, mrw has stationary increments and is non Gaussian. It exhibits scaling properties of the form of Eqs. (2.7) and (2.58) for the range $q_c^- < q < q_c^+$ (cf. Section 2.7.1-d), Eqs. (2.114) and (2.115)) [46, 47]:

$$\mathbb{E}L_X(a,t)^q \simeq c_q|a|^{\lambda(q)}, \tag{2.123}$$

$$\text{with } \lambda(q) = q + \varphi(qH), \tag{2.124}$$

where

$$\varphi(q) = \beta q - \beta^2 q^2/2. \tag{2.125}$$

The scaling exponents are given by: $\zeta^L(q) = \lambda(q)$ for $q \in [q_*^-, q_*^+]$,and behave as a linear function of q beyond the range $[q_*^-, q_*^+]$ (cf. Section 2.8, Eqs. (2.129) and (2.130)). Hence, $c_1 = H + \beta$, $c_2 = -\beta^2$ and $c_p \equiv 0$ for all $p \geq 3$. Its multifractal spectrum is given by the Legendre transform $D(h)$ of $\lambda(q)$ as: $\mathcal{D}(h) = D(h)$ if $D(h) \geq 0$, and $-\infty$ otherwise. Hence, as for CMC-LN, mrw is often seen as a simple multifractal process: The departure of its $\zeta^L(q)$ from a linear behavior in q simplifies to a 2nd order polynomial form and is fully controlled by the single parameter c_2, and its multifractal spectrum is given by a parabola.

2.7.2-c) Random Wavelet Cascades (RWC)

(Dyadic) Random Wavelet Cascades (RWC) [15] have been introduced to mimic the scale invariance properties of CMC. These specific cascade constructions define directly a process. They are based on a rather particular construction, since the scaling properties (as practically observed through for instance a wavelet coefficient based analysis, Eqs. (2.7) or (2.62)) are in some sense directly prescribed to the wavelet coefficients. It is the coefficients themselves that are the product of (i.i.d.) multipliers and define the cascade process. Starting from a first scale $j = 0$ at which there is only one coefficient $d_{j=0,k=1}$, the dyadic tree of wavelet coefficients is constructed by the following iteration: Draw, for each coefficient $d_{j-1,k}$ at scale $j-1$, two independent identically distributed random multipliers W_1 and W_2, giving rise to 2 new coefficients $d_{j,2k} = W_1 \cdot d_{j-1,k}$ and $d_{j,2k+1} = W_1 \cdot d_{j-1,k}$, and assign a random sign to them. Hence, the wavelet coefficient at scale j is given by:

$$d_{j,k} = \varepsilon_{j,k} \prod_{j'=0}^{j} W_{j'} \tag{2.126}$$

where $\varepsilon_{j,k}$ is a random sign and the $W_{j'}$ are the i.i.d. random multipliers corresponding to the position (j,k) in the cascade. The process is then obtained by the inverse wavelet transform. The convergence condition and the scaling and multifractal properties of the cascade are prescribed by the function:

$$\varphi(q) = -\log_2 \mathbb{E}W^q. \tag{2.127}$$

In this manuscript, we will use RWC exclusively for an analytic study of the dependence structure of wavelet coefficients in Section 11.2.1. For more details on the construction, for conditions and proofs of convergence and for further properties, see e.g. [15].

2.7.3 Practical implementation

The practical implementation of the synthesis procedures in MATLAB® is due to the members or alumni of the research group Sisyphe at the Laboratoire de Physique de l'Ecole Normale Supérieure de Lyon, in particular P. Chainais, P. Abry and S.G. Roux.

2.8 H-sssi Processes versus Multifractal Multiplicative Cascade Processes

The (monofractal) fv H-sssi processes and the multifractal multiplicative cascade processes amount to the two major scale invariant process classes and constitute the processes that are quasi-exclusively used in applications. Both classes satisfy the key relation[11] [92]:

$$\mathbb{E}L_X(a,t)^q \sim c_q|a|^{\lambda(q)} \tag{2.128}$$

In this section, we collect the key differences between these classes with respect to Eq. (2.128).

Scaling exponents, log-cumulants and multifractal spectrum. The first key difference lies in the fact that for fv H-sssi processes, $\lambda(q)$ is a *linear* function in q, whereas for MMC processes it is *nonlinear* in q:

$$\begin{aligned} \lambda(q) &= qH \quad \text{(fv } H\text{-sssi processes)} \\ \lambda(q) &\neq qH \quad \text{(MMC processes).} \end{aligned}$$

Therefore, whereas the scale invariance properties of H-sssi processes are entirely controlled by one *single* parameter H, this involves a *collection* of parameters for MMC processes. This translates to log-cumulants as:

$$\begin{aligned} c_1 = H,\ \forall p \geq 2: \quad c_p &\equiv 0 \quad \text{(fv } H\text{-sssi processes)} \\ \exists p \geq 2: \quad c_p &\neq 0 \quad \text{(MMC processes),} \end{aligned}$$

i.e., for fv H-sssi processes, all log-cumulants except c_1 are equal to the zero value, whereas for MMC processes, there exists at least one non-zero log-cumulant c_p of order $p \geq 2$. This fact greatly facilitates the practical design of test procedures aiming at discriminating these two process classes (cf. Chapter 9).
The difference is also reflected by the fact that the multifractal spectrum of fv H-sssi processes reduces to *one single point* ($\mathcal{D}(h) = d$ if $h = H$, and $-\infty$ otherwise), whereas for MMC processes, the multifractal spectrum is non trivial.

Integral scale. The second key difference is that for H-sssi processes, Eq. (2.128) is fulfilled for *all* (fine and coarse) scales,

$$0 < a < \infty,$$

whereas for MMC processes, this is only true for *fine* scales $a \to 0$, since the cascade construction has to start at a certain scale. In practice, this means that Eq. (2.128) is only verified for

$$0 < a < \tilde{a},$$

where $\tilde{a}$ is called the *integral scale* [75].
In this manuscript, we choose not to study the role of the integral scale: *All of the theoretical studies and numerical simulation studies in this manuscript suppose that the signal or process has only one integral scale.*

[11] We restrict ourselves here to processes with stationary increments.

Finiteness of moments. The third key difference resides in the range of q for which Eq. (2.128) is valid. For fv H-sssi processes, this is the case for[12] $q \in \mathbb{R}$ [92]. For MMC processes, Eq. (2.128) is only valid for the range $q \in [q_c^-, q_c^+]$ defined by Eqs. (2.114) and (2.115), with $-\infty < q_c^-$ and $q_c^+ < +\infty$.

Linearization effect. Finally, for MMC processes, the sample moments behave as[13]:

$$\frac{1}{n_a}\sum_t L_X(a,t)^q \sim c_q|a|^{\zeta^L(q)},$$

where $\zeta^L(q) \equiv \lambda(q)$ only within the further restricted range of statistical orders $q \in [q_*^-, q_*^+]$:

$$-\infty < q_c^- < q_*^- < q < q_*^+ < q_c^+ < +\infty,$$

The values of q_*^- and q_*^+ are conjectured in [108] to be given by the following expressions:

$$q_*^+ = \sup\{q \geq 0 :\ d + q\frac{d}{dq}\lambda(q) - \lambda(q) = 0\}, \tag{2.129}$$

$$q_*^- = \inf\{q \leq 0 :\ d + q\frac{d}{dq}\lambda(q) - \lambda(q) = 0\}, \tag{2.130}$$

Beyond the range $q \in [q_*^-, q_*^+]$, the function $\zeta^L(q)$ is necessarily *linear* in q. Therefore, sample moment estimators behave as if moments would only be finite for $q \in [q_*^-, q_*^+]$ [11, 108]. This is called the *linearization effect* and will be further discussed in Section 5.2.

We will throughout this manuscript – with the exception of Section 5.2 – only consider the range of statistical orders $q \in [q_*^-, q_*^+]$, and processes for which: $|q_*^-|, |q_*^+| > 2$.

[12] Note that when wavelet Leaders are replaced with wavelet coefficients in Eq. (2.128), the range of q is reduced to $-d < q < \infty$.

[13] For H-sssi processes, $\zeta^L(q) \equiv qH$ for $q \in \mathbb{R}$.

Chapter 3

Statistical Characterization of Wavelet Leaders

Contents

Wavelet Leaders are a recent development and have been shown to possess significant theoretical and practical advantages over other multiresolution quantities, such as wavelet coefficients, for performing multifractal analysis (see [89, 92] and Section 2.5.5). Theory for wavelet Leaders exclusively concentrates on their function space and local regularity characterization properties. Therefore, despite them being backed up by a solid mathematical framework from a functional analysis perspective, nothing is known on their *statistical properties*, although these are of key importance in applications. Even more elementary, the consequences of passing from the theoretical infinite resolution definition Eq. (2.54) of wavelet Leaders to practical implementations, necessarily based on *finite resolution* data and limited by a finest available scale, have never been clearly investigated.
Therefore, the first aim of this chapter is to investigate at which scale the $\sup$ over all finer scales in the definition of wavelet Leaders is practically attained — and hence wavelet Leaders are obtained — for finite resolution data, and to characterize the implications for multifractal attribute estimation. This is accounted for in Section 3.1.
Second, we aim at providing analytical results for the statistical properties of wavelet Leaders. To this end, we first state the general form of the marginal and bivariate distribution of wavelet Leaders. Motivated by the key role of wavelet coefficients for fBm, we further characterize these distributions for this specific process and propose marginal and bivariate distribution models. These studies make up the core of Section 3.2 and constitute the only analytical statistical characterization of wavelet Leaders available in the literature at present.
For convenience, we choose to limit the study in this chapter to the 1d case only. Similar results can be obtained for the 2d case.

3.1 Fine Scale Propagation and Finite Size Effects

The theoretical definition of the wavelet Leader $L_X(j,k)$ Eq. (2.54), involves taking the $\sup$ of wavelet coefficients at *all finer scales* $j' \leq j$, $j' \to -\infty$. In practical implementations, however, this $\sup$ can only be taken over finer scales until the *finest available scale*. With the notations used in this manuscript, the finest scale is given by $j = 1$. Therefore, the first issue to address in practice is: Where are wavelet Leaders obtained in practical implementations? Or, to put it more precisely: At which scale j is the number of available fine scales j', $1 \leq j' < j$, large enough such that the $\sup$ is practically attained, and the wavelet Leader $L_X(j,k)$ actually has the properties of a wavelet Leader? We choose to call this discrepancy between theoretical definition and practical implementation the *fine scale cutoff*. To characterize its implications, we propose in Section 3.1.1 to study the fine scale propagation probability, defined as the probability that a wavelet Leader $L_X(j,k)$ takes its value at a scale $j' \leq j$. Then, we investigate the impact on actual estimation quality by comparison of wavelet Leader and wavelet coefficient based structure functions in Section 3.1.2.

3.1.1 Fine scale propagation probability

The wavelet Leader $L(j,k)$ equals by definition the magnitude of one (the largest) of the wavelet coefficients $d(j',k')$ in the neighborhood $3\lambda_{j,k}$. We define the *fine scale propagation probability* as the probability $\pi_j(j')$ that a wavelet Leader $L(j,k)$ takes on its value at

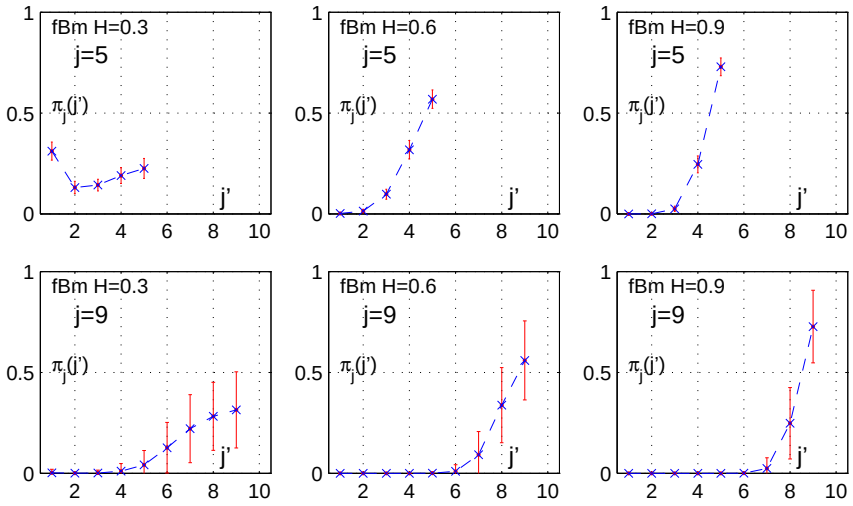


Figure 3.1: **Wavelet Leaders: Propagation of wavelet coefficients for fBm.** Fraction $\pi_j(j')$ of wavelet coefficients at scale j' that become wavelet Leaders at scale j, $\pi_j(j') = \mathsf{Pr}[L_X(j,\cdot) = |d_X(j',\cdot)|]$ for fBm with $H = 0.3, 0.6, 0.9$ (left, center, right column, respectively) at scales $j = 5$ (top row) and $j = 9$ (bottom row). Shown are the mean ('$\times$') and 1.96 standard deviation error bars (red solid vertical lines) of $\pi_j(j')$, obtained over 100 realizations of fBm of sample size $N = 2^{15}$, respectively.

scale j':

$$\pi_j(j') = \mathsf{Pr}[L(j,k) = |d(j',\cdot)|], \tag{3.1}$$

that is, as the probability that $L(j,k)$ equals the wavelet coefficient $|d(j',k')| \in 3\lambda_{j,k}$ at scale j'. In other words, $\pi_j(j')$ expresses the probability that a wavelet coefficient at scale j' becomes a wavelet Leader at scale j — or, how far are wavelet coefficients from fine scales likely to propagate towards coarser scales.
Empirically, the fine scale propagation probability Eq. (3.1) is evaluated by calculating the wavelet coefficients $d_X(j,k)$ and the wavelet Leaders $L_X(j,k)$, determining, for each $L_X(j,k)$, the scale $j'(j)$ of the corresponding wavelet coefficient $d_X(j',k') \rightarrow L_X(j,k)$, and calculating $\hat{\pi}_j(j') = \#j'(j)/n_j$, where n_j is the number of wavelet Leaders at scale j.

3.1.1-a) H-sssi processes

Fig. 3.1 shows fine scale propagation probabilities, obtained as means over 100 realizations of fBm of sample size $N = 2^{15}$ using Daubechies wavelets with $N_\psi = 3$, for scales $j = 5$ (top row) and $j = 9$ (bottom row). The left column corresponds to $H = 0.3$, the center column to $H = 0.6$, and the right column to $H = 0.9$.
First, we observe that for H fixed, $\pi_{j=5}(j') \approx \pi_{j=5+4}(j'+4)$ shifted to coarser scales[1], with the exception of $H = 0.3$, for which there is a significant increase of $\pi_{j=5}(j'=1)$ due to fine scale cutoff.

[1]The main difference is in the size of error bars: There are significantly less wavelet Leaders at scale $j = 9$ than at $j = 5$, hence decreased estimation quality at the larger scale.

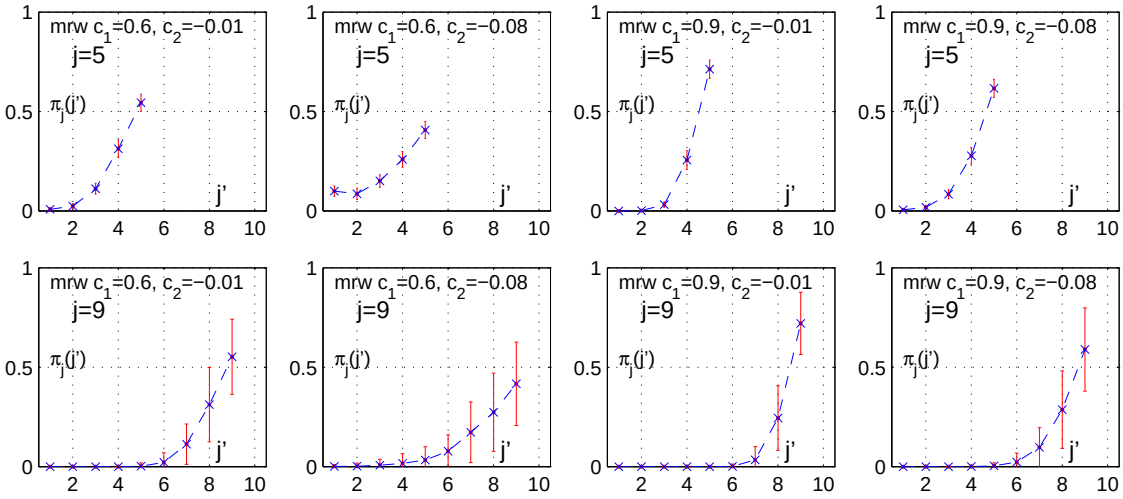


Figure 3.2: **Wavelet Leaders: Propagation of wavelet coefficients for mrw.** Fraction $\pi_j(j')$ of wavelet coefficients at scale j' that become wavelet Leaders at scale j, $\pi_j(j') = \Pr[L_X(j,\cdot) = |d_X(j',\cdot)|]$ for mrw with $(c_1,c_2) = (0.6,-0.01),(0.6,-0.08),(0.8,-0.01),(0.8,-0.08)$ (1st, 2nd, 3rd, 4th column, respectively) at scales $j=5$ (top row) and $j=9$ (bottom row). Shown are the mean ('$\times$') and 1.96 standard deviation error bars (red solid vertical lines) of $\pi_j(j')$, obtained over 100 realizations of mrw of sample size $N=2^{15}$, respectively.

Second, the probability for wavelet coefficients to propagate far into coarse scales decreases with increasing H, hence $\pi_j(j'=j) \gg \pi_j(j'<j)$ as $H \to 1$. In contrast, $\pi_j(j')$ spreads over a larger support of j' for H small: Indeed, whereas for $H=0.9$ $\pi_j(j') \approx 0$ for $j-j' \geq 3$, for $H=0.3$ $\pi_j(j') \approx 0$ only for $j-j' \geq 6$. This observation can be interpreted as follows: For fBm, wavelet coefficients are distributed as $d(j,\cdot) \sim \mathcal{N}(0,\sigma(j))$ with $\sigma(j) \sim 2^{jH}$ (cf. Section 2.3.4). Since $\sigma(j)$ increases faster with j for H large than for H small, the probability that the magnitude of a wavelet coefficient at scale j is larger than the magnitude of a wavelet coefficient at scale $j'<j$ is larger for large H. Therefore, it is more likely for a wavelet coefficient to propagate far into coarse scales when H is small.

3.1.1-b) Multifractal multiplicative cascade processes

Fig. 3.2 shows fine scale propagation probabilities, obtained as means over 100 realizations of mrw of sample size $N=2^{15}$ using Daubechies wavelets with $N_\psi=3$, for scales $j=5$ (top row) and $j=9$ (bottom row). The first column corresponds to mrw with $(c_1,c_2)=(0.6,-0.01)$ (weakly multifractal, small H), the second one to $(0.6,-0.08)$ (strongly multifractal, small H), the third one to $(0.8,-0.01)$ (weakly multifractal, large H) and the rightmost one to $(0.8,-0.08)$ (strongly multifractal, large H).
First, we observe that, as for fBm, $\pi_{j=5}(j') \approx \pi_{j=5+4}(j'+4)$ shifted to coarser scales, hence $\pi_j(j')$ does, for a fixed set of process parameters, not depend on the absolute value of j and depends only on the difference $j-j'$. Similar results can be obtained at other scales j.
Second, the probability for wavelet coefficients to propagate far into coarse scales decreases with increasing c_1, hence $\pi_j(j'=j) \gg \pi_j(j'<j)$ when $c_1=0.8$, whereas $\pi_j(j')$

spreads over a larger support of j' for c_1 small. This is consistent with the previous observations for fBm, since for multifractal processes, c_1 plays a role equivalent to that of H for H-sssi processes.
Third, the probability of propagating far into coarse scales increases when c_2 becomes more negative. This can be interpreted by the fact that a more negative c_2 implies more heavy-tailed marginals of wavelet coefficients, hence an increased probability of observing a wavelet coefficient with large magnitude, capable of propagating far into coarse scale.
Finally, let us note that the influence of c_1 on $\pi_j(j')$ is larger than that of c_2 (a c_2 of -0.01 representing very weak, one of -0.08 very strong departure from monofractal).

3.1.1-c) Conclusions

The study of fine scale propagation probabilities leads us to the following conclusions. Fine scale cutoff affects wavelet Leaders at fine scales. In practice, only wavelet Leaders above a certain scale $j > 3 - 6$ are free of the influence of limited fine scale resolution. The number of fine scales polluted by the cutoff depends on H for H-sssi, on c_1 for multiplicative cascade multifractal processes and, less significantly, also on c_2 for the latter. In particular, the smaller H or c_1, the larger the number of fine scales that are affected. Therefore, the number of scales concerned by this effect can in practice be circumvented and controlled by (fractionally) integrating the process by an order α, hence artificially increasing H or c_1 in a controlled way by α[2].

3.1.2 Practical implications of finite size effects

The previous section showed that finite size effects lead to pollution of wavelet Leaders for a certain number of fine scales: Fig. 3.1 and 3.2 suggest that in the best case, only the 3 finest scales are polluted, i.e. $\pi_j(j') \approx 0$ for $j - j' \geq 3$, but for other process parameter settings, this can concern up to 6 or 7 fine scales. However, the probabilities $\pi_j(j')$ only capture the difference between the scale at which a wavelet Leader lives, and the scale at which its $\sup$ is reached, and they do not measure the difference in *magnitude* of the $\sup$ due to fine scale cutoff. Therefore, $\pi_j(j')$ does not directly characterize the alteration of the distribution of wavelet Leaders or structure functions and, consequently, estimation quality. For this reason, Fig. 3.3 compares (the log of) structure functions $\log_2 S(j, q = 1)$, as involved in parameter estimation (top row), and structure functions minus their theoretical slope $\log_2 S(j, q = 1) - j\zeta(q = 1)$ (bottom row), based on wavelet coefficients (blue solid lines and '$\circ$') and Leaders (red solid lines and '$\times$'), to regression lines over large scales. Shown are results for fBm (left column) and mrw (right column) for typical process parameter settings.
We observe that, for both processes, wavelet coefficient based structure functions display linear behavior only from the second or third fine scale on. This phenomenon has nothing to do with fine scale cutoff – which does not exist for wavelet coefficients – but is a consequence of the fact that the initialization step in the calculation of the wavelet coefficients is omitted (cf. [3, 172]).
For both processes, wavelet Leader based structure functions are perfectly straight lines only from one scale higher on than wavelet coefficients based ones. This indicates that

[2]This can be practically performed using the procedures introduced in Section 5.1.3.

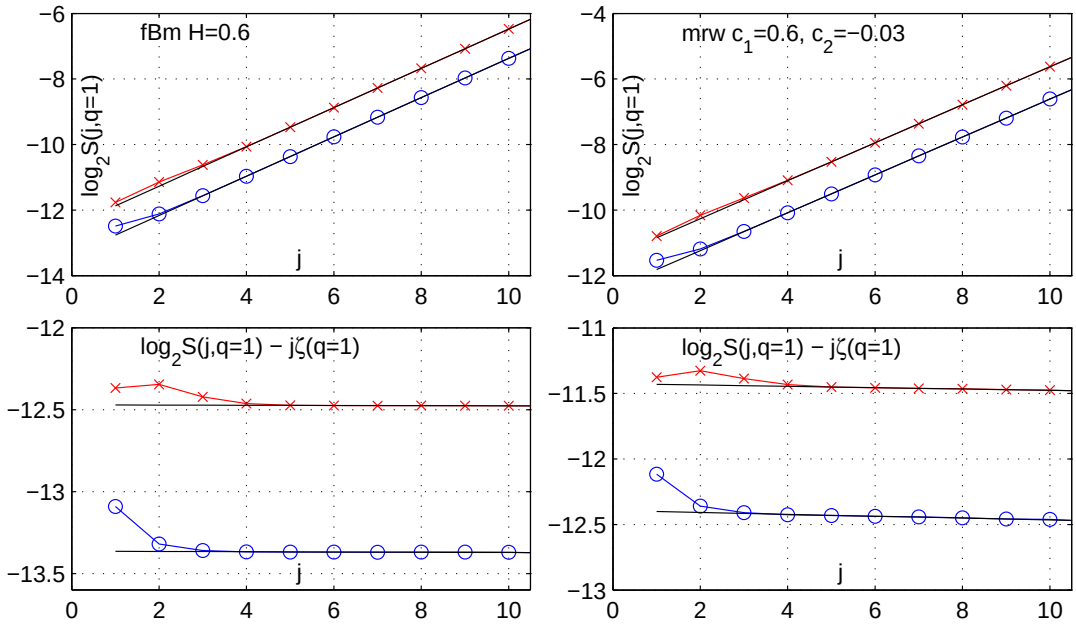


Figure 3.3: **Structure functions for fBm and mrw.** Mean over 1000 realizations of structure functions $\log_2 S(j, q = 1)$ (top) rectified by theoretical slope $\log_2 S(j, q = 1) - j\zeta(q = 1)$ (bottom) for wavelet coefficients (blue solid lines and 'o') and Leaders (red solid lines and '×') for fBm ($H = 0.6$, left) and mrw ($(c_1, c_2) = (0.6, -0.03)$, right) and sample size $N = 2^{18}$ (coarse scales are not shown). The solid black lines correspond to ordinary linear regressions over the scales $j \in [5, 10]$.

fine scale cutoff practically pollutes only one or two fine scales of wavelet Leaders more than are already polluted by omitting the initialization step of the wavelet transform. Similar results – not reported here for space reasons – are obtained for other processes and process parameter settings.
Complementary results and discussions for the 2d case are reported in Section 4.2.2.

3.1.3 Conclusions

We conclude that wavelet Leaders are in practice penalized by finite size effects, since roughly one or two fine scales more than for wavelet coefficients depart from linear behavior and should not be considered for estimation of multifractal attributes. Hence, the linear regression range for wavelet Leader based estimation should not start below the scale $j_1^L = j_1^d + 1$.
Although one additionally polluted scale for wavelet Leaders as compared to wavelet coefficients may seem a small price to pay, this can in practice constitute a severe limitation, in particular for small sample sizes, when only a very limited number of scales are available. Also, at scale j_1 there are roughly as many wavelet coefficients (Leaders) and hence statistical information as at all coarser scales $j > j_1$ together. The impact on estimation performance will be further quantified in Section 4.1.5.
Finally, the alteration of structure functions at fine scales introduces additional practical difficulties for the validation of the range of scales over which scale invariance actually exist.

3.2 Some Statistical Properties for Wavelet Leaders

3.2.1 Marginal and bivariate characterization

The characterization of wavelet Leaders considered in the literature is exclusively limited to their function space and local regularity analysis properties, and no theoretical result on their statistical properties is available in the literature at present. Therefore, in this section, we state the general form of the marginal and bivariate cumulative distribution functions of wavelet Leaders. This provides us with heuristic arguments concerning the dependence of wavelet Leaders with respect to the dependence of wavelet coefficients, suggesting that they are of the some order.

3.2.1-a) Marginal distributions

From the definition of wavelet Leaders Eq. (2.54), $L_X(j,k) = \sup_{\lambda' \in 3\lambda_{j,k}} |d_{X,\lambda'}|$, the general expression for the marginal cumulative distribution $F_{L_{j,k}}(\gamma) = \mathsf{Pr}\left[L_X(j,k) < \gamma\right]$ of the wavelet Leader $L_{j,k} = L_X(j,k)$ can immediately be written as:

$$F_{L_{j,k}}(\gamma) = F_{|d_{3\lambda_{j,k}}|}(\gamma, \gamma, \cdots, \gamma). \tag{3.2}$$

Here $F_{|d_{3\lambda_{j,k}}|}$ denotes the joint cumulative distribution function of the absolute value of all the wavelet coefficients $|d(j,k)|$ situated within the neighborhood $3\lambda_{j,k}$. Hence, the marginal distribution of the wavelet Leader $L(j,k)$ depends on the joint distribution of all wavelet coefficients within the neighborhood $3\lambda_{j,k}$. Note that in theory, it therefore involves the joint distribution of an infinity of wavelet coefficients, since the sup is taken over all finer scales $j' \leq j$ and hence till the limit $j' \to -\infty$. In practice, only a finite number of wavelet coefficients at scale $j \geq 1$ are involved, and the limit can not be reached. The influence of this fine scale cut off is discussed above in Section 3.1.

3.2.1-b) Bivariate distributions

In the same spirit, we can write the general form of the bivariate cumulative distribution

$$F_{L_{j,k},L_{j',k'}}(\alpha,\beta) = \mathsf{Pr}\left[L(j,k) < \alpha, L(j',k') < \beta\right]$$

of the two wavelet Leaders $L_{j,k} = L_{j,k}$ and $L_{j',k'} = L_{j',k'}$. Let us define the sets $\mathcal{A} = 3\lambda_{j,k} \cap 3\lambda_{j',k'}$, $\mathcal{B} = 3\lambda_{j,k} \backslash 3\lambda_{j',k'}$ and $\mathcal{C} = 3\lambda_{j',k'} \backslash 3\lambda_{j,k}$ as illustrated in Fig. 3.4. Then, the bivariate cumulative distribution is given by:

$$F_{L_{j,k},L_{j',k'}}(\alpha,\beta) = F_{|d_{\mathcal{A}}|;|d_{\mathcal{B}}|;|d_{\mathcal{C}}|}\left(\min(\alpha,\beta);\alpha;\beta\right). \tag{3.3}$$

Here $F_{|d_{\mathcal{A}}|;|d_{\mathcal{B}}|;|d_{\mathcal{C}}|}$ stands for the joint cumulative distribution function of the absolute value of the wavelet coefficients $|d(j,k)|$ within the union of the neighborhoods $3\lambda_{j,k} \cup 3\lambda_{j',k'}$, and $|d_{\mathcal{X}}|$ stands for the collection of wavelet coefficients $|d(j,k)|$ for which $(j,k) \in \mathcal{X}$.
We can hence distinguish three different cases for $F_{L_{j,k},L_{j',k'}}(\alpha,\beta)$: First, the neighborhoods do not overlap and $\mathcal{A}$ is empty. Second, one of the neighborhoods is a subset of the other neighborhood, in which case either $\mathcal{B}$ or $\mathcal{C}$ is empty, depending on whether $j < j'$ or $j' < j$. Third, the neighborhoods do overlap but none of them is a subset of the other (this is the situation illustrated in Fig. 3.4). In the last situation, there are only two possible configurations: The neighborhoods share either one or two wavelet coefficients at j' (if $j \geq j'$) or j (if $j' \geq j$).

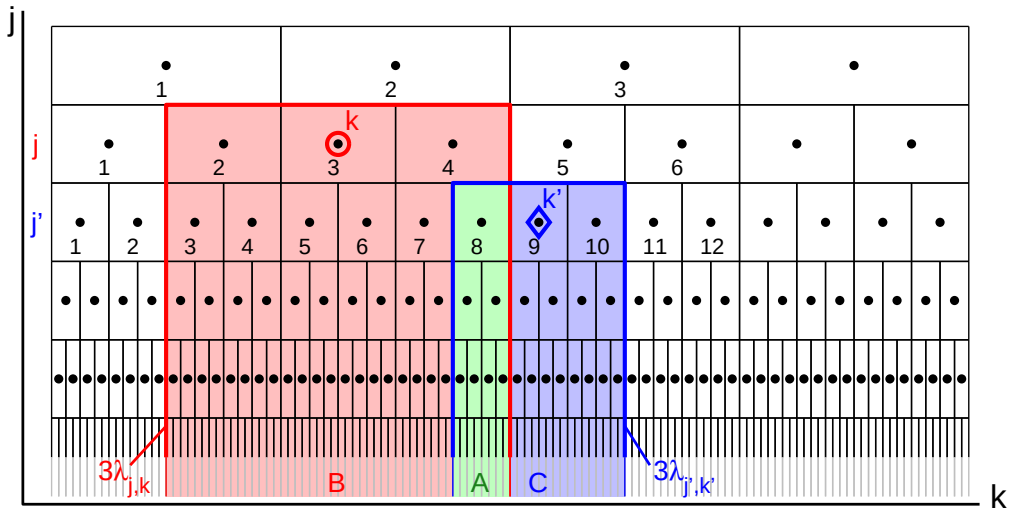


Figure 3.4: **Wavelet Leaders and neighborhoods $3\lambda_{j,k}$, $3\lambda_{j',k'}$.** Dyadic grid of wavelet Leaders ('•'), wavelet Leaders $L_X(j,k)$ ('○') and $L_X(j',k')$ ('◇'), neighborhoods $3\lambda_{j,k}$ (red frame) and $3\lambda_{j',k'}$ (blue frame), and $\mathcal{A} = 3\lambda_{j,k} \cap 3\lambda_{j',k'}$ (green region), $\mathcal{B} = 3\lambda_{j,k} \backslash 3\lambda_{j',k'}$ (red region) and $\mathcal{C} = 3\lambda_{j',k'} \backslash 3\lambda_{j,k}$ (blue region).

3.2.1-c) Dependence: Intuitions

Fig. 3.4 and Eq. (3.3) suggest the following heuristic: The dependence of the wavelet Leaders $L(j,k)$ and $L(j',k')$ is not larger than the largest dependence between a wavelet coefficient in $3\lambda_{j,k}$ and a wavelet coefficient in $3\lambda_{j',k'}$. Clearly, this intuition is not fruitful in the case when the two neighborhoods overlap, since then there are wavelet coefficients in $3\lambda_{j,k}$ and in $3\lambda_{j',k'}$ that are identical. It suggests, however, that the dependence of two wavelet Leaders whose neighborhoods are far apart in time decays roughly at the same speed as the dependence of wavelet coefficients at comparable distance.
Note that no mathematically rigorous results, such as an upper bound for the covariance of wavelet Leaders, has been obtained for this intuition. This demands for further theoretical work. The difficulty resides in the $\sup$ in the definition of wavelet Leaders, making standard bounds such as Cauchy-Schwartz ineffective.

3.2.1-d) Numerical evidence for correlation

Numerical evidence supports the intuition that dependence of wavelet Leaders is qualitatively similar to dependence of wavelet coefficients for large time lags. Fig. 3.5 shows, as a function of time lags, means of correlation coefficients for the absolute value of wavelet coefficients and for wavelet Leaders, $\rho_{|d|}(k-k') = \frac{\mathbb{E}|d_X(j,k)||d_X(j,k')| - \mathbb{E}|d_X(j,k)|\mathbb{E}|d_X(j,k')|}{\mathrm{Std}_{|d_X(j,k)|}\mathrm{Std}_{|d_X(j,k')|}}$ and $\rho_L(k-k') = \frac{\mathbb{E}L_X(j,k)L_X(j,k') - \mathbb{E}L_X(j,k)\mathbb{E}L_X(j,k')}{\mathrm{Std}_{L_X(j,k)}\mathrm{Std}_{L_X(j,k')}}$, at a common scale $j = 3$. Results are obtained as means over 100 realizations of fBm ($H = 0.7$, left) and mrw ($(c_1, c_2) = (0.7, -0.05)$, center and right) of sample size $N = 2^{14}$ (Daubechies' wavelet, $N_\psi = 3$).
We observe that for fBm, correlation of wavelet coefficients pertains only in a very narrow time neighborhood and then decays extremely fast. Correlation of wavelet Leaders is

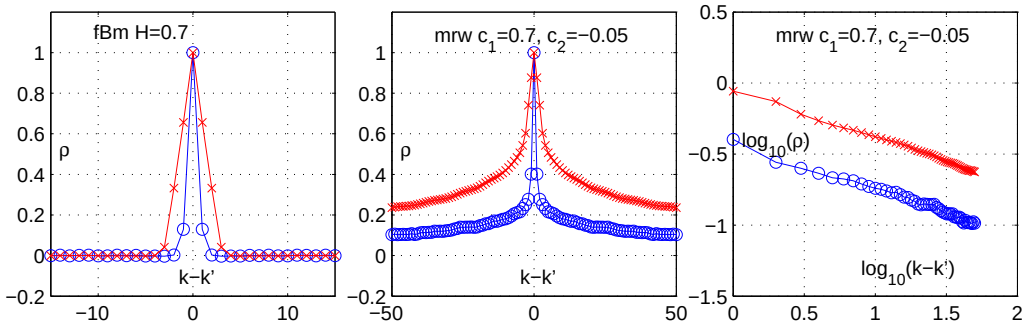


Figure 3.5: **Correlation coefficients of wavelet coefficients and Leaders.** Correlation coefficients $\rho(k-k')$ of absolute value of wavelet coefficients $|d_X(j,\cdot)|$ ('○') and Leaders $L_X(j,\cdot)$ ('×') as a function of lag $k-k'$ for fBm ($H=0.7$, left) and mrw ($(c_1,c_2)=(0.7,-0.05)$, center) at scale $j=1$, obtained as means over 100 realizations of sample size $N=2^{14}$. Log-log plot of correlation coefficients as in center plot (right).

significant in a time neighborhood only slightly larger than for wavelet coefficients, and goes to zero very fast with increasing time lag. For mrw, the magnitude of the correlation of wavelet Leaders is slightly larger than that of wavelet coefficients (cf. Fig. 3.5, center). However, the correlation of wavelet coefficients and Leaders are characterized by the same power law type behavior and decay, for large time lags, roughly at the same rate (cf. Fig. 3.5, right). Both observations support the intuition that for large time lags, wavelet coefficients and Leaders have qualitatively equivalent dependence.
The dependence structure of wavelet coefficients and Leaders for H-sssi and MMC processes will be subject to more detailed investigations in Chapter 11.

3.2.2 Marginal and bivariate characterization for fBm

The results in Eqs. (3.2) and (3.3) are interesting in themselves since they state the general form of the cumulative distributions of wavelet Leaders. In this subsection, we concentrate on such results for the specific case of fBm, with the following motivations: First, fBm represents the statistically most comfortable (since Gaussian) scale invariant model process. Hence, it is the most promising candidate process for the derivation of practically exploitable expressions for Eqs. (3.2) and (3.3). Second, fBm is the only scaling process for which exhaustive asymptotic theory on the statistics of wavelet coefficients is available [70] (cf. Section 2.3.4: stationarity (**P1**), weak correlation (**P3**), Gaussian (**P4**) and precise reproduction of scale invariance). These properties highly facilitate the statistical analysis of estimation procedures and make up for a major advantage of wavelet coefficients for the analysis of fBm, since they enable the derivation of statistical inference procedures [4, 26, 52, 84, 169]. Hence, the interest of investigating this specific case for wavelet Leaders.

From the properties **P1** and **P4**, we can directly state the marginal cumulative distri-

bution functions of the absolute value of the wavelet coefficients $|d(j,\cdot)|$,

$$\begin{aligned} F_{|d_j|}(\gamma) &= \mathsf{Pr}[|d_X(j,\cdot)| \leq \gamma] = 1 - \mathsf{Pr}[|d_X(j,\cdot)| > \gamma] = \\ &= 1 - 2 \cdot \mathsf{Pr}[d_X(j,\cdot) > \gamma] = 2 \cdot \mathsf{Pr}[d_X(j,\cdot) < \gamma] - 1 = \\ &= erf\left(\frac{\gamma}{\sqrt{2}\sigma(j)}\right), \end{aligned}$$

by making use of stationarity, symmetry around 0 of the marginals of $d_X(j,\cdot)$, and the fact that the marginals of $d_X(j,\cdot)$ are Gaussian. Hence, the absolute values of the wavelet coefficients have half-Normal distributions, and $\mathbb{E}|d_j| = \sqrt{\frac{2}{\pi}}\sigma(j)$ and $\mathsf{Var}|d_j| = \frac{\pi-2}{\pi}\sigma(j)$. As stated above (cf. Eq. (3.2)), the marginal distributions of the $L_X(j,k)$ involve the joint distribution of all wavelet coefficients in $3\lambda_{j,k}$. Whereas the joint distributions of the wavelet coefficients can be written in closed form if their covariance is explicitly known, this is not the case for their absolute value, since there is no closed form expression for joint half-Normal distributions. Therefore, we add the following simplifying assumption:

A1 The wavelet coefficients of fBm are independent.

Clearly, **A1** is a rather strong assumption, since it implies, together with **P1** and **P4**, *independent Gaussian wavelet coefficients*. We will, however, see that the resulting distribution model perfectly fits with empirical distributions for fBm.

3.2.2-a) Marginal distribution model

From property **P4** and under assumption **A1**, the marginal distributions of the $L_X(j,k)$ Eq. (3.2) simplify to:

$$F_{L_j}(\gamma) = \prod_{m=j_1}^{j} \left(F_{|d_m|}(\gamma)\right)^{3\cdot 2^{(j-m)}} = \prod_{m=j_1}^{j} \left(erf\left(\frac{\gamma}{\sqrt{2}\sigma(m)}\right)\right)^{3\cdot 2^{(j-m)}}, \qquad (3.4)$$

since at each scale $j' \leq j$, there are $3\cdot 2^{(j-j')}$ wavelet coefficients $d_X(j',\cdot)$ in $3\lambda_{j,k}$. Therefore, even in the simple model case of independent and Gaussian wavelet coefficients, the marginal distributions of wavelet Leaders do not have a simple closed form expression[3]. The form of the model distribution Eq. (3.4) indicates non Gaussian, but not heavy tailed or sub-exponential distributions for wavelet Leaders.

Fig. 3.6 plots numerical evaluations of Eq. (3.4) against the empirical distributions of $L_X(j,\cdot)$ from single realizations of fBm (sample size $N = 2^{20}$, Daubechies' wavelet with $N_\psi = 3$) with $H = 0.3, 0.6, 0.9$ (top, center and bottom rows, respectively) at the scales $j = 1, 3, 5, 7$ (first, second, third and fourth column, respectively). The finest scale is naturally $j_1 = 1$, and estimates of the standard deviations $\sigma(j)$ are used for evaluating Eq. (3.4). First, we observe that the fits are perfect, regardless of the scale j, and of the process parameter H. The latter observation is important in so far as H controls the (long range) dependence of the process. Even for very significant long range dependence ($H = 0.9$), the model Eq. (3.4) based on independent wavelet coefficients fits the data perfectly, hence indicating that the wavelet coefficients display only residual correlations, and that these residual correlations do not significantly alter the marginal distributions of the wavelet Leaders. Second, Fig. 3.6 shows that both the empirical distributions and the distribution model are *non Gaussian*, at all scales j and for any H.

[3] In particular, $j_1 \to -\infty$ in theory, hence an infinite product in Eq. (3.4).

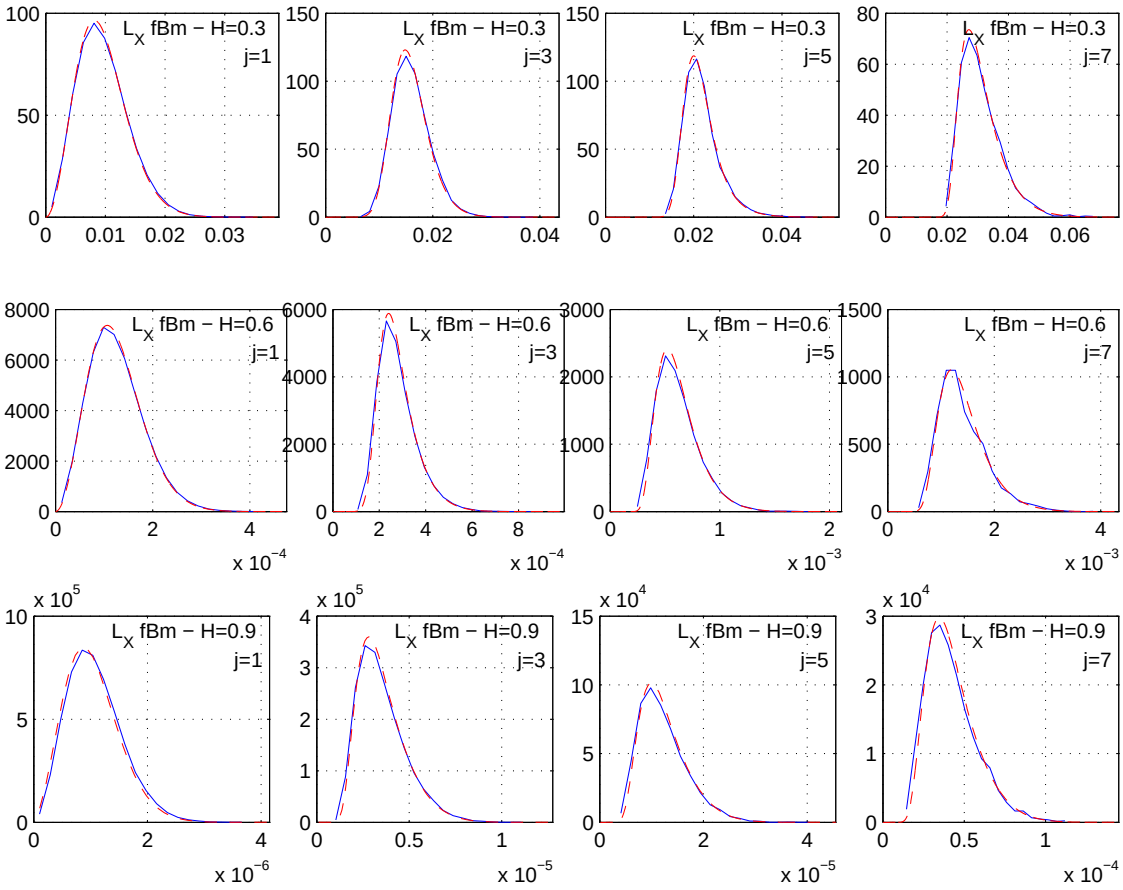


Figure 3.6: **Wavelet Leader marginal distribution model for fBm.** Wavelet Leader $L_X(j,\cdot)$ marginal empirical probability distributions (blue solid lines) and probability distribution model Eq. (3.4) (red dashed lines, numerically evaluated) for fBm $H = 0.3$ (top row), $H = 0.6$ (center row) and $H = 0.9$ (bottom row) at scales $j = 1, 3, 5, 7$ (1st, 2nd, 3rd and 4th column, respectively).

3.2.2-b) Bivariate distribution model

From tedious but straightforward geometrical considerations, mostly consisting of evaluations of the number of wavelet coefficients in the regions $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$ in Fig. 3.4, and under **P1**, **P4** and **A1**, the joint cdf Eq. (3.3) of $L(j,k)$ and $L(j',k')$ is derived as (assuming $j \geq j'$ without loss of generality):

$$F_{L_{j,k},L_{j',k'}}(\alpha,\beta) = \prod_{m=j_1}^{j'} \left(erf\left(\frac{\min(\alpha,\beta)}{\sqrt{2}\sigma(m)}\right)\right)^{\theta 2^{j'-i}} \cdot \left(erf\left(\frac{\alpha}{\sqrt{2}\sigma(m)}\right)\right)^{3\cdot 2^{j-m}-\theta 2^{j'-i}} \cdot \tag{3.5}$$

$$\left(erf\left(\frac{\beta}{\sqrt{2}\sigma(m)}\right)\right)^{(3-\theta)2^{j'-i}} \cdot \prod_{n=j'+1}^{j} \left(erf\left(\frac{\alpha}{\sqrt{2}\sigma(n)}\right)\right)^{3\cdot 2^{j-n}} \cdot \tag{3.6}$$

The first term in Eq. (3.5) involves wavelet coefficients in $\mathcal{A}$, the second and fourth term correspond to $\mathcal{B}$, the third term corresponds to $\mathcal{C}$ (cf. Fig. 3.4), and θ is given by:

$$\theta = \max\left(0, 2 + 3 \cdot 2^{j-j'-1} - \left|k' - \left(k - \frac{1}{2}\right) 2^{j-j'} - \frac{1}{2}\right|\right). \tag{3.7}$$

Hence, under assumption **A1**, the $L_{j,k}$ and $L_{j',k'}$ are independent if:

$$\left|k' - k \cdot 2^{j-j'} + 2^{j-j'-1} - \frac{1}{2}\right| \geq 3 \cdot 2^{j-j'-1} + 2$$

Then, $F_{L_{j,k},L_{j',k'}}(\alpha,\beta) = F_{L_j}(\alpha) \cdot F_{L_{j'}}(\beta)$. In other words, if wavelet coefficients (be they Gaussian or not) are *independent*, wavelet Leaders are m*-dependent*. This is a direct consequence of the fact that the neighborhoods $3\lambda_{j,k}$ and $3\lambda_{j',k'}$ of two wavelet Leaders that are far enough apart do not intersect[4].
Finally, we conclude that the form of both the Eqs. (3.4) and (3.5) do not suggest to be directly useful for further analytic statistical characterization of wavelet Leaders or estimation procedure. This needs to be further investigated.

3.2.3 Conclusions and perspectives

The results in this section lead us to the following conclusions. First, wavelet Leaders are in general *non Gaussian*. This suggests that for fBm, it may be advantageous to stay with (Gaussian) wavelet coefficients, which allow extensive statistical characterization of estimation procedures [4, 26, 52, 84, 169]. This does not take into account actual practical estimation performance, which will be the subject of the following Chapter 4. Eq. (3.2) indicates, nonetheless, that wavelet Leaders do in practice not introduce heavy-tailed or sub-exponential marginals if they are not already present for wavelet coefficients. Second, the covariance of wavelet Leaders has heuristically similar decay as the covariance of the absolute value of the wavelet coefficients from which they are constructed. Hence, qualitatively, wavelet Leaders do not seem to considerably increase practical difficulties resulting from dependence. A more precise statement demands for further theoretical investigations. Finally, the nonlinear nature of wavelet Leaders leads to (marginal and bivariate) distributions that appear not to be directly exploitable for further analytic characterization of wavelet Leaders or estimation procedures. Even in the simplest model case of independent and Gaussian wavelet coefficients, results as in Eqs. (3.4) and (3.5) do not seem to be directly useable for analytic statistical characterization of wavelet Leader based estimation.

[4] For instance, the neighborhoods of two wavelet Leaders $L_{j,k}$ and $L_{j,k'}$ at a common scale j do not intersect as soon as $|k - k'| \geq 3$.

Chapter 4

Estimation Performance

Contents

Although multifractal analysis has been extensively used in various applications involving 1d and 2d data of very different natures, the statistical properties and performance of estimation procedures remain poorly studied. Most of the literature concentrates on the estimation of the Hurst parameter for H-sssi processes based on (discrete) wavelet coefficients, for which important analytical results have been obtained. For fBm, for which the statistics of the wavelet coefficients can be widely characterized analytically (cf. Section 2.3.4), theoretical results for the estimator of H based on $\zeta(2)$ have been obtained. It can be shown that this estimator is quasi optimal, with performance close to the Cramer-Rao bound, and that estimates are Gaussian. Also, the estimates $\hat{\zeta}(q)$, $q > 0$ are (asymptotically) Gaussian, and their bias and standard deviation can be calculated analytically (e.g. [169]). However, no comparable results have been obtained for multifractal processes, due to the absence of theoretical statistical characterization of wavelet coefficients for such processes. What is more, no theoretical results have been obtained for wavelet Leader based estimation, be it for Gaussian H-sssi or for multifractal processes, since no analytical characterizations of wavelet Leaders other than the ones presented in the previous Chapter 3 are available at present.

Given this absence of analytical results for estimation procedures and the vast variety of applications, it is surprising that only very few numerical studies of the statistical performance of multifractal parameter estimation are available. For *1d signals*, some important practical characterizations have been obtained recently for wavelet coefficient and Leader based estimation, mainly in the works by Lashermes et al. [8, 107, 108, 109, 110, 148]. They demonstrate by means of numerical simulations for 1d signals that the two key functional analysis properties of wavelet Leaders, predicted by theory, hold in practice: Wavelet Leaders enable to estimate $\zeta(q)$, $q < 0$ and hence the decreasing part of the Legendre spectrum $D(h)$, and they can cope with oscillating singularities[1], whereas wavelet coefficients fail rather drastically in both cases. Nevertheless, these qualitative results leave most performance issues of practical importance unanswered: How do the respective performance of wavelet coefficient and Leader based estimation procedures compare? How do they depend on sample size? Do they depend on process parameters, and if yes, what is the nature of this dependence? The first aim of this chapter is precisely to study these and related issues in detail. This is reported in Section 4.1.

In this manuscript, we propose the extension of wavelet Leader based multifractal analysis to *2d images*, based on a practical implementation of 2d wavelet Leaders Eq. (2.55). Although the feasibility of practically implementing a 2d WLMF had been shown in [110], no relevant implementation of this 2d multifractal formalism existed until now. Hence, no results are available for investigating the practical relevance of the WLMF for image data: Can the proposed procedures provide relevant estimates of multifractal attributes for standard size images? Do they, for instance, practically enable the estimation of a non-zero c_2 or c_3? Therefore, the second aim of this chapter is to validate, by means of numerical simulations, the statistical performance of a 2d wavelet Leader based multifractal analysis, and to compare it to that obtained by a wavelet coefficient based one. This is addressed in Section 4.2, together with a discussion of practical limitations of the 2d WLMF.

Note that it would have been imaginable to combine the results and discussions in Sections 4.1 and 4.2 for 1d and 2d data, respectively, and to present them jointly. However, in an increasing number of applications, the data involved in analysis consists of images,

[1] See also Fig. 2.5.

due to significant progresses in high-resolution digital imaging. Also, the 2d WLMF has never been practically implemented before and constitutes an original contribution of this work. For these reasons, we choose to investigate the 1d and the 2d case separately.

Finally, the ultimate goal of the present chapter can be seen in providing an answer to the following questions: Do the estimation procedures proposed in Section 2.6 exhibit satisfactory statistical performance, and should one prefer wavelet coefficients or Leaders for the estimation of multifractal attributes of signals or images? The numerical results presented in Sections 4.1 and 4.2 enable us to conclude on this question and suggest that overall, wavelet Leader based estimation significantly outperforms wavelet coefficient based one and practically enables the precise multifractal characterization for both signals and images. Results presented in this chapter have been partly reported in [173, 175, 177, 179, 180, 182].

4.1 Statistical Performance of Estimation Procedures: Signals

4.1.1 Estimation performance: Preliminary results

Before diving into the details of the statistical performance and properties of the estimation procedures for 1d signals, we choose to present some qualitative overview results, similar to those that can be found in [8, 107, 108, 109, 110, 148]. These results allow us, on one hand, to appreciate the overall performance of wavelet coefficient and Leaders based estimation, and on other hand, to draw first preliminary conclusions which subsequently help us to concentrate on a more detailed analysis of the different aspects of statistical performance.

For convenience, we denote "wavelet coefficient (Leader) based estimation" briefly by d_X (L_X), respectively. Also, we will distinguish wavelet coefficient (Leader) related quantities by superscripts ^{d} (^{L}), respectively.

Fig. 4.1 summarizes mean estimations for fBm ($H = 0.8$, top row) and CPM-MF-fBm ($(c_1, c_2, c_3) = (0.8, -0.08, 0.031)$, bottom row) for sample size $N = 2^{14}$, for d_X (left column) and L_X (right column), obtained by numerical simulation with $N_{MC} = 1000$ realizations. In each subfigure, the subplots correspond to estimations for $\zeta(q)$ (top left), $D(q)$ $h(q)$ (top right), and c_1, c_2, c_3 (bottom), obtained with linear regressions with weights w^1 over scales $(j_1, j_2) = (3, 11)$. The figure compares the theoretical attributes (blue solid lines) to mean over Monte Carlo realizations of estimates (black solid lines) and shows the upper and lower 5% quantiles of the Monte Carlo distributions of the estimates (red solid lines). The subplots for c_p additionally show boxplots of these Monte Carlo distributions (black solid lines: median, upper and lower quartiles; black dashed-dotted lines: support of the distributions).

Fig. 4.1 indicates that the estimations of $\zeta(q)$, $D(q)$ and $h(q)$ for $q > 0$ (this corresponds to the increasing part of the spectrum $D(h(q))$) for both d_X and L_X are very close to each other and to the theoretical values, for both the mono- (fBm) and multifractal (CPM-MF-fBm) process. Similarly, the estimations of c_1 are roughly equivalent for both processes.
In contrast, the estimations for c_2 and c_3, capturing the deviation from monofractal, are significantly better for L_X for both processes. In particular, Monte Carlo confidence inter-

vals for c_2 and c_3 of CPM-MF-fBm exclude the zero value. This indicates that non-zero c_2 and c_3 can indeed be estimated from data by the wavelet Leader based procedure. Also, for d_X, the estimates of $\zeta(q)$, $D(q)$ and $h(q)$ for $q < 0$ (corresponding to the decreasing part of the spectrum $D(h(q))$) are disconnected from the theoretical values and meaningless, whereas for L_X, estimates closely follow the theoretical values. This behavior for d_X for $q < 0$ is systematic and due to numerical instabilities of structure functions for negative orders q (see e.g. [107] and Section 2.5.5). Therefore, we will not further discuss d_X for $q < 0$ here.

First conclusions. These results suggest the following preliminary conclusion. The fact that d_X can not be used for exploring negative statistical orders q, together with them having significantly inferior performance for the estimation of c_2 and c_3, makes them rather inadequate for practically performing a multifractal analysis of multifractal processes. It would, for instance, be very difficult to decide, based on d_X, whether the multifractal spectrum of CPM-MF-fBm is symmetric or not (either from $D(h(q))$ directly, or from c_3), whereas the Monte Carlo based 90% confidence interval for c_3 for L_X excludes the value zero, giving strong evidence to the fact that the spectrum is not symmetrical.
Let us now proceed to a more detailed analysis.

4.1.2 Numerical simulation study

4.1.2-a) Monte Carlo simulations

To evaluate the performance of both the wavelet coefficient and Leader based estimation procedures Eq. (2.65), Eq. (2.71) and Eqs. (2.74–2.75), we apply them to a large number N_{MC} of 1d realizations of sample size N of mono- and multifractal processes X with known and a priori controlled multifractal properties and attributes. The estimations $\hat{\theta}$ are performed for all possible combinations of ranges of scales $[j_1, j_2]$, $j_2 - j_1 \geq 2$ for linear regressions, and for both ordinary (w^0) and weighted (w^1) linear regressions (cf. Section 2.6). For each multifractal attribute θ and parameter setup, we evaluate the *bias* (b):

$$\hat{b}_\theta = \widehat{\mathbb{E}}_{MC}\,\hat{\theta} - \theta, \tag{4.1}$$

the *standard deviation* (s):

$$\hat{s}_\theta = \widehat{\mathrm{Std}}_{MC}\,\hat{\theta} = \sqrt{\widehat{\mathbb{E}}_{MC}\,\hat{\theta}^2 - \left(\widehat{\mathbb{E}}_{MC}\,\hat{\theta}\right)^2} \tag{4.2}$$

and the *(root) mean squared error* (mse):

$$\widehat{mse}_\theta = \sqrt{\hat{b}_\theta^2 + \hat{s}_\theta^2}, \tag{4.3}$$

of the estimation, where $\widehat{\mathbb{E}}_{MC}$ stands for mean over Monte Carlo realizations. Notice that $\hat{b}_\theta$, $\hat{s}_\theta$ and $\widehat{mse}_\theta$ depend on j_1, j_2, w_i, N, d or L, the process and the process parameters (explicit notation taking into account these dependences are omitted here for convenience of notation).
Finally, we define the *minimum (root) mean squared error* ($mmse$) for given w^i, N, d or L, process and process parameters:

$$\widehat{mmse}_\theta = \inf_{[j1,j2]} \widehat{mse}_\theta, \tag{4.4}$$

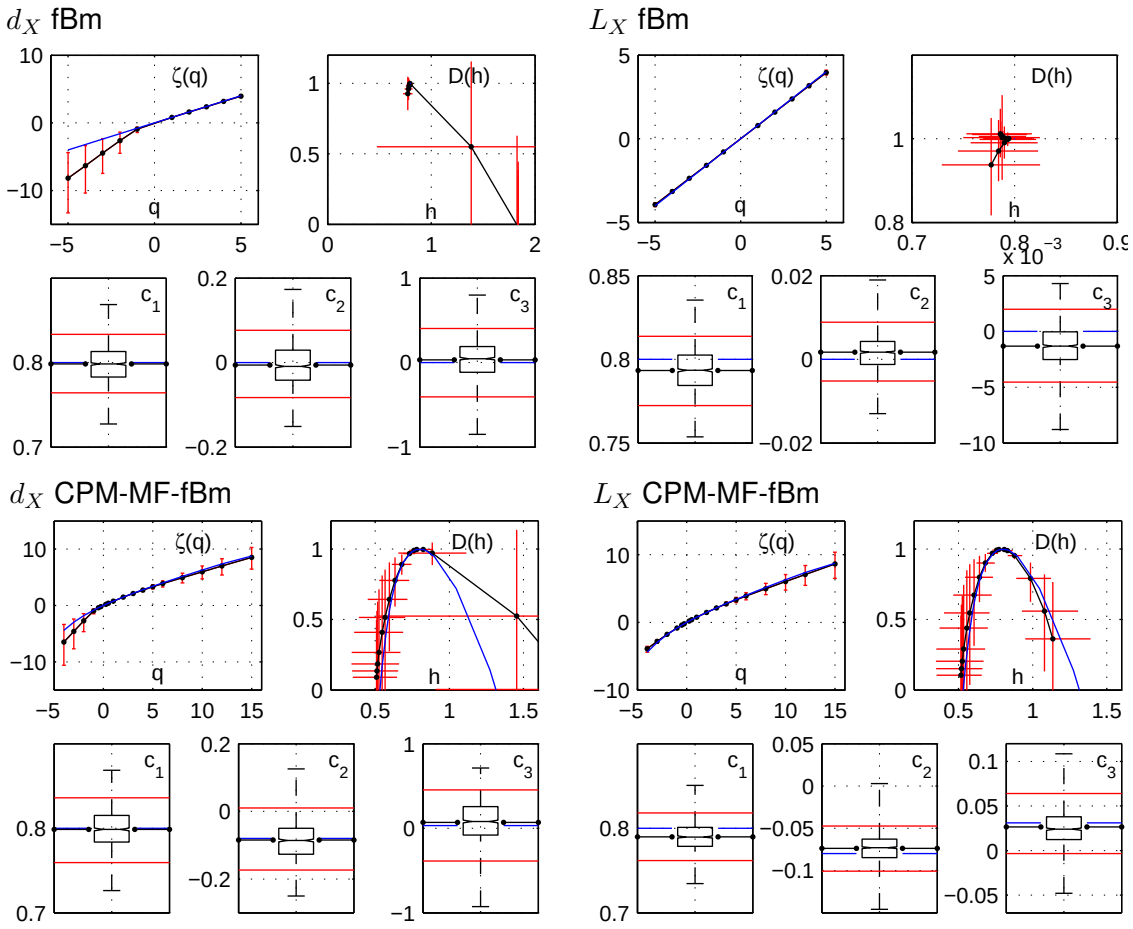


Figure 4.1: **Estimation performance.** Wavelet coefficient (left column) and Leader (right column) based estimation for fBm (top row) and CPM-MF-fBm (bottom row) for $N = 2^{14}$ and w^1 over $(j_1, j_2) = (3, 11)$. The subplots correspond to $\zeta(q)$ (to left), $D(q)$, $h(q)$ (top right) and c_p (bottom). Shown are the theoretical values of the attributes (solid blue), means of estimates over $N_{MC} = 1000$ Monte Carlo realizations (solid black, '•') and upper and lower 5% quantiles of Monte Carlo distributions (solid red vertical and horizontal bars).

which quantifies the best possible estimation performance based on d_X or on L_X for given process setup and regression weights. The $mmse$ is a quantity that can not be assessed in practice apart from by numerical simulations on synthetic processes, since the optimal (j_1, j_2) are in general unknown in practice: The couple (j_1, j_2) is usually fixed by eye based on the experience of the practitioner. By precisely avoiding the practical regression range selection issue, the $mmse$ therefore provides a basis for a fair and objective performance comparison.

4.1.2-b) Simulation and process parameters

Simulation parameters. For the numerical simulation presented here, we use Daubechies wavelets with $N_\psi = 3$ vanishing moments (the influence of N_ψ on estimation has been studied elsewhere, e.g. [4, 6, 12, 70], and will not be considered here). The simulation parameters are set to $N_{MC} = 1000$ and $N \in \{2^{10}, 2^{14}, 2^{18}\}$, and the multifractal

attributes considered are $\theta \in \{c_1, c_2, c_3, \zeta(q), D(q), h(q)\}$ for $q \in \{-3, -2, -1, 1, 2, 3\}$.

fBm. The self-similarity parameter of (monofractal) Gaussian fractional Brownian motion is set to $H \in \{0.1, 0.2, \cdots, 0.9\}$ and therefore includes $H = 0.5$ (ordinary Brownian motion with independent increments) and the range $H > 0.5$ for which the process is long range dependent.

ROS. The self-similarity parameter of the non Gaussian fv H-sssi Rosenblatt process is set to $H \in \{0.5, 0.6, 0.7, 0.8, 0.9\}$.

CPM. The process parameters for compound Poisson motion are set to $\mu = -0.005$ and $\sigma^2 \in \{0.01, 0.03, 0.05, 0.07, 0.09\} - \mu^2$, such that it has second log-cumulant $c_2 \in \{-0.01, -0.03, -0.05, -0.07, -0.09\}$.

CPM-MF-fBm. The parameters of CPM-MF-fBm are set to $H \in \{0.5, 0.6, 0.7, 0.8, 0.9\}$, $\mu = -0.005$ and $\sigma^2 \in \{0.01, 0.03, 0.05, 0.07, 0.09\} - \mu^2$, such that c_1 is in the range $[0.51, 0.95]$, $c_2 \in \{-0.01, -0.03, -0.05, -0.07, -0.09\}$, and c_3 is in the range $[0.0001, 0.0012]$.

mrw. The process parameters (H, β) of mrw are set such that $c_1 \in \{0.6, 0.7, 0.8, 0.9\}$ and $c_2 \in \{-0.01, -0.03, -0.05, -0.07, -0.09\}$.

The definitions and the scaling and multifractal properties of the above fv H-sssi processes are stated in Section 2.2.3, and those of MMC processes CPM, CPM-MF-fBm and mrw in Sections 2.7.1-b), 2.7.2-a) and 2.7.2-b), respectively.

4.1.2-c) A note on the presented results

The simulation study we conduct includes a total of 64 different process (parameter) settings. For each of these settings, 3 sample sizes N and two different linear regression weights (w^0) and (w^1) are considered, and for each of these now $64 \times 3 \times 2$ settings, we calculate the bias, standard deviation and mean squared error for a total of 21 estimates. If we in addition consider that for the three sample sizes, the largest available scales are $j_{2,max} \in \{7, 11, 15\}$ and thus there are $\{15, 45, 91\}$ combinations of (j_1, j_2) for performing the linear regressions, it is obvious that it would be rather meaningless and, for space reasons, as well impossible to present these results in all detail within this manuscript (the author leaves it to the reader to calculate the total number of table entries this would require).

The goal of this numerical study is to characterize the impact of the different factors that potentially influence estimation performance: The precise range of scales chosen for linear regressions, the weights used in linear regressions, the precise nature of the estimate, the sample size, the nature of the processes or the precise value of process parameters. The presentation of the numerical results in the following sections therefore aims at systematically disentangling the effects of the various factors that potentially influence estimation performance. Hence, we will proceed in the following way: For each issue evoked above and grouped in the following sections, we will present a selection of results that are representative and illustrate the discussions and conclusions, valid for the entire simulation study. For the reader who is interested in more details, a large collection of results is available upon request.

Also, since the multifractal attributes $\zeta(q)$, $D(h)$ and c_p are closely related to each other and essentially measure the same aspects of data, we concentrate on results for $\zeta(q)$ and c_p only in the following sections.

fBm	$\zeta(2)$		c_1		c_2	
H	d_X	L_X	d_X	L_X	d_X	L_X
0.1	0.085	0.019	0.050	0.011	0.024	0.005
0.2	0.055	0.009	0.033	0.004	0.026	0.004
0.3	0.045	0.017	0.026	0.022	0.025	0.003
0.4	0.037	0.026	0.022	0.020	0.025	0.003
0.5	0.032	0.040	0.021	0.011	0.025	0.005
0.6	0.032	0.039	0.020	0.006	0.025	0.003
0.7	0.031	0.033	0.017	0.007	0.024	0.003
0.8	0.027	0.031	0.016	0.009	0.025	0.004
0.9	0.025	0.031	0.015	0.010	0.025	0.005
ROS	$\zeta(2)$		c_1		c_2	
H	d_X	L_X	d_X	L_X	d_X	L_X
0.6	0.036	0.034	0.020	0.016	0.027	0.008
0.7	0.053	0.062	0.024	0.025	0.027	0.011
0.8	0.070	0.082	0.032	0.033	0.028	0.015
0.9	0.081	0.089	0.039	0.030	0.033	0.022

Table 4.1: **Minimum mse for H-sssi processes.** Best case estimation performance ($mmse$) for Gaussian fBm (top) and non Gaussian ROS (bottom) for different values of the self-similarity parameter H. Estimations are performed for sample size $N = 2^{14}$ using weighted linear regressions with weights w^1.

4.1.3 Estimation performance: Minimum mean squared error

The goals of this section are to provide guidelines for deciding for which processes and for which process parameters wavelet coefficient or Leader based procedures are to be preferred, and to study which performance these procedures can achieve in practice, depending on the process nature and parameters. To this end, we compare the $mmse$ obtained for Gaussian H-sssi, non Gaussian H-sssi processes, and multifractal processes for various different process parameter settings.

4.1.3-a) H-sssi processes

We choose to base performance comparisons for fBm and ROS on the multifractal attributes $\zeta(2)$, c_1 and c_2: $\zeta(2)$ is classically considered for the estimation of H, c_1 is the log-cumulant equivalent of H (for H-sssi, $c_1 = H$), and c_2 measures first departure from H-sssi (i.e., $c_2 = 0$ for fv H-sssi). Results are summarized in Tab. 4.1 for sample size $N = 2^{14}$ and weighted regression w^1 for fBm (top) and ROS (bottom).

Relative performance of wavelet coefficients and Leaders. The following observations are valid for both Gaussian fBm and non Gaussian ROS. First, for estimating $\zeta(2)$, d_X perform slightly better than L_X for $H \geq 0.5$, and L_X better than d_X for $H < 0.5$. The first observation is consistent with the fact that d_X are close to optimal for the estimation of H for H-sssi processes. The latter observation, L_X better than d_X for $H < 0.5$, therefore seems surprising at first. It can be interpreted in the light of results presented

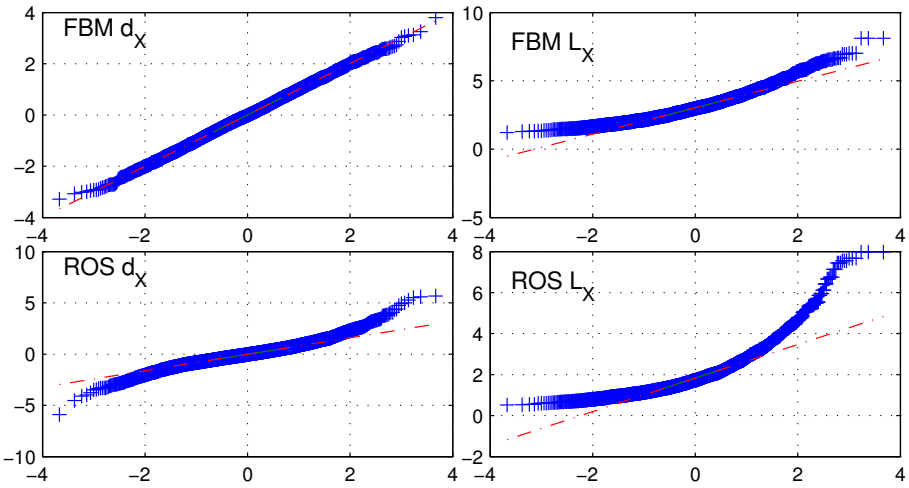


Figure 4.2: **qq plots fBm versus ROS.** Quantile-quantile plots of the empirical distributions of wavelet coefficients $d_X(j,\cdot)$ (left column) and Leaders $L_X(j,\cdot)$ (right column) against standard Normal for fBm (top row) and ROS (bottom row) ($j = 3$, $H = 0.7$, $N = 2^{15}$).

in Section 3.1 as follows: It is known that the estimation of H can be biased for small and medium sample sizes, and the bias of d_X and L_X are in principle comparable if the same range of scales are chosen for regressions. Results in Section 3.1 suggest that wavelet Leader based structure functions do not behave linearly in log-log coordinates at fine scales due to fine scale cutoff. This non-linear behavior of the structure functions can, though not meaningful, be used for adjusting for bias by choosing an appropriate regression range at medium or fine scales. For wavelet coefficients, this is not the case, and bias can only be reduced by performing linear fits at coarse scales, hence increasing standard deviation. Thus, in this particular setting, the $mmse$ measure allows L_X to "cheat" on d_X by choosing a regression range that no practitioner would select since the structure functions in this range are curved.
For the estimation of c_1, the performance of d_X and L_X are similar, with slightly smaller overall $mmse$ for L_X. The gain in performance is, however, small. The situation is completely different for the estimation of c_2, for which L_X significantly outperforms d_X, with $mmse$ gains of up to nearly one order of magnitude. Note, however, that when it is known that the process underlying the analysis is H-sssi, there is no particular interest in estimating c_2, since $c_2 \equiv 0$.

Relative estimation performance for fBm and ROS. We observe that the non Gaussian nature of ROS significantly decreases estimation performance as compared to fBm, for both d_X and L_X. This additional estimation difficulties for ROS can be interpreted in the following ways: First, the marginals of wavelet coefficients $d_X(j,\cdot)$ and wavelet Leaders $L_X(j,\cdot)$ of ROS are strongly non Gaussian and skewed (cf. Figure 4.2). Second, whereas the absolute values of wavelet coefficients $|d_X(j,\cdot)|$, on which estimations are based, are only weakly correlated for fBm, this turns out not to be the case for ROS (cf. analysis and discussion in Chapter 11.1).

Influence of H on estimation performance for H-sssi processes. Tab. 4.1 sug-

gests that, for ROS, there is a general increase of $mmse$ with increasing H, for both wavelet coefficient and Leader based estimation, and for all estimates. For fBm, the situation is more complicated: Whereas wavelet Leader based estimation performance degrades with increasing H for $\zeta(2)$ and stays constant for c_1 and c_2, wavelet coefficient based estimation performance increases with H for both $\zeta(2)$ and c_1 and stays constant for c_2. Hence, there is a clear impact of the precise value of H on the best case estimation performance. This is studied in more detail in Subsection 4.1.7.

4.1.3-b) Multifractal processes

For comparing performance for multifractal processes, we investigate estimates for $\zeta(-2)$ and c_3 in addition to the attributes chosen above for H-sssi processes: $\zeta(-2)$ represents estimates involving negative statistical moments q, and c_3 measures the second order departure from monofractal by capturing the asymmetry of the multifractal spectrum, when c_2 captures its width. Results are summarized in Tab. 4.2 for sample size $N = 2^{14}$ and weighted regression w^1 for CPM (top), CPM-MF-fBm (center) and mrw (bottom).

Relative performance of wavelet coefficients and Leaders. Tab. 4.2 suggests that for the estimation of $\zeta(2)$, d_X perform in general slightly better than L_X, with the exception of CPM-MF-fBm and mrw when c_1 is close to 0.5, for which L_X are preferable. Similarly, the performance of d_X and L_X are practically equivalent for the estimation of c_1, with slightly better overall performance for L_X, in particular for mrw. The gain in performance is, however, small.
In contrast, for the estimation of c_2 and c_3, wavelet Leader based estimation significantly outperforms wavelet coefficient based one, with $mmse$ gains of up to one order of magnitude. These observations are consistent with those obtained for H-sssi processes. Nevertheless, Tab. 4.2 seem to suggest that measurements of non-zero c_3 (for CPM and CPM-MF-fBm) are difficult to obtain in practice. Even with wavelet Leaders, $mmse$ are one order of magnitude larger than the c_3 to be estimated. However, the values of c_3 considered here are very small, and if c_3 and the sample size N are sufficiently large, a non-zero c_3 can be statistically significantly measured: This is illustrated in Fig. 4.1, where a $c_3 \approx 0.03$ is non-ambiguoulsy distinguished from zero value for sample size $N = 2^{15}$ using wavelet Leader based estimation. This is not feasible with coefficient based estimation, since the MC confidence interval has span three orders of magnitude larger than the value of c_3.
Let us also appreciate the performance of wavelet Leader based estimations of $\zeta(-2)$, which display $mmse$ close to those of $\zeta(2)$ for all processes and parameter settings. Hence, estimation for negative statistical moments does not pose any particular problems for wavelet Leaders as compared to positive statistical moments.

Influence of c_p on estimation performance. Tab. 4.2 suggests that for CPM and CPM-MF-fBm, an increase of $|c_p|$ and hence stronger departure from monofractal leads overall to increased $mmse$, for both d_X and L_X. Note that it is impossible to disentangle the influence of increase or decrease of one particular c_p for these processes, since the c_p, $p \geq 1 \neq 0$ can not all be independently controlled by two (three) process parameters, respectively.
Similarly, for mrw, $mmse$ are significantly larger for strongly negative c_2 values and thus strong departure from monofractal. The precise value of c_1 has little to no effect on $mmse$.

CPM			$\zeta(-2)$		$\zeta(2)$		c_1		c_2		c_3	
c_1	c_2	c_3	d_X	L_X	d_X	L_X	d_X	L_X	d_X	L_X	d_X	L_X
1.005	−0.01	0.0001	−	0.067	0.034	0.041	0.032	0.027	0.042	0.011	0.186	0.007
1.015	−0.03	0.0004	−	0.077	0.039	0.053	0.033	0.031	0.042	0.016	0.187	0.011
1.025	−0.05	0.0007	−	0.087	0.045	0.066	0.035	0.034	0.043	0.019	0.185	0.017
1.035	−0.07	0.0010	−	0.101	0.056	0.077	0.036	0.037	0.044	0.023	0.193	0.024
1.046	−0.09	0.0013	−	0.110	0.065	0.086	0.038	0.040	0.044	0.027	0.193	0.032
CPM-MF-fBm			$\zeta(-2)$		$\zeta(2)$		c_1		c_2		c_3	
c_1	c_2	c_3	d_X	L_X	d_X	L_X	d_X	L_X	d_X	L_X	d_X	L_X
0.552	−0.05	0.0004	−	0.067	0.058	0.034	0.024	0.021	0.058	0.012	0.287	0.014
0.643	−0.05	0.0004	−	0.058	0.049	0.047	0.021	0.023	0.041	0.012	0.256	0.014
0.736	−0.05	0.0005	−	0.043	0.043	0.054	0.019	0.020	0.027	0.011	0.171	0.012
0.832	−0.05	0.0006	−	0.042	0.041	0.050	0.017	0.018	0.029	0.011	0.153	0.012
0.928	−0.05	0.0007	−	0.044	0.046	0.051	0.018	0.018	0.028	0.012	0.133	0.013
0.707	−0.01	0.0001	−	0.037	0.032	0.035	0.017	0.007	0.026	0.004	0.135	0.004
0.722	−0.03	0.0003	−	0.041	0.037	0.046	0.018	0.014	0.026	0.008	0.128	0.008
0.736	−0.05	0.0005	−	0.043	0.043	0.054	0.019	0.020	0.027	0.011	0.171	0.012
0.752	−0.07	0.0007	−	0.057	0.055	0.051	0.020	0.020	0.039	0.014	0.199	0.017
0.767	−0.09	0.0009	−	0.077	0.067	0.059	0.020	0.018	0.055	0.022	0.305	0.032
mrw			$\zeta(-2)$		$\zeta(2)$		c_1		c_2		c_3	
c_1	c_2	c_3	d_X	L_X	d_X	L_X	d_X	L_X	d_X	L_X	d_X	L_X
0.600	−0.05	0	−	0.054	0.058	0.050	0.028	0.019	0.029	0.011	0.134	0.011
0.700	−0.05	0	−	0.064	0.054	0.061	0.030	0.009	0.028	0.009	0.130	0.012
0.800	−0.05	0	−	0.066	0.051	0.060	0.030	0.009	0.028	0.011	0.129	0.012
0.900	−0.05	0	−	0.062	0.048	0.058	0.027	0.010	0.028	0.012	0.130	0.012
0.700	−0.01	0	−	0.040	0.034	0.038	0.019	0.007	0.025	0.005	0.128	0.004
0.700	−0.03	0	−	0.051	0.046	0.050	0.026	0.007	0.026	0.008	0.129	0.008
0.700	−0.05	0	−	0.064	0.054	0.061	0.030	0.009	0.028	0.009	0.130	0.012
0.700	−0.07	0	−	0.071	0.065	0.065	0.034	0.012	0.030	0.012	0.133	0.017
0.700	−0.09	0	−	0.074	0.075	0.060	0.039	0.015	0.034	0.015	0.140	0.023

Table 4.2: **Minimum mse for multifractal processes.** Best case estimation performance ($mmse$) for CPM (top), CPM-MF-fBm (center rows) and mrw (bottom) for different combinations of log-cumulants c_p. Estimations are performed for sample size $N = 2^{14}$ using weighted linear regressions with weights w^1.

Relative estimation performance for CPM, CPM-MF-fBm and mrw. The impact of process nature – $A(t)$ (CPM) vs. $V_H(t)$ (CPM-MF-fBm) vs. $Z_H(t)$ (mrw) – should ideally be evaluated by comparisons for one common setup of values for c_p. This is, however, not possible since $c_p \equiv 0$, $p \geq 3$ for mrw whereas these log-cumulants are non-zero for CPM and CPM-MF-fBm, and since $c_1 > 1$ for CPM whereas $c_1 < 1$ for CPM-MF-fBm and mrw. Nevertheless, Tab. 4.2 suggests that for common sample size N, estimation performance is overall practically comparable for the three processes and vary, for each estimate, less than one order of magnitude among different processes, regardless of the

Process		$\zeta(2)$		c_1		c_2	
		d_X	L_X	d_X	L_X	d_X	L_X
fBm	$c_1 = 0.7$	0.031	0.033	0.017	0.007	0.024	0.003
ROS	$c_1 = 0.7$	0.053	0.062	0.024	0.025	0.027	0.011
mrw	$c_1 = 0.7\ c_2 = -0.05$	0.054	0.061	0.030	0.009	0.028	0.009

Table 4.3: **Minimum mse for H-sssi and multifractal processes.** Best case estimation performance ($mmse$) for Gaussian fBm (first line) and non Gaussian ROS (second line) and non Gaussian multifractal mrw (third line) with common $c_1 = 0.7$. Estimations are performed for sample size $N = 2^{14}$ using weighted linear regressions with weights w^1.

precise values of process parameters.

4.1.3-c) Mono- versus multifractal processes

Let us finally compare $mmse$ performance for monofractal fBm and ROS to that for multifractal mrw for one common $c_1 = 0.7$ and $c_2 = -0.05$ for mrw, and with estimation parameters and sample size set as above. Tab. 4.3 summarizes $mmse$ for estimation of $\zeta(2)$, c_1 and c_2 based on wavelet coefficients and Leaders for these three processes. It suggests that, first, estimation performance is better for fBm than for ROS and mrw for all estimates and both wavelet coefficients and Leaders. Second, for wavelet coefficient based estimation, $mmse$ are larger for multifractal mrw than for non Gaussian H-sssi ROS, whereas for wavelet Leader based estimation, the situation is converse: estimation performance is better for multifractal mrw than for ROS.

4.1.3-d) Conclusions

The minimum mean squared error results presented in this section lead us to the following conclusions. The best case estimation performance is practically equivalent for wavelet coefficient and Leader based estimation of $\zeta(2)$ and c_1 for both (Gaussian or non Gaussian) H-sssi and for multifractal processes. Therefore, if it is a priori known that the process under analysis is H-sssi, the estimation of the self-similarity parameter H should be based on wavelet coefficients, since the comparable performance of wavelet coefficient and Leader based procedures does not justify the additional complexity of wavelet Leaders. In contrast, wavelet Leader based estimation of c_p, $p \geq 2$ significantly outperform wavelet coefficient based one. Also, estimates of $\zeta(q < 0)$ involving negative statistical orders are relevantly estimated when using wavelet Leaders, whereas they can not be measured using wavelet coefficients. Therefore, wavelet Leaders based estimations have to be employed for a relevant multifractal analysis in any situation in which it can not be a priori ensured that data are H-sssi. Also, their use is practically mandatory for the discrimination of mono- vs. multifractal models, and among different multifractal models.

Estimation is significantly more difficult for non Gaussian H-sssi processes than it is for Gaussian fBm. Similarly, estimation is more difficult for the multifractal processes considered, and estimation performance decreases with increasing departure from monofractal. These results can be interpreted in the light of result presented in Chapter 11, which

suggest that statistical properties of wavelet coefficients are more difficult for these processes than they are for fBm.

The conclusions for H-sssi processes being practically coherent, we use fBm with moderate $H = 0.7$ as a model example for monofractal processes in the following sections and only resort to ROS when its non Gaussian nature is explicitly a concern for the topic studied. Also, no significant difference in estimation performance for the multifractal processes considered being observed, we employ mrw as our multifractal model process in the following sections: It is easy to simulate and ensures independent control of c_1 and c_2. We set them to (moderate) values $c_1 = 0.7$ and $c_2 = -0.05$.

4.1.4 Weighted versus ordinary linear regressions

The weighted regressions w^1 have been suggested for wavelet coefficient based estimations for fBm and are asymptotically optimal for this process (see e.g. [4]). It is, however, not clear whether they should be preferred over ordinary regressions w^0 for other processes or for wavelet Leader based estimation. Tab. 4.4 summarizes $mmse$ results for fBm (top, $H = 0.7$) and mrw (bottom, $c_1 = 0.7$, $c_2 = -0.05$) of sample size $N = 2^{14}$ for ordinary (w^0) and weighted (w^1) linear regression (for additional results and alternative choices for weights, see [173, 180]). It suggests that for fBm, estimations using weighted regression clearly and systematically outperform their ordinary regression based counterparts, for both wavelet coefficient and Leader based estimation. The performance gains of w^1 over w^0 reach up to a factor of 2 for fBm. Tab. 4.4 (bottom) also indicates similar conclusions for multifractal mrw: Weighted regression results in general in better performance than ordinary regression and achieves performance gains over w^0 of up to a factor of 2. This is remarkable since for such processes, the assumptions for the derivation of w^1 (notably Gaussian and independent coefficients at each scale) are not valid and therefore, weights w^1 are clearly suboptimal.

Conclusions. We conclude that for both mono- and multifractal processes, and for both wavelet coefficient and Leader based estimation, weighted linear regressions with weights w^1 are preferable over ordinary linear regressions.

fBm			$\zeta(-2)$		$\zeta(2)$		c_1		c_2		c_3	
H		weights	d_X	L_X	d_X	L_X	d_X	L_X	d_X	L_X	d_X	L_X
0.70		w^0	$-$	0.042	0.038	0.044	0.024	0.013	0.040	0.005	0.205	0.002
		w^1	$-$	0.037	0.031	0.033	0.017	0.007	0.024	0.003	0.118	0.002
mrw			$\zeta(-2)$		$\zeta(2)$		c_1		c_2		c_3	
c_1	c_2	weights	d_X	L_X	d_X	L_X	d_X	L_X	d_X	L_X	d_X	L_X
0.7	-0.05	w^0	$-$	0.049	0.062	0.065	0.034	0.010	0.049	0.014	0.236	0.009
		w^1	$-$	0.064	0.054	0.061	0.030	0.009	0.028	0.009	0.130	0.012

Table 4.4: **Estimation performance for ordinary versus weighted linear regressions.** Performance ($mmse$) of wavelet coefficient and Leader based estimation for fBm (top) and mrw (bottom) using ordinary (w^0, first line) and weighted (w^1, second line) linear regressions and sample size $N = 2^{14}$.

fBm - $H = 0.7$												
	$\zeta(2)$				c_1				c_2			
	$j_1 = 2$		$j_1 = 3$		$j_1 = 2$		$j_1 = 3$		$j_1 = 2$		$j_1 = 3$	
	d_X	L_X	d_X	L_X	d_X	L_X	d_X	L_X	d_X	L_X	d_X	L_X
b	-0.023	-0.053	-0.011	-0.018	-0.009	-0.029	-0.002	-0.011	-0.004	0.000	-0.006	0.003
s	0.021	0.019	0.032	0.028	0.014	0.009	0.022	0.013	0.035	0.003	0.053	0.004
mse	0.031	0.056	0.034	0.033	0.017	0.030	0.022	0.017	0.035	0.003	0.054	0.005
mrw - $c_1 = 0.7$, $c_2 = -0.05$												
	$\zeta(2)$				c_1				c_2			
	$j_1 = 2$		$j_1 = 3$		$j_1 = 2$		$j_1 = 3$		$j_1 = 2$		$j_1 = 3$	
	d_X	L_X	d_X	L_X	d_X	L_X	d_X	L_X	d_X	L_X	d_X	L_X
b	-0.041	-0.062	-0.030	-0.038	-0.029	-0.034	-0.019	-0.025	0.008	0.000	0.004	0.006
s	0.035	0.034	0.052	0.049	0.015	0.011	0.023	0.016	0.040	0.010	0.060	0.013
mse	0.054	0.070	0.060	0.061	0.033	0.036	0.030	0.029	0.041	0.010	0.060	0.015

Table 4.5: **Performance for fixed regression ranges.** Estimation performance for fixed regression ranges $j_1 = \{2, 3\}$, $j_2 = 11$ for fBm (top) and mrw (bottom) with sample size $N = 2^{14}$, and weighted regressions w^1. Shown are bias (first line), standard deviation (second line) and mean squared error (third line) for wavelet coefficient (left) and Leader (right) based estimation procedures for both regression ranges.

4.1.5 Estimation performance: Fixed regression range

The results in Chapter 3.1 suggest that for d_X and L_X, inspection of the structure functions do not necessarily lead to the same conclusions on the range of scales that should be involved in linear regressions. Indeed, wavelet Leaders need a number of fine scales for initialization before practically becoming Leaders that are uninfluenced by the fine scale cutoff. This translates to a non-linear behavior of structure functions at fine scales for wavelet Leaders, whereas wavelet coefficients are not subject to such limitations if the DWT has been properly initialized [3, 172]. Thus, j_1 would in practice typically be chosen larger for L_X than for d_X.
The goal of this section is to account for this discrepancy and to study the relative performance of coefficient and Leader based estimation for their "natural" regression range setup, i.e. regression ranges as they would be chosen by a practitioner, rather than theoretically best achievable performance as discussed above. Tab. 4.5 summarizes results for fBm (top) and mrw (bottom), for which inspection of the structure functions (not shown here) suggest $j_1 = 2$ for coefficients and $j_1 = 3$ for Leaders. Results are obtained for sample size $N = 2^{14}$, weighted regressions w^1 and for linear regressions over scales $(j_1, j_2) = (2, 11)$ and $(j_1, j_2) = (3, 11)$. Shown are bias (first line), standard deviation (second line) and mean squared error (third line) for wavelet coefficient (left) and Leader (right) based estimation.

Bias vs. j_1. Tab. 4.5 suggests that the bias for any estimate and for both processes is consistently smaller for $j_1 = 3$ than for $j_1 = 2$. This is in agreement with observations reported in [171] for wavelet coefficients and H-sssi processes.
The results indicate that when the same regression range is chosen for both d_X and L_X, i.e. $j_1^d = j_1^L = 2$ or 3, the bias b^d of coefficient based estimation is (with the exception of c_2) consistently smaller than the bias b^L for L_X for both processes, and significantly so

for fBm. However, if the structure function inspection based scales $j_1^d = 2$ and $j_1^L = 3$ are chosen for linear regressions, $b^d \approx b^L$. Therefore, in practice, Leader based estimation would usually not be more biased than coefficient based one.

Standard deviation vs. j_1. Tab. 4.5 suggests that no matter what j_1 is chosen for d_X or L_X, L_X has consistently smaller standard deviation than d_X for both the Gaussian H-sssi fBm and the multifractal mrw. This difference is especially pronounced for the estimation of c_2, for which L_X achieve standard deviation gains of up to one order of magnitude.
Also, standard deviations are obviously smaller for $j_1 = 2$ than for $j_1 = 3$, for any estimate and for both processes, since the effective number of samples actually used in estimation is larger. Therefore, the choice of j_1 can be interpreted as a bias-variance tradeoff, as has been discussed in [171] for wavelet coefficients and H-sssi processes.

Mean squared error vs. j_1. Overall estimation performance in terms of mean squared error, capturing the interplay of bias and standard deviation, suggests that for d_X, mse are smaller for $j_1 = 2$, for L_X, mse are smaller for $j_1 = 3$. This confirms the structure function based practical choice of scales for linear regressions. Also, when choosing equal regression range for d_X and L_X, i.e. $j_1^d = j_1^L$, d_X have smaller mse than L_X for the estimation of $\zeta(2)$ and c_1 for $j_1 = 2$, whereas for $j_1 = 3$, the situation is converse, and L_X have overall slightly better performance than d_X. For the estimation of c_2, L_X consistently perform better. For scales $j_1^d = 2$ and $j_1^L = 3$, practically suggested by structure functions, the estimation of $\zeta(2)$ and c_1 based on wavelet coefficients and Leaders display comparable performance, whereas L_X achieve performance gains of up to one order of magnitude for the estimation of c_2. These observations are valid for both monofractal and Gaussian fBm and multifractal mrw.

Conclusions. The results, obtained for a practically realistic estimation setup, are in good agreement with the best case performance study discussed above: Wavelet coefficient and Leader based procedures have approximately equivalent performance for the estimation of $\zeta(q), q > 0$ and c_1, and wavelet Leaders bring substantial performance gains for the estimation of c_2, for both monofractal and multifractal processes. Therefore, conclusions are similar: Wavelet coefficient based estimation is to be preferred over a wavelet Leader based one exclusively for the estimation of the self-similarity parameter of H-sssi processes. In general, for the multifractal analysis of real-world data of unknown nature, the use of wavelet Leaders is indispensable.

4.1.6 Sample size and statistical performance

The goal of this subsection is to study the behavior of wavelet coefficient and Leader based estimation procedures with respect to sample size N. To this end, Tab. 4.6 summarizes estimation performance for fBm (first and second part) and mrw (third and fourth part) for d_X (first and third part) and L_X (second and fourth part) and sample sizes $N \in \{2^{10}, 2^{14}, 2^{18}\}$. Shown are bias b (first line), standard deviation s (second line) and mean squared error mse (third line). Estimations are performed using linear regressions with weights w^1 over the scales $j_1 = 3$ to $j_2 \in \{7, 11, 15\}$, respectively.

Bias and sample size. We observe that the bias decreases systematically with in-

fBm	d_X											
H	$\zeta(-2)$			$\zeta(2)$			c_1			c_2		
0.7	2^{10}	2^{14}	2^{18}	2^{10}	2^{14}	2^{18}	2^{10}	2^{14}	2^{18}	2^{10}	2^{14}	2^{18}
b	$-$	$-$	$-$	-0.061	-0.011	-0.004	-0.005	-0.002	-0.001	-0.056	-0.006	-0.001
s	$-$	$-$	$-$	0.172	0.032	0.007	0.118	0.022	0.005	0.284	0.053	0.012
mse	$-$	$-$	$-$	0.182	0.034	0.008	0.118	0.022	0.005	0.289	0.054	0.012

fBm	L_X											
H	$\zeta(-2)$			$\zeta(2)$			c_1			c_2		
0.7	2^{10}	2^{14}	2^{18}	2^{10}	2^{14}	2^{18}	2^{10}	2^{14}	2^{18}	2^{10}	2^{14}	2^{18}
b	0.029	0.030	0.029	-0.051	-0.018	-0.013	-0.018	-0.011	-0.010	-0.005	0.003	0.004
s	0.122	0.025	0.006	0.143	0.028	0.007	0.064	0.013	0.003	0.022	0.004	0.001
mse	0.126	0.039	0.029	0.152	0.033	0.014	0.066	0.017	0.010	0.022	0.005	0.004

mrw		d_X											
c_1	c_2	$\zeta(-2)$			$\zeta(2)$			c_1			c_2		
0.7	-0.05	2^{10}	2^{14}	2^{18}	2^{10}	2^{14}	2^{18}	2^{10}	2^{14}	2^{18}	2^{10}	2^{14}	2^{18}
b		$-$	$-$	$-$	-0.103	-0.034	-0.024	-0.022	-0.019	-0.019	-0.057	0.001	0.006
s		$-$	$-$	$-$	0.209	0.050	0.015	0.123	0.022	0.005	0.272	0.056	0.013
mse		$-$	$-$	$-$	0.232	0.061	0.028	0.125	0.030	0.019	0.278	0.056	0.015

mrw		L_X											
c_1	c_2	$\zeta(-2)$			$\zeta(2)$			c_1			c_2		
0.7	-0.05	2^{10}	2^{14}	2^{18}	2^{10}	2^{14}	2^{18}	2^{10}	2^{14}	2^{18}	2^{10}	2^{14}	2^{18}
b		0.045	0.056	0.059	-0.092	-0.039	-0.031	-0.033	-0.024	-0.023	-0.011	0.005	0.007
s		0.168	0.040	0.013	0.176	0.048	0.015	0.074	0.016	0.004	0.046	0.013	0.004
mse		0.174	0.069	0.060	0.199	0.062	0.034	0.081	0.029	0.024	0.047	0.014	0.008

Table 4.6: **Sample size and estimation performance.** Estimation performance for fBm (first and second parts) and mrw (third and fourth parts) for wavelet coefficient (first and third parts) and Leader (second and fourth parts) based estimation procedures and sample sizes $N \in \{2^{10}, 2^{14}, 2^{18}\}$. Shown are bias b (first line), standard deviation s (second line) and (root) mean squared error mse (third line) for weighted (w^1) linear regressions from $j_1 = 3$ to $j_2 \in \{7, 11, 15\}$.

creasing sample size, for both d_X and L_X, and for both mono- and multifractal processes. For $\zeta(2)$ for d_X and fBm, this is consistent with the fact that the coefficient based estimator of H of H-sssi processes is asymptotically unbiased [4, 169]. These results thus suggest that this is as well true for L_X and for multifractal processes. We note, however, that for d_X, the decrease in bias is less significant for multifractal mrw, and also that it is less significant for L_X than for d_X.

Standard deviation and sample size. The results indicate that for both d_X and L_X, and for both mono- and multifractal processes, standard deviation decreases with increasing sample size N and is in a first approximation roughly proportional to $\sqrt{N}$.
Looking more into detail, standard deviations decrease overall more slowly for multifractal mrw than for fBm, indicating stronger correlation and hence smaller effective sample size increase with increasing N, for both d_X and L_X.

Mean squared error and sample size. The mse decreases systematically with increasing sample size for all estimates, for both d_X and L_X, and for both processes. This

decrease is in general more slowly for wavelet Leader based estimation than for coefficient based one: Whereas for $N \in \{2^{10}, 2^{14}\}$ and for both processes, L_X is more efficient than d_X for estimation of $\zeta(2)$ and c_1, this is the converse for $N = 2^{18}$, for which d_X display smaller mse. For the estimation of c_2, L_X perform consistently better for both processes and regardless of sample size N, however with decreasing performance gain over d_X when N is growing larger.
Finally, results and conclusions for different regression ranges not reported here for space reasons, notably for $j_1 = 2$, are similar and consistent with what is reported here for $j_1 = 3$, the main difference being that mse gains for Leader based estimation of $\zeta(2)$ and c_1 for sample size $N \in \{2^{10}, 2^{14}\}$ are smaller for $j_1 = 2$ than for $j_1 = 3$.

Conclusions. The results indicate that performance gain with increasing sample size is less pronounced for L_X than for d_X: Whereas for medium and small sample size, L_X outperforms d_X consistently, for large sample size, estimation of the self-similarity attributes $\zeta(2)$ or c_1 is more relevant for d_X. Also, although estimation of the multifractal attribute c_2 remains significantly better for all sample sizes reported here when based on wavelet Leaders, the difference in performance between d_X and L_X decreases with increasing N. One could hence imagine (very) large sample sizes (and specific process parameter setups) for which wavelet coefficients and Leaders would eventually display equivalent performance for the estimation of c_2. However, note that the largest sample size, $N = 2^{18}$, considered here is rarely available in practice, where signals are typically shorter[2]. Still, a more detailed account for the influence of sample size on estimation performance would demand for exploration of more different and larger sample sizes N

4.1.7 Process parameters and statistical performance

It has been observed [4, 169] that the standard deviation of wavelet coefficient and $\zeta(2)$ based estimation of H for (Gaussian) H-sssi processes does practically not depend on the precise value of H (although it has been shown theoretically that there is a very small dependence, cf. [58, 59]). This is indeed a strong and desirable property, since it highly facilitates the design of statistical procedures, and has been exploited for instance for the design of statistical tests in [170]. For wavelet Leader based estimation, this property has not yet been analyzed for H-sssi processes. Neither are there studies for multifractal processes or other estimates $\zeta(q)$ or c_p, for both Leader and coefficient based estimation. This section aims at numerically investigating these issues.
To this end, we study the standard deviations Eq. (4.2) of wavelet coefficient and Leader based estimation of $\theta \in \{\zeta(q), c_1, c_2, c_3\}$ as a function of process parameters λ ($\lambda = H$ for fBm and ROS, $\lambda = \{c_1, c_2\}$ for mrw). Let us denote these standard deviations by $s_{\theta,fBm}(H)$, $s_{\theta,ROS}(H)$, $s_{\theta,mrw}(c_1)$ and $s_{\theta,mrw}(c_2)$ respectively, and use the superscripts ^{d} and ^{L} to distinguish between wavelet coefficients and Leaders.
As observed above in Sections 4.1.3 and 4.1.6, the standard deviations s_θ depend both on the precise estimate θ and on the sample size N. We expect, however, that varying N does not change the *nature* of the function $s_\theta(\lambda)$, but results only in a change of scale. What is more, we will see below that $s_\theta(\lambda)$ is of the same nature for all $\zeta(q)$ and for c_1. Therefore, in order to be able to compare the nature of the functions $s_{\theta,\cdot}(\lambda)$ for different sample sizes N and different estimates θ, we define the following rescaled

[2] One example for which very large real-world samples are available is turbulence, where state-of-the-art measurements reach 2^{24} samples.

standard deviations:

$$\tilde{s}_{\theta,\cdot}(\lambda) = s_{\theta,\cdot}(\lambda) / \frac{1}{N_\lambda} \sum_{i=1}^{N_\lambda} s_{\theta,\cdot}(\lambda_i), \tag{4.5}$$

where N_λ is the number of values taken by the parameter and numerically investigated.

4.1.7-a) H-sssi processes: fBm and ROS

Fig. 4.3 summarizes rescaled standard deviations $\tilde{s}_\theta$ for fBm (Subfigure (a)) and ROS (Subfigure (b)). Shown are $\tilde{s}_\theta$ for wavelet coefficient (top rows, blue solid lines) and Leader (bottom rows, red solid lines) based estimation of $\theta = c_2$ and $\theta = c_3$ (center and right columns, respectively), and superpositions of $\tilde{s}_\theta$ for estimates $\theta \in \{\zeta(q), c_1\}$ ($q^d \in \{1,2,3\}$, $q^L \in \{-3,-2,-1,1,2,3\}$) (left column). In addition, $\tilde{s}_\theta$ are superposed for different sample sizes ($N \in \{2^{10}, 2^{14}, 2^{18}\}$ for fBm, $N \in \{2^{10}, 2^{14}\}$ for ROS).
The superpositions show that the nature of $\tilde{s}_\theta$ does indeed not depend on sample size: $\tilde{s}_\theta$ for different sample sizes N, as shown in each of the subplots, superpose perfectly well. What is more, $\tilde{s}_\theta$ is of the same nature for $\theta = \zeta(q)$ for all q and for c_1: Superpositions of $\tilde{s}_\theta$ for these estimates (left column) display the same behavior with H, for both fBm and ROS. In contrast, the curves $\tilde{s}_{c_2}$ and $\tilde{s}_{c_3}$ are different.
A key observation lies in the fundamental difference between $\tilde{s}^d_\theta(H)$ and $\tilde{s}^L_\theta(H)$ as functions of H: Heuristically, the $\tilde{s}^d_\theta(H)$ vary only weakly with H (they are "flat"), whereas the $\tilde{s}^L_\theta(H)$ have significant slopes and hence systematic variation with H, for both processes and any estimate θ. In order to further explore its nature, we propose the following model for $\tilde{s}_\theta(H)$:

$$f(\lambda) = \alpha \cdot \lambda^\beta, \tag{4.6}$$

with $\lambda = H$. This arbitrary model is used to heuristically evaluate the order of the dependence on λ. The parameter α is a proportionality factor between $s_\theta(\lambda)$ and λ^β and thus depends both on sample size and θ and is of no further particular interest here, whereas the parameter β quantifies the (non linear) dependence of $s_\theta(\lambda)$ on λ. Notice that β is invariant under rescaling $C \cdot s_\theta(\lambda)$. The model Eq. (4.6) is least-square fitted to the raw standard deviations $s_\theta(\lambda)$. It has been verified to show good fit for both fBm, ROS (and mrw), and it is observed that the fit parameter β does not vary significantly with sample size N. What is more, β is consistently equal for $\theta = \zeta(q)$ for any q and for c_1. Hence, this confirms the statements on the nature of dependence on sample size and parameters from above. We translate these fits to the rescaled standard deviations $\tilde{s}_\theta(\lambda)$ by using for β the mean of the β from fits to the raw standard deviations s_θ. The parameter α is chosen in analogy with the definition of $\tilde{s}_\theta$ in Eq. (4.5) by setting $1/N_\lambda \sum_{i=1}^{N_\lambda} f(\lambda_i) = 1$.

The resulting models are shown as thick black lines in Fig. 4.3. For fBm, $\beta^d_{\theta,fBm} \approx 0$ is close to zero for all estimates θ, hence confirming that $s^d_{\theta,fBm}(H)$ is practically independent of H. In contrast, $\beta^L_{\theta,fBm}$ is up to two orders of magnitude larger than $\beta^d_{\theta,fBm}$. This indicates much stronger dependence of estimation performance on H for L_X than for d_X.
However, $\beta^d_{\theta,ROS} \gg \beta^d_{\theta,fBm}$, hence dependence of standard deviation on H for d_X is much stronger for ROS than for fBm. This is in particular the case for estimates of $\zeta(q)$ and c_1. Similarly, $\beta^L_{\theta,ROS} > \beta^L_{\theta,fBm}$, but increase of dependence on H for ROS with respect to fBm is less pronounced for L_X than for d_X. Nevertheless, $\beta^L_{\theta,ROS} > \beta^d_{\theta,ROS}$ consistently, and significantly so for c_2 and c_3.

Conclusions. The results illustrate that wavelet Leader based estimation performance for H-sssi processes displays in general a strong dependence on H, be the process Gaussian or not, and regardless of the precise estimate θ. Also, wavelet coefficient based estimation is practically independent of H for Gaussian H-sssi processes only. For non Gaussian H-sssi processes, $\tilde{s}^d_\theta(H)$ is a nonlinear function of H for estimation of $\zeta(q)$ and c_1. We suggest that this dependence on H is due to dependence between wavelet coefficients d^2_X: Results presented in Chapter 11 indicate indeed stronger residual dependence of coefficients for non Gaussian H-sssi than for fBm.

4.1.7-b) Multifractal process: mrw

Fig. 4.4 shows superpositions of rescaled standard deviations $\tilde{s}_{\theta,mrw}(c_1)$ (Subfigure (a)) and $\tilde{s}_{\theta,mrw}(c_2)$ (Subfigure (b)) for mrw and sample sizes $N \in \{2^{10}, 2^{14}, 2^{18}\}$. Shown are $\tilde{s}_{\theta,mrw}$ for wavelet coefficient (top rows, blue solid lines) and Leader (bottom rows, red solid lines) based estimation of $\theta = c_2$ and $\theta = c_3$ (center and right columns, respectively), and superpositions of $\tilde{s}_\theta$ for estimates $\theta \in \{\zeta(q), c_1\}$ ($q^d \in \{1, 2, 3\}$, $q^L \in \{-3, -2, -1, 1, 2, 3\}$) (left column). The solid black lines correspond to models Eq. (4.6) with $\lambda = |c_p|$. As before, we obtained β by least square fits to raw standard deviations $s_{\theta,mrw}(c_1)$ and $s_{\theta,mrw}(c_2)$. The results indicate that $s^d_{\theta,mrw}(c_1)$ does only weakly depend on c_1 for $\theta = \zeta(q)$ and $\theta = c_1$, and that it is practically independent of c_1 for $\theta = c_2$ and $\theta = c_3$. In contrast, the $s^L_{\theta,mrw}(c_1)$ show more significant dependence on c_1 for any parameter θ. In comparison with $s^L_{\theta,fBm}(H)$ and $s^L_{\theta,ROS}(H)$ for H-sssi processes, the influence of c_1 on wavelet Leader based estimation performance for mrw is, however, moderate, with $\beta^L_{\theta,mrw} \approx 0.25$ as compared to $\beta^L_{\theta,\cdot} \approx 0.6 - 2.2$ for fBm and ROS. Also, $s^d_{\theta,mrw}(c_2)$ is only weakly dependent on c_2 for $\theta = \zeta(q)$ and $\theta = c_1$, and practically independent of c_2 for $\theta = c_2$ and $\theta = c_3$. In contrast, $s^L_{\theta,mrw}(c_2)$ display more significant dependence on c_2, moderately so for $\theta = \zeta(q)$ and $\theta = c_1$, and very strongly for $\theta = c_2$ and $\theta = c_3$.

Conclusions. We conclude that, first, $s^L_{\theta,mrw}$ display stronger dependence on c_1 and c_2 than $s^d_{\theta,mrw}$. Second, whereas $s^d_{c_2,mrw}$ and $s^d_{c_3,mrw}$ are practically independent of the precise values of c_1 and c_2, c_1 and c_2 have strong (non linear) influence on $s^L_{c_2,mrw}$ and $s^L_{c_3,mrw}$. Hence, in this perspective, one could conclude that statistical procedures or tests based on c_2 or c_3 should use wavelet coefficient based estimates. However, as discussed above, it is precisely for these estimates that d_X display poor performance, whereas wavelet Leader based estimates are excellent.

4.1.7-c) Conclusions

Overall, the results reported above suggest that the independence of process parameters of coefficient based estimation performance can not in general be validated: It is in fact reminiscent only to Gaussian H-sssi processes. In this particular case, wavelet coefficients may therefore substantially ease the design of statistical procedures as compared to wavelet Leaders, which show a strong non linear dependence on H. In the general case, the dependence of performance on process parameters makes the design of such procedures intricate, for both d_X and L_X, demanding for either precise models or estimates of the dependence, which are unavailable at present. These facts justify the use of nonparametric methods such as bootstrap, as proposed in Chapter 7.

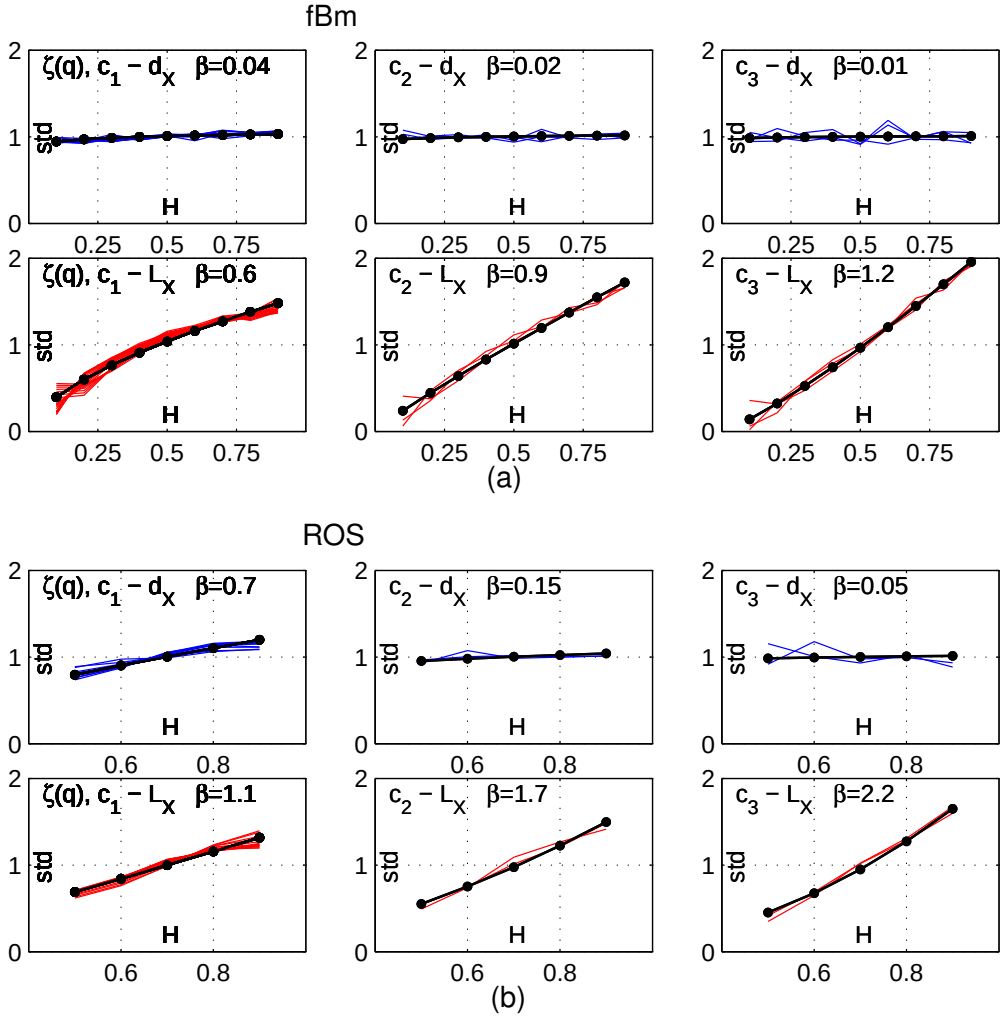


Figure 4.3: H **of fBm and ROS and estimation performance.** Rescaled standard deviations $\tilde{s}_\theta(H)$ for fBm (a) and ROS (b) for d_X (first rows) and L_X (second rows). Each subplot shows superpositions of $\tilde{s}_\theta(H)$ for sample sizes $N \in \{2^{10}, 2^{14}, 2^{18}\}$ (fBm) and $N \in \{2^{10}, 2^{14}\}$ (ROS), for $\theta = c_2$ (center columns), $\theta = c_3$ (right columns) and $\theta \in \{\zeta(q), c_1\}$ ($q^d \in \{1, 2, 3\}$, $q^L \in \{-3, -2, -1, 1, 2, 3\}$) (left columns), and the fitted model Eq. (4.6) (black line, '•'). Estimations of θ are obtained using $j_1 = 3$, $j_2 = 7$, and weighted regressions w^1.

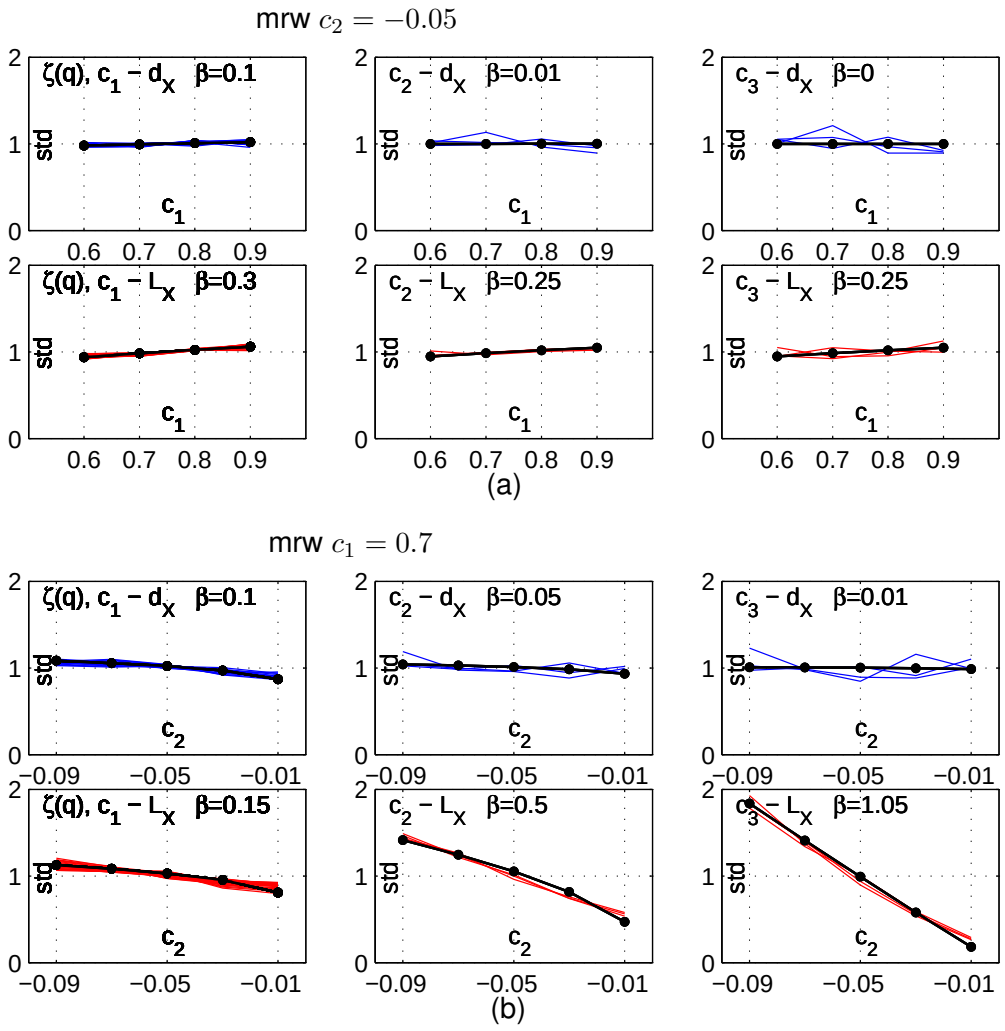


Figure 4.4: c_1, c_2 **of mrw and estimation performance.** Rescaled standard deviations $\tilde{s}_{\theta,mrw}(c_1)$ (a) and $\tilde{s}_{\theta,mrw}(c_2)$ (b) for d_X (first rows) and L_X (second rows). Each subplot shows superpositions of $\tilde{s}_\theta(H)$ for sample sizes $N \in \{2^{10}, 2^{14}, 2^{18}\}$ for $\theta = c_2$ (center columns), $\theta = c_3$ (right columns) and $\theta \in \{\zeta(q), c_1\}$ ($q^d \in \{1, 2, 3\}$, $q^L \in \{-3, -2, -1, 1, 2, 3\}$) (left columns), and the fitted model Eq. (4.6) (black line, '•'). Estimations of θ are obtained using $j_1 = 3$, $j_2 = 7$, and weighted regressions w^1.

4.1.8 Gaussianity of estimates

The fact that wavelet coefficients are Gaussian for fBm (**P4** in Section 2.3.4, cf. e.g. [4]) highly facilitates their statistical analysis for this specific process. This property has inspired detailed analytical work on the statistical properties of the estimation procedures in this setting, and has been used in the design of statistical inference procedures (e.g. [4, 70, 169, 170]). In Section 3.2, we reported results that suggest that wavelet Leaders of fBm do not inherit this comfortable property from wavelet coefficients, and that their distributions are non Gaussian[3]. This constitutes one of the reasons why analytical results on statistical properties of wavelet Leaders are difficult to obtain, even in the case of Gaussian fBm. Also, as soon as one leaves the special case of fBm, there is no reason for wavelet coefficients – and even less for wavelet Leaders – to be Gaussian: Fig. 4.2, for instance, clearly indicates that for the non Gaussian H-sssi Rosenblatt process, the empirical distributions of both coefficients and Leaders are highly non Gaussian, with stronger non Gaussianity for Leaders than for coefficients. This is in general as well the case for multifractal processes and makes – together with potentially complicated dependence structure of wavelet coefficients and Leaders – the analytical study of estimation procedures difficult.

Independently of these issues, it is, in practice, important to have knowledge on in how far the distributions of the *estimates* θ are Gaussian or not: Indeed, many (asymptotical) statistical tests and procedures are based on Gaussian theory. In the context of multifractal analysis, estimates are obtained by linear regressions over a range of scales that increases as sample size N increases. Hence, we expect the distributions of the estimates to approach a Normal limiting distribution as N grows large if structure functions behave nicely and are sufficiently independent at different scales. To investigate this issue, Fig. 4.5 summarizes quantile-quantile plots of the empirical distributions of the wavelet coefficient and Leader based estimates of $\zeta(2)$ and c_2 (centered and normalized to unit variance) against the standard Normal distribution for fBm (Subfigure (a)) and mrw (Subfigure (b)), and for sample sizes $N = 2^{10}$ (top rows) and $N = 2^{18}$ (bottom rows). Estimates are obtained using weighted (w^1) linear regressions with $j_1 = 3$ and $j_2 \in \{7, 15\}$, respectively.

Fig. 4.5 (a) shows that for fBm, the distributions of $\hat{\zeta}(2)$ are Gaussian for both small and large N, for both wavelet coefficients and Leaders. In contrast, the distributions of both wavelet coefficient and Leader based estimates of c_2 display some degree of non Gaussianity (in kurtosis for d_X and skewness for L_X) for small sample size, and approach Gaussian only for the large sample size.

For mrw, Fig. 4.5 (b) suggests that conclusions are similar to fBm: The distributions of $\hat{\zeta}(2)$ are close to Gaussian already for the small sample size, and distributions of $\hat{c}_2$ are non Gaussian for small N and approach Gaussian for large N, for both coefficient and Leader based estimation.

Conclusions. For the sample sizes and process parameters considered here, there is no clear evidence that estimates based on wavelet coefficients have distributions significantly closer to Gaussian than wavelet Leader based ones, although the distributions of the wavelet coefficients themselves are in general closer to Gaussian than those of wavelet Leaders.

[3] Clearly, the $\sup$ of Gaussian r.v.s is not Gaussian.

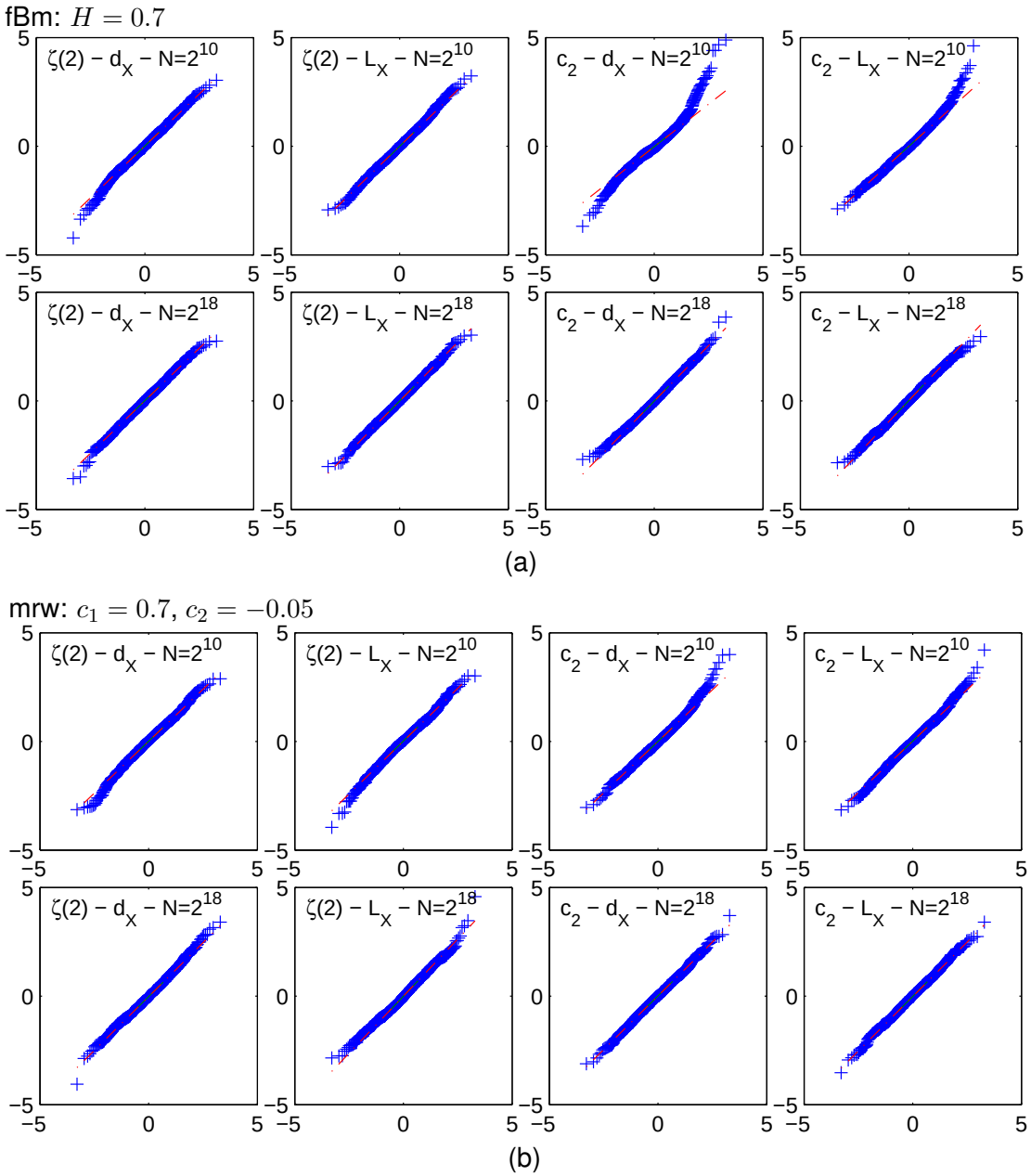


Figure 4.5: **Gaussianity of estimates.** Quantile-quantile plots against standard normal of Monte Carlo distributions of estimates of $\zeta(2)$ (first (d_X) and second (L_X) column) and c_2 (third (d_X) and fourth (L_X) column) for fBm (a) and mrw (b) and sample size $N = 2^{10}$ (first rows) and $N = 2^{18}$ (second rows), respectively.

4.2 Statistical Performance of Estimation Procedures: Images

4.2.1 Numerical simulations

The goal of this section is to practically validate the 2d multifractal analysis based on practical implementations of 2d wavelet Leaders Eq. (2.55), and to compare the performance to that obtained for wavelet coefficient based estimation. To this end, we apply the respective estimation procedures Eq. (2.65), Eq. (2.71) and Eqs. (2.74–2.75) to a large number N_{MC} of realizations of synthetic stochastic 2d processes of size $N \times N$ with a priori known and controlled multifractal properties.

4.2.1-a) Synthetic multifractal processes

The synthetic reference processes chosen for numerical validation of 2d multifractal analysis estimation procedures are fractional Brownian motion (fBm) and canonical Mandelbrot cascades with log-Normal multipliers (CMC-LN) and with log-Poisson multipliers (CMC-LP) (cf. Sections 2.2.3 and 2.7.1-a), respectively). They provide us with simple yet representative examples of a Gaussian monofractal process, and of log-Normal ($c_3 = 0$) and non log-Normal ($c_3 \neq 0$) multifractal processes, respectively. The CMC-LN and CMC-LP multifractal reference processes are obtained by fractionally integrating directly the cascades $Q_r(t)$ by an order $\alpha = 0.5$ to ensure minimum regularity. This is a necessary prerequisite for the validity of the wavelet Leader multifractal formalism (WLMF), since $Q_r(t)$ of CMC are not uniform Hölder (cf. Definition 2.16). A more precise theoretical and practical account for the minimum regularity conditions assumed to hold for the images will be given in Section 5.1.

4.2.1-b) Simulation setup

Parameters for numerical simulations are set to $N_{MC} = 500$ and $N = 1024$. All results are obtained with Daubechies' wavelets with $N_\psi = 2$ vanishing moments. It has been checked that using wavelets with larger N_ψ yields identical results and conclusions. Linear regressions are performed over the scales $2^3 \leq 2^j \leq 2^7$ with weights w^1, as suggested by Section 4.1.4 for 1d signals. The process parameters are fixed to $H = 0.7$ for fBm, and for CMC such that $(c_1, c_2, c_3) = (0.513, -0.025, 0)$ (for CMC-LN, i.e. $m = 0.0125$)) and $(c_1, c_2, c_3) = (0.538, -0.080, 0.014)$ (for CMC-LP, i.e. $\beta = 0.8395$, $\gamma = 0.4195$).

4.2.2 Structure functions

Fig. 4.6 shows, for fBm (left) and CMC-LP (right), means over Monte Carlo realizations of structure functions $\log_2 S(2^j, q)$ for $q = 2$ (top row) and for $q = -2$ (second row), and the deviations of the structure functions from their theoretical slope, $\log_2 S(2^j, q) - j \cdot \zeta(q)$, for $q = 2$ (third row) and $q = -2$ (bottom row). The 95% asymptotic confidence limits are obtained by Monte Carlo simulation. For wavelet coefficients, only positive qs are shown, as structure functions diverge for negative qs.

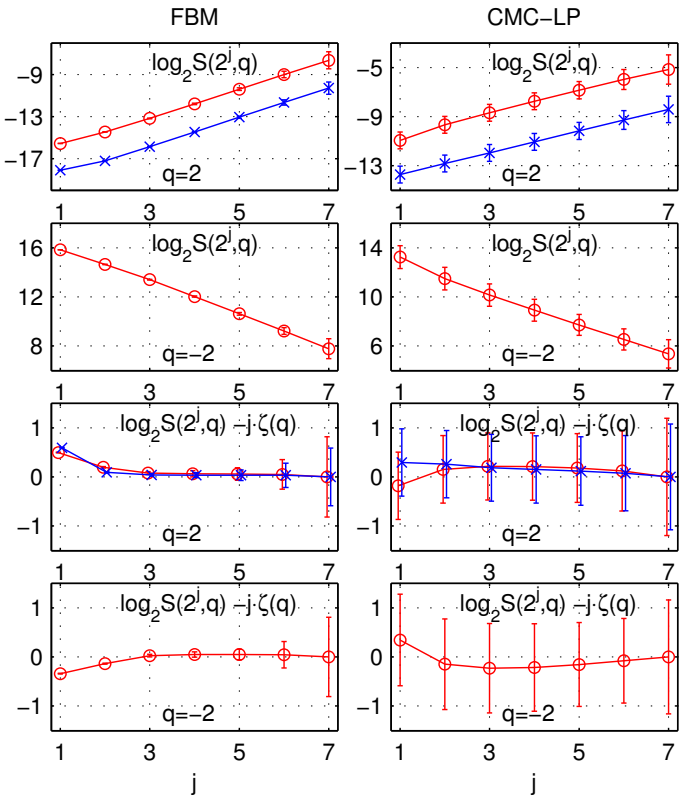


Figure 4.6: **Structure functions for 2d synthetic processes.** Structure functions $\log_2 S(2^j, q)$ for $q = 2$ (top row) and $q = -2$ (second row), and deviation of structure functions from theoretical slope, $\log_2 S(2^j, q) - j \cdot \zeta(q)$, for $q = 2$ (third row) and $q = -2$ (bottom row), obtained by mean over Monte Carlo realizations for fBm (left) and CMC-LP (right); 'o' correspond to wavelet Leader, '$\times$' to wavelet coefficient based estimations. The error bars correspond to 1.96 Monte Carlo standard deviation.

4.2.2-a) Scaling range

A first investigation (top and second row) suggests that for both wavelet coefficients and Leaders, structure functions display scaling behavior as in Eqs. (2.7) and (2.58) over the range of scales $2^2 \leq 2^j \leq 2^7$. A closer look at the deviations from the theoretical slope (third and last row) confirms this observation for wavelet coefficients. However, it reveals that for wavelet Leaders, $\log_2 S(2^j, q)$ becomes a linear function of j only for $2^j \geq 2^3$. As for 1d signals, this can be interpreted through the fact that, theoretically, a Leader is defined as the $\sup$ of coefficients at all finer scales down to infinitely fine scales, whereas practically, the $\sup$ can be taken only down to the finest available, first scale. Hence, in practice, wavelet Leaders need fine scales for initialization, whereas wavelet coefficients do not (cf. Section 3.1).

4.2.2-b) Projection step

The non-scaling behavior of $S(2^j, q)$ of coefficients at the first scale $j = 1$ is due to the fact that the pre-filtering or projection step theoretically necessary for a clean wavelet analysis has been omitted (cf. [3, 172]).

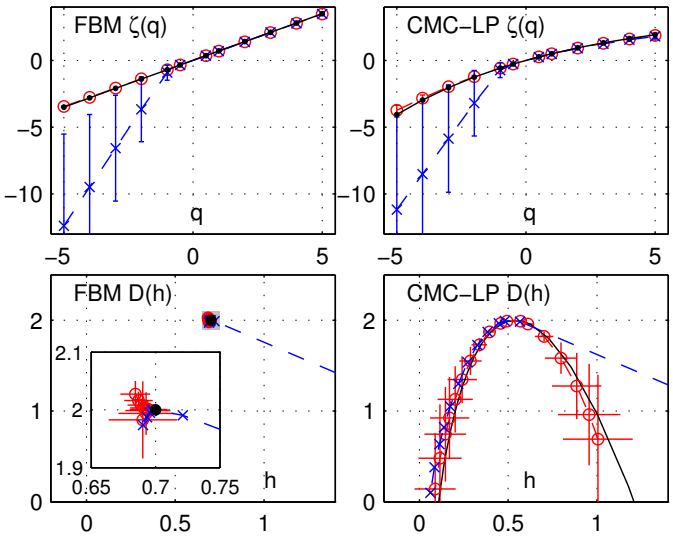


Figure 4.7: **Estimates $\zeta(q)$ and $D(h)$ for 2d synthetic processes.** Estimates $\zeta(q)$ (top) and $D(h)$ (bottom) for fBm (left) and CMC-LP (right), obtained from wavelet Leaders ('$\circ$') and coefficient ('$\times$'), by mean over Monte Carlo simulation, and theoretical attributes (solid black, '$\bullet$'). The error bars correspond to 1 Monte Carlo standard deviation. The error bars for coefficient based estimates for $D(h)$ are not shown for better visibility of the spectra. The insert for $D(h)$ of fBm corresponds to the area shaded around the theoretical spectra location.

4.2.2-c) Regressions

We observe further that the error bars for structure functions for CMC-LP are substantially larger than those for fBm, suggesting a smaller variability of the $S(2^j, q)$ and better subsequent estimation performance for the latter, monofractal process. Moreover, it is interesting to note that whereas the error bars for fBm behave approximately as $1/\sqrt{n_j}$, this is not the case for CMC-LP, where the size of the confidence intervals varies only slightly with j. This confirms that the choice of weights for weighted linear fits, as proposed in [179], is appropriate for fBm. Though not optimal for multiplicative cascades, it has been shown to perform better than non-weighted regression for 1d signals (cf. Section 4.1). Additional results not reported here indicate that this remains valid for 2d images.

4.2.3 Performance of parameter estimation

We quantify the performance of the estimators $\hat{\theta} \in \{\hat{\zeta}(q), \hat{c}_p\}$ by their *bias* (b):

$$\hat{b}_\theta = \widehat{\mathbb{E}}_{MC}\,\hat{\theta} - \theta, \tag{4.7}$$

standard deviation (s):

$$\hat{s}_\theta = \widehat{\mathrm{Std}}_{MC}\,\hat{\theta} = \sqrt{\widehat{\mathbb{E}}_{MC}\,\hat{\theta}^2 - \left(\widehat{\mathbb{E}}_{MC}\,\hat{\theta}\right)^2} \tag{4.8}$$

and *(root) mean squared error* (mse):

$$\widehat{mse}_\theta = \sqrt{\hat{b}_\theta^2 + \hat{s}_\theta^2}, \tag{4.9}$$

where $\widehat{\mathbb{E}}_{MC}$ stands for mean over Monte Carlo realizations.

Estimation performance results are summarized in Tab. 4.7 and illustrated in Fig. 4.7, for both wavelet coefficients and Leaders, for medium size images ($N = 1024$). In addition, Tab. 4.7 shows complementary results for large images of CMC-LP ($N = 2048$). Overall, we observe that Leader based estimations are more efficient than coefficient based ones.

4.2.3-a) Positive statistical moments q

For $q > 0$, both d_X and L_X have approximately equal mean squared error and standard deviation. Wavelet coefficient based estimations have slightly better performance for monofractal fBm and weakly ($c_2 = -0.025$ being relatively small) multifractal CMC-LN while wavelet Leader based ones have better performance for multifractal CMC-LP.

4.2.3-b) Negative statistical moments q

For $q \leq -1$, the $\hat{\zeta}^d(q)$ are not meaningful as they exhibit very large mse. Therefore, a wavelet coefficient based multifractal analysis of images allows to explore exclusively the range $q > -1$ and thus, in practice, only the increasing part of the multifractal spectrum $D(h)$. In contrast, the wavelet Leader based procedure permits a complete analysis of the multifractal properties of an image, and notably of the decreasing part of the multifractal spectrum (cf. Fig. 4.7).

4.2.3-c) Log-cumulants

Tab. 4.7 shows that the $\hat{c}_p^L$ exhibit consistently smaller standard deviation and mean squared error than the $\hat{c}_p^d$. Whereas the difference in performance is only small for the estimation of c_1, it becomes more significant for c_2 and c_3, with gains in mse of up to more than one order of magnitude for L_X. This is of crucial importance, since non-zero c_2 and c_3 discriminate fv H-sssi from MMC processes. Furthermore, for $N = 2048$, Tab. 4.7 (last line) enables to deduce that an asymptotic 85% confidence interval for the Leader based estimation of c_3 excludes zero, hence that the real c_3 is different from zero with high probability. To the best of our knowledge, this is the first time that the statistically significant estimation of non-zero c_3 has been achieved on multifractal images. It also clearly shows that for images of smaller size, the estimation of the parameter c_3, a fortiori of higher order c_p, should be used with care: Confidence intervals might be so large that they may not exclude 0, even if the parameters are non zero.

4.2.3-d) Self-similar vs. multiplicative cascade processes

For both the wavelet Leader and coefficient based procedures, estimation is more difficult for multifractal CMC-LN and CMC-LP than for fBm, resulting in larger mse for estimations on the multifractal process.

4.2.3-e) Dependence on process parameters

Results for (2d) images reported here are complementary to those obtained for 1d signals in Section 4.1.7, and are obtained by the same procedure detailed therein. Hence, we

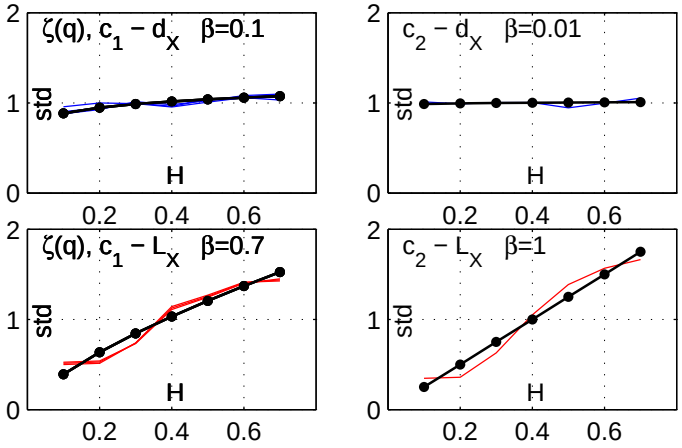


Figure 4.8: H **of 2d fBm and estimation performance.** Rescaled standard deviations $\tilde{s}_\theta(H)$ for fBm for d_X (first row) and L_X (second row) for $\theta = c_2$ (right column), and $\theta \in \{\zeta(q), c_1\}$ ($q^d \in \{1,2,3\}$, $q^L \in \{-3,-2,-1,1,2,3\}$) (left columns), and the fitted model Eq. (4.6) (black line, '•'). Estimations of θ are obtained using $j_1 = 3$, $j_2 = 7$, and weighted regressions w^1 ($N = 1024$). A clear dependence is visible for wavelet Leaders, and less pronounced for wavelet coefficients.

compare standard deviations through the (arbitrary) model Eq. (4.6):

$$f(\lambda) = \alpha \cdot \lambda^\beta,$$

where λ is the parameter as a function of which the dependence is explored.
Figure 4.8 confirms and extends the results obtained for 1d signals: For fBm, the standard deviations of wavelet coefficient based estimates of $\zeta(q), q \geq 0$ or c_p do not (or only very weakly) depend on H. In contrast, for wavelet Leader based estimation, there is a clear and strong dependence on the parameter H.
Figure 4.9 illustrates the corresponding results for standard deviations for estimation of $\zeta(q)$, c_1 and c_2 for CMC-LN, as a function of the parameters c_1 (top) and c_2 (bottom). Conclusions are essentially similar: The dependence on c_1 and c_2 of the standard deviation of $\hat{\zeta}^d(q)$, $\hat{c}_1^d$ et $\hat{c}_2^d$ is weak, whereas it is much more pronounced for wavelet Leader based estimation.

4.2.4 Discussion

Numerical simulations, equivalent to those reported above, have been performed on other multifractal processes and lead to similar conclusions: Wavelet coefficients do not allow to meaningfully explore negative qs and thus to measure the decreasing part of the spectrum $D(h)$ whereas wavelet Leaders do; Estimation of c_p is better when based on Leaders than on coefficients, and significantly so for c_p, $p \geq 2$.

4.2.4-a) Selection of scaling range

We have seen in Sections 3.1 and 4.2.2 that the range of scales over which the linear fits are to be performed are likely to be narrower (requiring the use of a larger j_1) for wavelet Leaders than for wavelet coefficients. For practical multifractal analysis and real-world

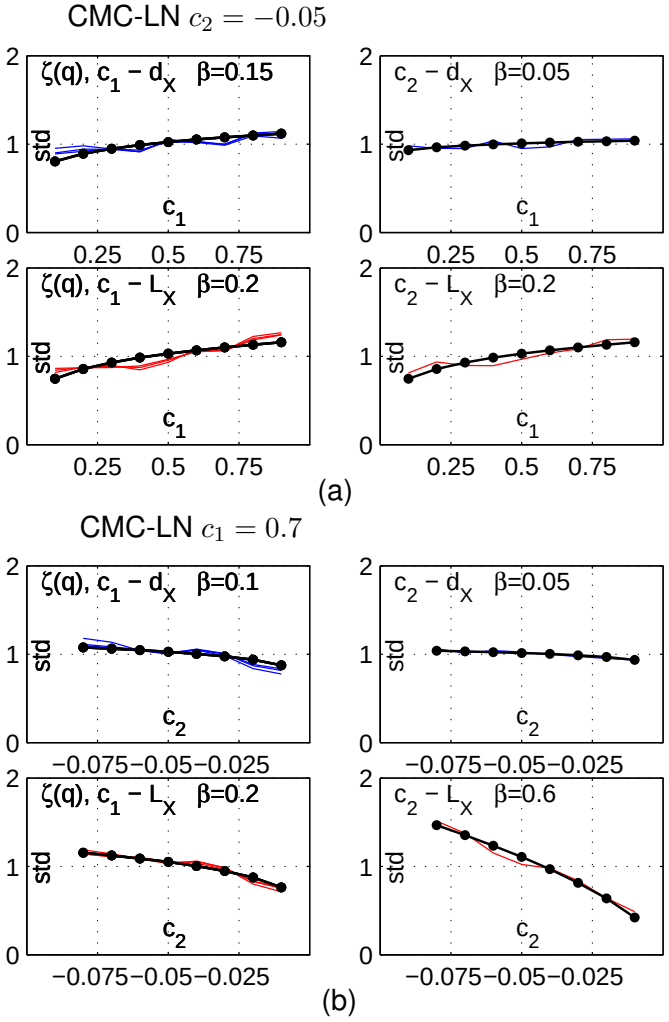


Figure 4.9: c_1, c_2 **of 2d CMC-LN and estimation performance.** Rescaled standard deviations $\tilde{s}_{\theta,CMC-LN}(c_1)$ (a) and $\tilde{s}_{\theta,CMC-LN}(c_2)$ (b) for d_X (first rows) and L_X (second rows), for $\theta = c_2$ (right column) and $\theta \in \{\zeta(q), c_1\}$ ($q^d \in \{1,2,3\}$, $q^L \in \{-3,-2,-1,1,2,3\}$) (left column), and the fitted model Eq. (4.6) (black line, '•'). Estimations of θ are obtained using $j_1 = 3$, $j_2 = 7$, and weighted regressions w^1 ($N = 1024$).

images, the choice of this regression range is a crucial, difficult and controversial issue, further complicated by the use of Leaders.

4.2.4-b) Vanishing moments of the wavelet

Another critical practical issue is the choice of the number of vanishing moments N_ψ of the wavelet with which the data are analyzed. Condition $N_\psi > h$, where h is the largest singularity exponent present in the data, is expected to be sufficient for a relevant multifractal analysis (cf. Eq. (2.56)). However, in practice, the choice of N_ψ results from a trade-off: A larger N_ψ stabilizes the estimates of the negative q structure functions and enables to get rid of potentially superimposed smooth trends such as polynomial, hence improves estimation and brings robustness; A larger N_ψ also implies a larger time support for the wavelet and thus produces border effects of wider size, such that no coefficients

may remain unaffected at large scales, hence in itself degrades estimation performance. Therefore, a reasonable practical rule of thumb is to choose the smallest N_ψ for which the estimated multifractal attributes do not significantly change when N_ψ is increased (i.e, remain within confidence intervals, emphasizing again the need for and the importance of such confidence intervals).

4.2.4-c) Real-world data and multifractal analysis

In Section 2.5, we recalled how the estimates of $\zeta(q)$, c_p, $D(q)$ and $h(q)$ are theoretically intimately tied to the multifractal spectrum $D(h)$ of the field X and hence to the analysis of its local regularity fluctuations. This relation is known to hold theoretically for all the synthetic images used here for numerical simulations and illustrations. For real-world images, this interpretation in terms of multifractal spectrum and singularity description might not always be completely relevant. However, this, in no way, prevents practitioners to make use of the measured $\zeta(q)$, c_p, $D(q)$ and $h(q)$ to analyze the data in terms of a less mathematically stringent formulation of scale invariance property, or to perform standard image processing tasks such as classification or retrieval based on such quantities.

4.2.4-d) Analysis of real-world images and computational costs

The performance of the wavelet Leader based estimation procedures are illustrated on a real-world image of size 1024×768, taken by the author of this manuscript with a standard digital camera. The image and estimation results are shown in Fig. 8.2 in Chapter 8, together with confidence intervals obtained with the method proposed therein. Results demonstrate that the procedure can be readily applied to real-world images for their complete multifractal characterization. The procedures have also been applied to a large database of real-world texture images in a classification task. This is reported in Chapter 13 and further demonstrates the practical relevance of the 2d WLMF.
The 2d WLMF is simple both conceptually (2d DWT, Leaders and linear regressions) and practically (very low computational cost, with respect to both time and memory, of the order of a 2d DWT). The entire estimation procedure with wavelet Leaders takes less then a minute on a standard PC, whereas the same estimation with 2d MMWT (cf. Section 2.5.5) increases computation time by a factor larger than 20.

4.2.5 Conclusions

The present section shows that the 2d Wavelet Leader Multifractal Formalism outperforms significantly the previous propositions based on 2d DWT in estimation performance, and 2d MMWT in computation time, memory cost and implementation complexity. It is fast and efficient since simply based on a 2d orthogonal DWT. The procedure is backed up by a strong mathematical framework and enables – thanks to the use of wavelet Leaders – an accurate and complete characterization of the (ir)regularities of the texture of an image, under the minimum regularity conditions assumed to hold for the images.

fBm - $N = 1024$							
		DWT			LWT		
θ	theo	b	s	mse	b	s	mse
$\zeta(-2)$	-1.400	-2.061	2.315	3.100	0.012	0.022	0.025
$\zeta(-1)$	-0.700	-0.195	0.544	0.578	0.005	0.011	0.012
$\zeta(1)$	0.700	-0.001	0.010	0.010	-0.004	0.012	0.013
$\zeta(2)$	1.400	-0.002	0.021	0.021	-0.007	0.025	0.026
c_1	0.700	-0.001	0.012	0.012	-0.004	0.011	0.012
c_2	0.000	0.000	0.029	0.029	0.001	0.003	0.003
c_3	0.000	0.001	0.150	0.150	-0.001	0.001	0.001
CMC LN - $N = 1024$							
		DWT			LWT		
θ	theo	b	s	mse	b	s	mse
$\zeta(-2)$	-1.075	-2.081	2.514	3.264	0.036	0.035	0.050
$\zeta(-1)$	-0.525	-0.169	0.576	0.600	0.018	0.017	0.025
$\zeta(1)$	0.500	-0.017	0.014	0.023	-0.020	0.018	0.027
$\zeta(2)$	0.975	-0.035	0.032	0.047	-0.042	0.032	0.059
c_1	0.512	-0.018	0.015	0.024	-0.018	0.017	0.026
c_2	-0.025	0.003	0.031	0.031	-0.001	0.007	0.007
c_3	0.000	0.006	0.152	0.152	-0.000	0.004	0.004
CMC LP - $N = 1024$							
		DWT			LWT		
θ	theo	b	s	mse	b	s	mse
$\zeta(-2)$	-1.256	-1.859	2.507	3.121	0.029	0.064	0.070
$\zeta(-1)$	-0.580	-0.138	0.546	0.563	0.010	0.025	0.027
$\zeta(1)$	0.500	-0.016	0.017	0.023	-0.007	0.021	0.022
$\zeta(2)$	0.933	-0.031	0.036	0.047	-0.013	0.045	0.046
c_1	0.538	-0.019	0.018	0.027	-0.008	0.022	0.023
c_2	-0.080	$+0.007$	0.032	0.033	0.003	0.015	0.015
c_3	0.014	-0.009	0.152	0.152	-0.003	0.013	0.013
CMC LP - $N = 2048$							
		DWT			LWT		
θ	theo	b	s	mse	b	s	mse
c_1	0.538	-0.017	0.010	0.019	-0.006	0.012	0.013
c_2	-0.080	0.007	0.016	0.018	0.002	0.009	0.009
c_3	0.014	-0.002	0.075	0.075	-0.002	0.008	0.009

Table 4.7: **Estimation performance for 2d synthetic processes.** Bias b, standard deviation s and root mean squared error mse of wavelet coefficient (center columns) and Leader (columns on the right) based estimation for 2d fBm (top), CMC LN (center top) and CMC LP (center bottom) for $N = 1024$, and for CMC LP (bottom) for $N = 2048$. The columns on the left identify the parameters and their theoretical values.

4.3 Conclusions and Perspectives

In this chapter, numerical simulation results for the wavelet coefficient and Leader based estimation of multifractal attributes for both (Gaussian and non-Gaussian) monofractal and different multifractal processes, and for both 1d and 2d data, have been presented and discussed. The aims were manifold: First, analyzing and comparing statistical performance of wavelet coefficient and Leader based estimation for various processes and multifractal attributes; Second, disentangling the impact of estimation parameters (regression range and weights) and sample size on estimation performance; Third, quantifying the influence of process nature (Gaussian vs. non Gaussian H-sssi, mono- versus multifractal processes) and parameters on estimation performance, and fourth, characterizing their impact on the distributions of the estimates (standard deviation and departure from Gaussian).
The contributions of this chapter lie, on one hand, in the systematic empirical statistical performance analysis for wavelet coefficient and wavelet Leader based multifractal analysis, and, on other hand, in the validation of the 2d WLMF. These numerical results illustrate effects that have never been clearly stated elsewhere before and provide practitioners with performance references and clear practical guidelines for the multifractal analysis of signals and images in applications.
Let us state the key findings and the main conclusions of this chapter.
2d WLMF. The proposed 2d WLMF is practically operational and enables a relevant multifractal characterization of images with reasonable computational cost. It constitutes the first and only procedure that is based on a solid mathematical ground and enables to practically perform a multifractal analysis of images. The performance is significantly better than those of DWT or CWT coefficient based methods formerly proposed in the literature.
H-sssi and MMC processes. For the estimation of the self-similarity parameter in H-sssi data, wavelet coefficient and Leader based procedures have comparable performance. Hence, if it can be a priori ensured that the process under analysis is H-sssi, wavelet coefficients are to be preferred over wavelet Leaders, which add extra (implementational) complexity and practical difficulties (selection of regression range).
For performing multifractal analysis in any other situation, however, wavelet Leader based procedures are to be clearly preferred over wavelet coefficient based ones, since wavelet Leader based procedures achieve significant performance gains over wavelet coefficient based ones for the estimation of attributes characterizing the departure from monofractal, and in particular so for $c_p, p \geq 2$. This constitutes an original and important result that has, to our knowledge, never been clearly stated elsewhere before and opens the way towards the design of efficient statistical tests (cf. Chapters 9 and 10). Also, estimates involving negative statistical orders q can only be obtained with Leaders, as has been reported by Lashermes et al. (e.g. [107]). The results presented in this chapter demonstrate in addition that such estimates display performance comparable to those for positive statistical orders. Consequently, only wavelet Leader based estimation can be used to precisely characterize the multifractal nature of data, to discriminate monofractal from multifractal models, or to discriminate among different multifractal models. Hence, Leader based estimations have to be consulted as soon as it is not known a priori that the data are H-sssi.
Also, estimation performance is significantly decreased for multifractal processes as opposed to Gaussian monofractal processes. The results heuristically indicate that the

further from monofractal the processes (i.e., the larger the values of $|c_p|$ for $p \geq 2$), the more difficult the estimation of multifractal attributes. Similarly, multifractal attribute estimation is more difficult for non Gaussian than for Gaussian H-sssi processes.

Gaussianity of estimates and dependence on process parameters. The distributions of estimates based on wavelet Leaders were not found to be significantly further from Gaussian than those based on wavelet coefficients for the sample sizes considered here. It is conjectured that for smaller sample size, they may be. This point requires further numerical or analytical investigations.

The standard deviation for both wavelet coefficient and Leader based estimation does in general depend on the precise value of the process parameters in a non linear manner. The only exception is the particular case of wavelet coefficient based estimation of H for fBm, for which standard deviations are practically independent of H. The dependence of standard deviations is in general significantly larger for wavelet Leaders than for wavelet coefficients. This difference – the variance of the estimates for wavelet coefficient based estimation depending weakly on the quantities to be estimated, those for wavelet Leader based estimation significantly – constitutes an original and important result that has, to the best of our knowledge, never been reported before, be it for 1d signals or for 2d images. At a first glance, this comes to the disadvantage of wavelet Leaders, since the dependence complicates confidence interval and test constructions for the estimates. The difficulty is, however, put in perspective by the other results reported in this section: Statistical performance of wavelet Leader based estimation is usually significantly better than for wavelet coefficient based estimation. This dependence also justifies the use of techniques such as bootstrap, as proposed in Chapter 7.

This statistical disadvantage is – together with the practically more difficult validation of ranges of scales over which scale invariance occurs, hence determination of scales to be involved in linear regressions – the main price to pay for the use of Leaders. The issue demands for further numerical and analytical analysis, in particular in the perspective of models that could be exploitable for the design of statistical procedures and tests.

Overall, the results lead us to the conclusion that due to their significantly better multifractal attribute estimation performance, wavelet Leaders are to be practically preferred for the multifractal analysis of signals and images. Wavelet coefficient based estimates, which can be obtained as a side product without noteworthy extra computational effort, should nonetheless be consulted in a complementary way for the validation of scaling in the data and for the selection of the scaling range, and for the estimation of $\zeta(2)$ and c_1.

Chapter 5

Measures, Fractional Integration, Linearization Effect and Quantization Impacts

Contents

In this chapter, we are interested in three theoretical questions of major importance for practical multifractal analysis through the multifractal formalism developed in Section 2.5.

First, we account for one limitation of the wavelet Leader multifractal formalism that is due to the fact that it can only be meaningfully applied to data that belong to the *bounded function* class. This constitutes indeed a severe practical limitation, in particular for the analysis of images, which can be naturally interpreted as approximations of positive measures and do hence not fulfill this requirement. We theoretically state the problem and propose methods for practically evaluating to which function class data belong, and whether they fulfill the bounded function requirement. Based on these results, we elaborate on a modification of the WLMF (proposed earlier for the MMMF [16, 17, 57] and involving a pseudo-fractional integration) that can be applied to data that do not fulfill this requirement, and discuss its properties and limitations. This is addressed in Section 5.1. The proposed methods as well as the theoretical characterization of the problem are original and are seen as a major contribution to practical multifractal analysis. This work has been reported in [9, 91, 90, 178, 182].

The goal of Section 5.2 is to study the *linearization effect*, a phenomenon in multifractal analysis that determines the order of statistical moments that can be involved in the practical multifractal analysis of data. Though of fundamental practical importance, the origins of this phenomenon are still not well understood theoretically. Based on empirical studies, we obtain original results that indicate that the linearization effect can be explained through the combination of arguments involving extreme values, heavy tailed marginal distributions and the dependence structure of multifractal processes. We see these results as an important contribution and as an empirical basis for theoretically linking the linearization effect to the multifractal properties of the data. This work has been reported in [11].

In an important number of applications, notably those involving image data, the data under analysis can be subject to (severe) quantization in amplitude. Though of crucial importance in practice, the characterization of the impact of such quantizations on practical multifractal analysis, and in particular of the respective robustness of different multiresolution quantities against quantization effects, has, to our knowledge, never been accounted for in the literature. This is addressed in Section 5.3, and has been reported in [181].

5.1 Measures and Fractional Integration

The wavelet Leader multifractal formalism is well-defined for bounded functions only (cf. Section 2.5). However, a digital image can naturally be seen as the approximation, at a given resolution level, of a positive measure. There is hence a priori no guarantee that images fall into the class of bounded functions, and it is indeed commonly reported in the literature dedicated to empirical multifractal analysis that images are not in the class of bounded functions [134]. Answering the question whether an image is a bounded function or not is hence a mandatory prerequisite for practical application of the wavelet Leader multifractal analysis. Therefore, the first goal of the present section is to propose a method that enables to practically decide whether this condition is fulfilled or not. This is described in Section 5.1.2. To our knowledge, this is the first time that such a method is obtained.

Along the same line, there is a priori no guarantee that actual digital images should belong

to any $L^q(\mathbb{R}^2)$ space, $q \geq 1$, nor to the space of bounded variation (BV) functions (i.e., its gradient is a bounded measure). Many image processing algorithms are, nonetheless, based on models that assume that the image belongs to one or combinations of such function spaces. Thus, the second goal of this section is to propose a procedure for practically evaluating to which function space a given real-world image belongs. This is presented in Section 5.1.1.
The fact that the WLMF can only be applied to images within the bounded function class is rather restrictive in practice. Therefore, the third goal of this section is to address the bounded function requirement. We propose a solution – consisting of a modified WLMF that is based on heuristics reported in [16, 17, 57] – which enables a characterization of images not belonging to the class of bounded functions. The characterization is deeply related to the intrinsic regularity properties of these images. At the same time, the approach can be used to account for the function space issue evoked above. It makes use of a pseudo-fractional integration in the wavelet domain and can be obtained without noteworthy additional computational effort. The solution is detailed in Section 5.1.3. It has, to our knowledge, never been obtained and reported before for the WLMF.
Finally, we discuss the theoretical and practical implications of such approaches with respect to the multifractal characterization of images. These have never been clearly stated elsewhere before.
The work presented in this section has been obtained in collaboration with Stéphane Jaffard (Univerité Paris XII) and published in [9, 90, 91, 178, 182].

5.1.1 Classes of images

Let $S^d(2^j, q)$ denote the wavelet coefficient based structure functions (Eq. (2.62)):

$$S^d(2^j, q) = \frac{1}{3n_j} \sum_{m=1}^{3} \sum_{k_1,k_2} |d_X^{(m)}(j, k_1, k_2)|^q,$$

and let us define, for $q > 0$,

$$\zeta^d(q) = \liminf_{2^j \to 0} \ln S^d(2^j, q) / \ln 2^j. \tag{5.1}$$

Then, $\zeta^d(q) > 0$, $q \geq 1$, indicates that the image X is in $L^q(\mathbb{R}^2)$ [182]. Hence, estimates $\hat{\zeta}^d(q)$ (cf. Eq. (2.65)) contain important information on the precise function space to which the image belongs: For instance, with $q = 1$, $\zeta^d(1) > 0$ implies that X is in $L^1(\mathbb{R}^2)$. In addition, when $\zeta^d(1) > 1$, this indicates that X belongs to the class of bounded variations images. Conversely, when X is a measure, this necessarily yields $\zeta^d(1) \geq 0$. For $q = 2$, $\zeta^d(2) > 0$ implies that X is in $L^2(\mathbb{R}^2)$ (e.g. [178, 182]).
Therefore, estimates of $\zeta^d(q)$ help practitioners to decide whether or not the image X can be embedded in models underlying many image processing algorithms, such as the Osher-Rudin-Fatemi model (or any of its declinations) [125, 149]. Indeed, all these models rely on the underlying assumption that the image considered belongs to a certain function space (or a sum of two functions spaces).
Examination of large sets of real world images reveals that a small proportion of images are characterized with $\zeta^d(2) < 0$, with confidence intervals (computed with the method proposed in Chapters 7 and 8) clearly validating the negativity for the estimate. Also, $\zeta^d(1)$ is positive for most images (as expected for positive measures), but $0 < \zeta^d(1) < 1$

for a large proportion of images, which are hence not within the bounded variation class. These statements are further supported by the examination of a large image database, consisting of 1000 texture images and described in Chapter 13. Results obtained for these 1000 images are presented in Tab. 5.2 and Fig. 5.2 and clearly confirm the conclusions drawn from the examination of other real-world images.
For illustration purposes, we also note that the images in Fig. 5.1 are characterized by: $\zeta^d(1) = 0.03 \pm 0.02$ and $\zeta^d(2) = 0.05 \pm 0.04$ (Fig. 5.1, left), $\zeta^d(1) = -0.21 \pm 0.03$ and $\zeta^d(2) = -0.49 \pm 0.04$ (Fig. 5.1, center), and $\zeta^d(1) = -0.08 \pm 0.03$ and $\zeta^d(2) = -0.17 \pm 0.08$ (Fig. 5.1, right).

Estimation performance of $\hat{\zeta}^d(q)$ have been largely addressed in Chapter 4 and will not be further discussed here. In Section 5.1.3, we will propose a practical procedure for increasing $\zeta^d(q)$ such that $\zeta^d(1) > 1$ and/or $\zeta^d(2) > 0$.

5.1.2 Uniform regularity

As noted in the introduction, there is no guarantee that real-world images should fall into the class of bounded functions, which would permit a straightforward application of the wavelet Leader multifractal formalism. A digital image consists of local averages (for each pixel) of light intensity, and can hence be naturally seen as the approximation, at a given resolution level, of a positive measure, and not of a bounded function. We recall that a (two-dimensional) measure μ is a linear form on $\mathcal{C}(\mathbb{R}^2)$: $\exists C > 0$, such that $\forall f$ with compact support (see e.g. [178, 182]):

$$|\langle \mu | f \rangle| \leq C \sup |f|.$$

The wavelet coefficients of μ satisfy:

$$|d_\mu^{(m)}(j, \mathbf{k})| = 2^{-2j} |\langle \mu | \psi_{j,\mathbf{k}}^{(m)} \rangle| \leq C 2^{-2j},$$

and can hence grow exponentially when $j \to -\infty$. Then, the wavelet Leaders can be infinite, and the WLMF can not be directly applied. We note that this difficulty is not specific to images, but also appears for signals (1d) that can be seen as discretized versions of measures (such as dissipation signals in turbulence [121], and rainfall precipitation in meteorology [155]). Also, all multifractal formalisms proposed at present suffer from this same problem and can not be applied to unbounded measures.
The difficulty can be related to the existence of *negative Hölder exponents* in the measures. The precise definition of negative Hölder exponents requires theoretical developments beyond the scope of this section (this is addressed in [178, 182]). We only recall here that the underlying singularity heuristic yet remains the same: $|X(\mathbf{t}) - X(\mathbf{t_0})| \sim_{|\mathbf{t}-\mathbf{t_0}| \to 0} |\mathbf{t} - \mathbf{t_0}|^h$. The presence of negative Hölder exponents in data has been reported in, for instance, [134].
WLMF and unbounded functions. In practice, any image – since obtained through a practical acquisition system and hence necessarily finite valued – will yield empirical wavelet Leaders with finite value, be the bounded function condition satisfied or not. For functions that do not belong to the bounded function class, they do, however, not have any theoretical significance and take on arbitrary values, since the signature of negative Hölder exponents persists. Therefore, when the wavelet Leader based multifractal formalism is applied to images that do not respect the bounded function requirement, it

		CMC-LN	CMC-LP
h_{min}		-0.304	-0.391
$\hat{h}_{min}$	mean	-0.342	-0.395
	std	0.126	0.117
	bias	-0.038	-0.004
	mse	0.132	0.117

Table 5.1: **Estimation performance for h_{min}.** Mean estimation performance of $\hat{h}_{min}$ (Eq. (5.2)), obtained for 500 realizations of CMC-LN (left column) and CMC-LP (right column) of size $N \times N = 1024 \times 1024$. The first line gives the theoretical value for h_{min}, the second, third, fourth and last line summarize the mean, standard deviation (std), bias and (root) mean squared error (mse) of the estimate, respectively.

produces results that are meaningless. Moreover, practitioners have little or no means to decide a posteriori whether wavelet Leaders are meaningful or not and this needs to be checked a priori.

A sufficient condition for the WLMF to be well-defined is that the image is *uniform Hölder* (cf. Def. 2.16) and hence that the minimum regularity exponent $h_{min} > 0$. The notation h_{min} (cf. Eq. (2.53)) is justified by the fact that $h_{min} > 0$ for X implies for the Hölder exponents $h_X(\mathbf{t}_0)$ of X that $\forall \mathbf{t}_0 : h_X(\mathbf{t}_0) \geq h_{min}$. If $X(\mathbf{t})$ is uniformly Hölder, it is a continuous function, hence bounded, hence possesses finite wavelet Leaders in the limit of fine scales.

Estimation of h_{min}. The exponent h_{min} can practically be obtained through the quantities [91, 178, 182]:

$$d_X^M(j) = \sup_{m,k_1,k_2} |d_X^{(m)}(j,k_1,k_2)|,$$

consisting of the largest wavelet coefficient at the scale j. Then, h_{min} is given by:

$$h_{min} = \liminf_{2^j \to 0} \frac{\ln d_X^M(j)}{\ln 2^j}.$$

This is a direct consequence of Eqs. (2.52) and (2.53) in the definition of uniform regularity. Practically, this amounts to measuring linear slopes in log-log coordinates, yielding the estimate:

$$\hat{h}_{min} = \sum_j w_j \log_2 d_X^M(j), \tag{5.2}$$

where w_j are the linear regression weights as defined in Section 2.6.

Performance evaluation. To evaluate the performance of the estimator $\hat{h}_{min}$, we apply it to a large number N_{MC} of 2d CMC-LN and CMC-LP cascades Q_r, for which the theoretical h_{min} is known and controlled a priori. The process parameters are fixed as in Section 4.2.1-b). Estimation parameters are set to $N_{MC} = 500$, $N \times N = 1024 \times 1024$, Daubechies wavelets with $N_\psi = 2$, and ordinary (w^0) regression involving the scales $j \in [1, 3]$. Results are reported in Tab. 5.1. They indicate that the estimator Eq. (5.2) has satisfactory statistical performance to practically enable determination of whether for a real-world image h_{min} is close to zero or negative.

Real-world images and h_{min}. Examination of real-world images shows that they often exhibit negative h_{min}. This is illustrated in Fig. 5.1 for three example images, taken by

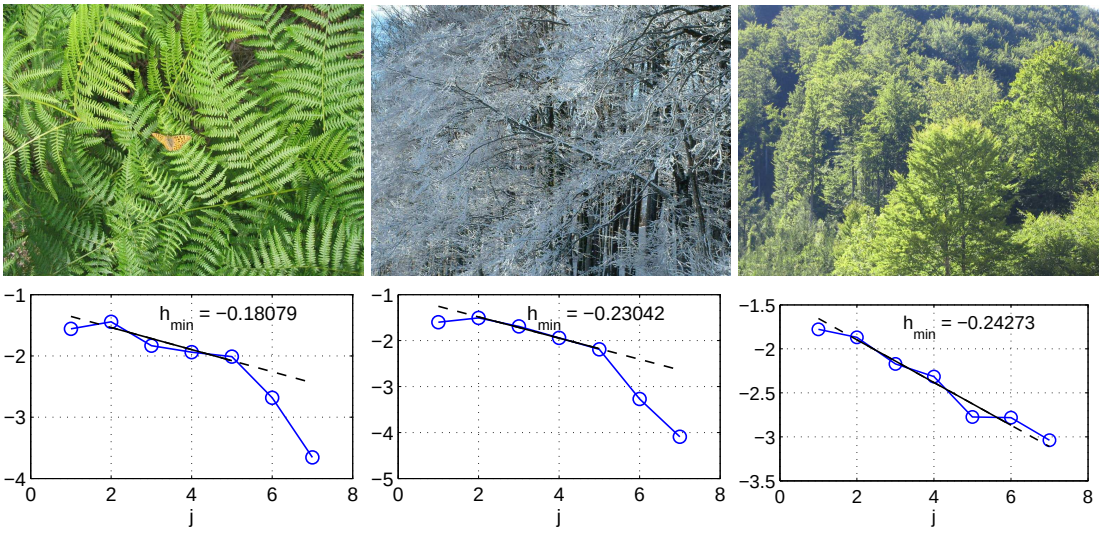


Figure 5.1: **Estimation of h_{min} for real-world images.** Estimation of h_{min} for (gray level versions of) images of fern leaves (left), trees in winter (center) and a forest in summer (right). Shown are the images (top, taken by ourselves with a standard digital camera), the coefficients $\log_2 d_X^M(j)$ as a function of scale j (bottom, blue solid and '$\circ$'), linear fits (bottom, black solid and dashed lines) and the estimated values for $\hat{h}_{min}$.

	median	std	< 0	< 1
$\hat{\zeta}^d(1)$	0.61	0.29	5.5%	96.5%
$\hat{\zeta}^d(2)$	0.95	0.59	9.2%	-
$\hat{h}_{min}$	-0.03	0.27	53.4%	-

Table 5.2: **Function space and uniform regularity.** Mean estimates of $\hat{\zeta}^d(q=1)$ (first line), $\hat{\zeta}^d(q=2)$ (second line) and $\hat{h}_{min}$ (bottom line) for the 1000 gray level texture images in the database in [185] (cf. Chapter 13): Median (left column) and standard deviations (second column) of estimates, fraction (in %) of estimates smaller than 0 and smaller than 1 (third and fourth column, respectively).

ourselves with a standard digital camera. The estimates of h_{min} for their gray level versions clearly indicate that these three images are not uniform Hölder.
A systematic examination of the texture image database described in Chapter 13 further reveals that negative h_{min} do actually occur systematically in real-world images (cf. Tab. 5.2 and Fig. 5.2): For the 1000 textures images in this database, the minimum regularity exponents h_{min} are distributed around the zero value, and a large fraction of the images has negative h_{min}. The results suggest that more than one half of them are not uniform Hölder.
This calls for a modification of the wavelet Leader based multifractal formalism proposed in Section 2.5 for bounded functions. This is the subject of the next section.

5.1.3 Fractional and pseudo-fractional integration

One way to overcome the negative h_{min} issue and to return to the study of bounded (actually uniform Hölder) functions, to which the WLMF formalism described in Section

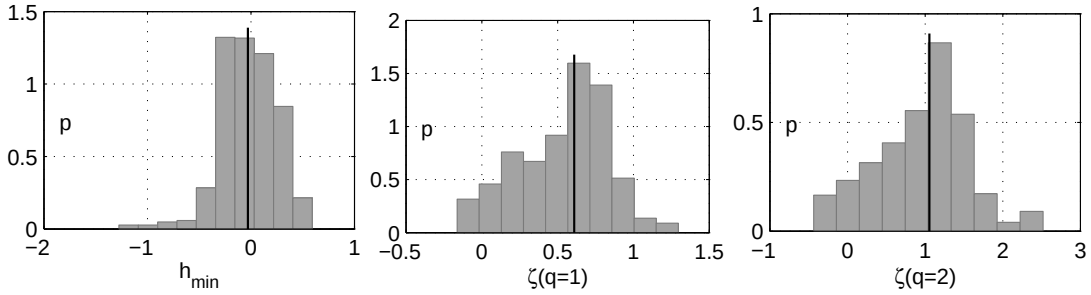


Figure 5.2: **Function space and uniform regularity evaluation: Histograms.** Histograms of $\hat{h}_{min}$ (left), $\hat{\zeta}^d(q=1)$ (center), and $\hat{\zeta}^d(q=2)$ (right), obtained for the 1000 gray level texture images in the database in [185] (cf. Chapter 13). The black vertical line corresponds to the median of the estimates.

2.5 can be applied, is to fractionally integrate the image with a sufficient order. This has been abundantly used for instance in [16, 17, 57] for the MMMF (cf. Section 2.5.5).

The fractional integration (of order α) of a function or measure X is defined in the Fourier domain as:

$$(\widehat{I^\alpha X})(\xi) = (1+|\xi|^2)^{\alpha/2}\hat{X}(\xi). \tag{5.3}$$

If X is a positive measure on $\mathbb{R}^2$, fractional integration of order $\alpha > 2$ is always sufficient for $I^\alpha X$ to be uniform Hölder. In practice, a smaller order can be sufficient [90, 178, 182]:

Proposition 5.1 *If X is a measure with $h_{min} < 0$, then $I^\alpha X$ is uniform Hölder as soon as $\alpha > -h_{min}$.*

5.1.3-a) Pseudo-fractional integration and multifractal formalism

From a practical point of view, the fractional integration of an image can be difficult and disadvantageous in applications: It demands non negligible computational effort and is likely to introduce numerical errors due to border effects. Instead of actually computing the fractionally integrated version of X and then applying the wavelet Leader multifractal formalism to $I^\alpha X$, we propose a modified version of the WLMF that combines both operations into a single one, as follows.

i) First compute the 2d wavelet coefficients $d_X^{(m)}(j,k_1,k_2)$ (Eq. 2.35) and replace them with:

$$d_X^{(m),\alpha}(j,k_1,k_2) = 2^{\alpha j} d_X^{(m)}(j,k_1,k_2).$$

This amounts to computing the wavelet coefficients of $\tilde{I}^\alpha X$, a pseudo-fractionally integrated version of X, whose local and global regularity properties are identical to that of $I^\alpha X$ as soon as $\alpha > -h_{min}$.

ii) Then, calculate the wavelet Leaders of these new wavelet coefficients:

$$L_X^\alpha(j,k_1,k_2) = \sup_{m,\,\lambda' \subset 3\lambda_{j,k_1,k_2}} |d_X^{(m),\alpha}(\lambda')|. \tag{5.4}$$

Such modified wavelet Leaders L_X^α are equivalent to wavelet Leaders $L_{I^\alpha X}$ computed from $I^\alpha X$, in the sense that if $I^\alpha X$ has Hölder exponent h at $\mathbf{t}_0$ then, when $2^{-j}\mathbf{k}$ is the

closest dyadic point of $\mathbf{t}_0$, $L_X^\alpha(j,k_1,k_2) \sim_{2^j \to 0} 2^{jh}$ (see Eq. 2.56).

iii) Finally, compute new structure functions (cf. Eq. (2.57)):

$$S_\alpha^L(j,q) = \frac{1}{n_j} \sum_{k_1,k_2} L_X^\alpha(j,k_1,k_2)^q \tag{5.5}$$

They behave as power law w.r.t. the analyzing scale 2^j, in the limit of fine scales $2^j \to 0$:

$$S_\alpha^L(j,q) \approx G_q^\alpha 2^{j\zeta_\alpha(q)}. \tag{5.6}$$

The spectrum of $I^\alpha X$ is therefore given by:

$$D_\alpha(h) = \min_{q \neq 0}(1 + qh - \zeta_\alpha(q)). \tag{5.7}$$

5.1.3-b) Spectrum of the original image

The method we proposed consists in fixing $\alpha > -h_{min}$ and then applying the WLMF to the image $\tilde{I}^\alpha X$, which is uniform Hölder, to obtain the spectrum $D_\alpha(h)$. Therefore, it is natural to ask how $D_\alpha(h)$ can be interpreted, and what the link is between $D_\alpha(h)$ and the original image X.

Commonly, the spectrum $D_\alpha(h)$ of $I^\alpha X$, for a fixed and a priori chosen $\alpha > -h_{min}$, is used as a characterization of a translated version of the multifractal spectrum of the image X. In other words, it amounts to associating with the original image X a spectrum $D(h)$ via:

$$D(h) = D_\alpha(h+\alpha), \quad \alpha > -h_{min}. \tag{5.8}$$

It is important to be aware of the fact that this heuristic faces two limitations. First, a function $D(h)$ is not systematically obtained by this procedure since translations $D_\alpha(h+\alpha)$ for different values of $\alpha > -h_{min}$ might not superimpose and can hence not define a function $D(h)$. We note that in this case, the original image X contains *oscillating singularities* (cf. Eq. (2.47)). The discrepancy between translations $D_\alpha(h+\alpha)$ for different values of $\alpha > -h_{min}$ can be theoretically related to the oscillation parameter β in Eq. (2.47). A precise analysis of these issues is beyond the scope of this section and can be found in [90, 91, 178]. Nevertheless, we note that the modified WLMF therefore potentially opens perspectives for the practical detection and measurement of oscillating singularities in real-world data by comparing translated versions $D_\alpha(h+\alpha)$ for different values of $\alpha > -h_{min}$. This demands for further theoretical and empirical investigations and will not be discussed here.

Second, the function $D(h)$ can not in general be related to the multifractal properties of X. Let us suppose that the multifractal formalism holds for $I^\alpha X$ when $\alpha > -h_{min}$, and that it contains only cusp-type singularities (cf. Eq. (2.46)) and no oscillating singularities. Then, the multifractal spectra of $I^\alpha X$ and $I^\beta X$, $\alpha, \beta > -h_{min}$, can be obtained from each other by a simple translation of $\alpha - \beta$, and the same is true for their Legendre spectra. Therefore, one can naturally associate a Legendre spectrum $D(h)$ with them through Eq. (5.8). However, in contrary to the above heuristic, it can not in general be interpreted as the spectrum of X when $h_{min} < 0$, since X is not uniform Hölder and therefore, its multifractal spectrum (Def. 2.11) is not defined.

Despite such restrictions, the function $D(h)$ obtained by Eq. (5.8) for large enough values of $\alpha > -h_{min}$, is deeply related to the intrinsic properties the original image X and can be

used as such. In particular, it corresponds exactly to the spectrum of the original image X if X is uniform Hölder and does not contain oscillating singularities.

Finally, we note that fractional integration also increases the $\zeta^d(q)$ (the counterpart of the heuristic translation in Eq. 5.8 reads $\zeta^d_\alpha(q) = \zeta^d(q) + \alpha q$), hence ensuring, for large enough α, that $\zeta^d_\alpha(1) > 1$ and/or $\zeta^d_\alpha(2) > 0$. Similarly, it increases H or c_1 by α and can hence reduce fine scale cutoff effects (cf. Section 3.1).

5.1.4 Conclusions and perspectives

In this section, we proposed a practical solution for determining which function class an image belongs to, and for verifying whether a real-world image actually fulfills the bounded function condition for meaningful application of the WLMF or not. We proposed a modification of the WLMF that can be applied to images for which the bounded function condition is not fulfilled. Combined together, these methods enable the practical multifractal characterization of images. We note that the modified WLMF opens new perspectives for the detection and measurement of oscillating singularities in real-world data, an important issue in applications that has, to our knowledge, never been addressed before. This has not been explored in this section and demands for further theoretical and empirical work[1].
The issues discussed in this section also show that although wavelet coefficients offer a restricted analysis only of the multifractal properties of an image (see Sections 2.5.5 and 4.1), a number of useful pieces of information can still be extracted from the wavelet coefficient based structure functions and from the largest wavelet coefficients at each scale. This leads to the conclusion that wavelet coefficients are providing preliminary information regarding the regularity properties of X and should hence be used before applying the wavelet Leader multifractal formalism, and in a complementary manner, rather than with the usual competition perspective.

[1] A first practical approach and first empirical results (with Stéphane G. Roux, ENS Lyon) are preliminary and not reported here.

5.2 Linearization Effect

5.2.1 Position of the problem

In practice, multifractal analysis essentially amounts to measuring a collection of scaling law exponents, which are then related to the multifractal spectrum via the multifractal formalism (cf. Sections 2.5.2 and 2.5.3). Practical multifractal analysis is based on the structure functions[2] and states that they behave as (cf. Eqs. (2.58) and (2.62)):

$$S_N(a,q) = \frac{1}{n_a}\sum_{k=1}^{n_a} |T_X(a,ak)|^q \simeq c_q a^{\zeta(q)}, \quad \text{as } a \to 0, \tag{5.9}$$

where, for the present section, $T_X(a,t) = \delta_{a\tau_0}X(t) = X(t+a\tau_0) - X(t)$ are the increments[3] of the data under analysis $X(t)$ at scale a, N is the sample size of $X(t)$ and $n_a = N/a$ the number of increment coefficients available at scale a. Estimation of the scaling exponents $\zeta(q)$ is commonly performed by linear fits in log-log coordinates as in Eq. (2.65).
It can be further shown that for a number of commonly used multifractal processes (notably within the class of MMC processes with stationary increments):

$$\mathbb{E}|T_X(a,t)|^q \simeq C_q a^{\lambda(q)}, \quad \text{as } a \to 0, \tag{5.10}$$

for $0 < q < q_c^+ = \sup\{q : E|T_X(a,t)|^q < \infty\}$, where:

$$\lambda(q) = q + \varphi(q). \tag{5.11}$$

The function $\varphi(q)$ is strongly related to the details of the multiplicative construction underlying the definitions of such processes (see Section 2.7, Eqs. (2.81), (2.102), (2.125), (2.127)).

Since sample averages Eq. (5.9) are naturally used as estimates for the ensemble averages, it has long and largely been believed in the applied multifractal literature that the functions $\zeta(q)$ and $\lambda(q)$ in Eq. (5.9) and Eq. (5.10) were identical, at least for $0 < q \leq q_c^+$, hence that the scaling exponents $\zeta(q)$ were related to the details of the definitions of the processes $X(t)$. However, after the seminal works of Molchan [127], Ossiander and Waymire [134] on Mandelbrot multiplicative cascades [121], it is now being realized that the two functions $\lambda(q)$ and $\zeta(q)$ coincide, surprisingly, only on the narrow range of powers $0 < q \leq q_*^+$ with $q_*^+ < q_c^+$ as defined in Eq. (2.129). Moreover, $\zeta(q)$ is known to behave as a linear function for $q > q_*^+$. This is referred to as the *linearization effect*. These observations have been confirmed in a comprehensive empirical study by Lashermes et al. [108] who conjectured that this phenomenon is intrinsic to all multifractal processes and measures and that the value of q_*^+ is given by Eq. (2.129). This confusing association has often been misleading in the use of scaling exponents for real-life data analysis.
In a number of significant contributions, whose most prominent are [123] (Chapter 9) and [122], Mandelbrot relates negative singularity observation and *supersampling* issues, intimately tied to the linearization effect, to the intrinsically heavy tail nature of multiplicative

[2] We prefer in this section to keep explicit reference to the dependence of quantities on the sample size N, denoted by the subscript $_N$.

[3] Here τ_0 stands for an (arbitrary) time unit which is, by consistency with the definition of wavelet coefficients, chosen as the inverse of the sampling frequency.

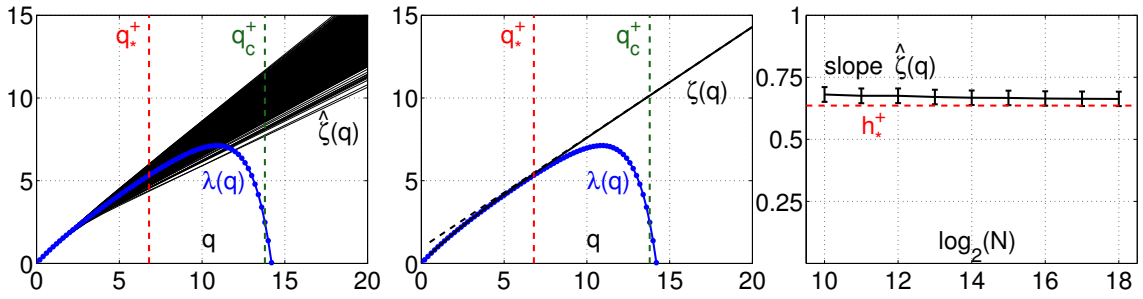


Figure 5.3: **Linearization effect.** $\hat{\zeta}(q)$ versus q, observed over 1000 realizations (left) and averaged over all realizations (middle), together with the curve $\lambda(q)$ (solid dotted curve), the dashed line expands the linear behavior observed at large q. Right: averaged slope, characterizing this linear behavior, as a function of $\log_2 N$. Observe that it does not depend on N and is found to be very close to $h_*^+ \simeq 0.64$ (red dashed line).

cascades. The present section aims at contributing to the analysis and a better understanding of the origins and causes of the differences in nature of these two different functions of q: $\lambda(q)$ and $\zeta(q)$. We propose to explain this effect through an argument involving extreme values, the intrinsic heavy tail nature of marginal distributions and the dependence structure of multifractal processes. The work in this section has been obtained in collaboration with Vladas Pipiras (University of North Carolina, Chapel Hill) and reported in [11].

5.2.2 Numerical simulations and linearization effect

We choose to consider here a particular class of multifractal processes, the compound Poisson motions (CPM) $A(t)$, chosen because their increments $\{T_A(a,t), t \in \mathbb{R}\}$ form stationary processes, for each analysis scale a, and since their multifractal properties are well known ([30], see Section 2.7.1-b)). Notably, the function $\lambda(q)$ in Eqs. (5.10) and (5.11) is given by Eqs. (2.108) and (2.102).

Numerical simulations. All numerical simulations reported below were conducted over $N_{MC} = 1000$ independent realizations of CPM, with various $\varphi(q)$ and various data lengths ($N = 2^{10}, \ldots, 2^{18}$), within a single integral scale. Plots and results are presented for the specific $\varphi(q)$ based on lognormal multipliers W (cf. Section 2.7.1-b), Eq. (2.112)), yielding numerically $q_c^+ \simeq 13.8$, $q_*^+ \simeq 6.8$ (obtained by Eq. (2.129)) and $h_*^+ \simeq 0.64$ (as computed from Eq. (5.14), see below). However, the results presented here are conjectured to hold for all choices of $\varphi(q)$.

Linearization effect. The estimation procedure Eq. (2.65) based on increments has been applied to N_{MC} realizations of CPM. First, we observe that, for each and every realization of CPM, $\hat{\zeta}(q)$ is close to $\lambda(q)$ at small qs, i.e., $0 \leq q \leq q_N$ while it behaves linearly in q, for large qs, i.e., $\hat{\zeta}(q) = \alpha_N + \beta_N q$, for $q \geq q_N$:

$$\hat{\zeta}(q) \simeq \begin{cases} \lambda(q), & \text{if } q \leq q_N, \\ \alpha_N + \beta_N q, & \text{if } q \geq q_N, \end{cases} \tag{5.12}$$

where α_N, β_N and q_N are r.v.s whose means are found not to depend on N [108]. This is illustrated in Fig. 5.3, left plot. Second, averaging over the N_{MC} realizations, we observe

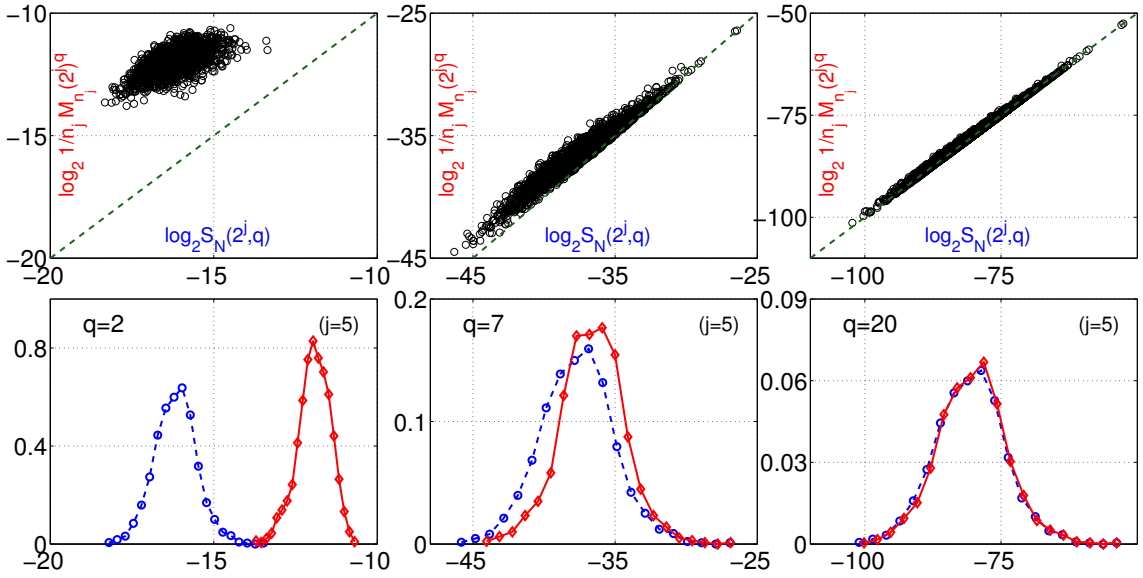


Figure 5.4: **Structure functions versus extrema.** Two sides of Relation (5.16): Scatter plots (top row) and empirical distribution functions (bottom row; '$\circ$': left side of Eq. (5.16); '$\diamond$': right side of Eq. (5.16)), for $j = 5$ with (left to right) $q = 2, 7(\simeq q_*^+), 20$.

(Fig. 5.3, middle plot) that[4] $\hat{\mathbb{E}}_{MC}\hat{\zeta}(q)$ is close to $\lambda(q)$ at small qs but behaves linearly in q, at large qs. Third, we observe that the slope and intercept of this average linear behavior do not vary (or only extremely slowly vary) when N is increased (estimated slopes as a function of N are reported in Fig. 5.3, right plot). Such observations can be gathered as follows: $\hat{\mathbb{E}}_{MC}\hat{\zeta}(q) \simeq \zeta(q)$, where

$$\zeta(q) = \begin{cases} \lambda(q), & \text{if } q \leq q_*^+, \\ 1 + qh_*^+, & \text{if } q > q_*^+, \end{cases} \tag{5.13}$$

$$h_*^+ = \min_h \{D_A(h) = 0\}, \quad q_*^+ = (dD_A/dh)_{h=h_*^+}. \tag{5.14}$$

It is worth mentioning again that one necessarily has $q_*^+ \leq q_c^+$, and q_*^+ is often far smaller than q_c^+. The equations above are fully consistent with the results in [127, 134] that were previously obtained for the specific case of Mandelbrot cascades. It is formulated as a general conjecture for multifractal processes in [108]. It can appear paradoxical as ensemble averages (in Eq. (5.10)) and time averages (in Eq. (5.9)) differ.

5.2.3 Extreme values and heavy tails

Structure functions and extreme values. Simple algebra yields that the structure functions $S_N(2^j, q)$ are driven by the largest increment at scale $a = 2^j$,

$$M_{n_j}(2^j) = \max\{|T_A(2^j, 2^j k)|,\ k = 1, \ldots, n_j\} \tag{5.15}$$

for fixed N, in the limit $q \to +\infty$:

$$S_N(2^j, q) \simeq \frac{1}{n_j}(M_{n_j}(2^j))^q,$$

[4] $\hat{\mathbb{E}}_{MC}$ stands for means over N_{MC} realizations.

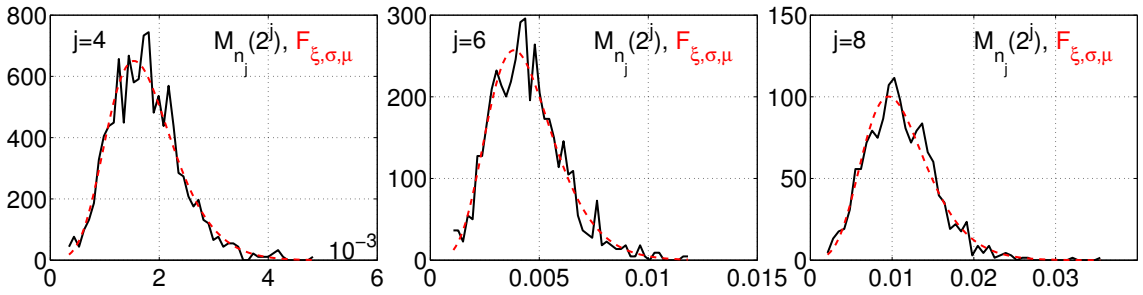


Figure 5.5: **Extreme value distribution fits.** Empirical distributions of the maxima $M_{n_j}(2^j)$ as in (5.15) (solid black line) and their best GEV fits (dashed red lines) for scales $a = 2^j$ with $j = 4, 6, 8$.

or, equivalently:

$$\log_2 S_N(2^j, q) \simeq -\log_2 n_j + q \log_2 M_{n_j}(2^j). \tag{5.16}$$

The moments of $T_A(a,t) = A(t + a\tau_0) - A(t)$ are finite only up to order $0 < q < q_c^+$, and therefore the variables $T_A(a,t)$ have heavy tails (see e.g. [160] or [69], Chapter 8, for an introduction to heavy tailed distributions). Be they independent, the order q for which $M_{n_j}(2^j)$ takes the control of $S_N(2^j, q)$ should be such that $T_A(a,t)^q$ has infinite mean, i.e., when $q \geq q_c^+$. Fig. 5.4 illustrates that the relevance of (5.16) actually starts for $q \simeq q_*^+ \leq q_c^+$.

Extreme value distributions. It is well-known that the distributions of maxima of i.i.d. random variables are modeled by extreme value distributions [69]. In the present study, the variables $T_A(a,t)$ have heavy tails, hence so do the $|T_A(a,t)|^q$, $q > 0$. Therefore, the maximum taken over *independent* $|T_A(a,t_k)|$, $k = 1, ..., n_a$, would theoretically follow a Frechet distribution with a power law tail $x^{-q_c^+}$ as $x \to +\infty$ [69]. For a given realization of CPM, the $|T_A(a,t_k)|^q, k = 1, ..., n_a$, entering the sums $S_N(2^j, q)$, are, by construction of CPM, dependent so that the limit distribution of their maxima is not theoretically known. Therefore, we chose to fit the distribution of $M_{n_j}(2^j)$, separately at each scale $a = 2^j$, using the generalized extreme value (GEV) probability density distribution, whose cumulative distribution function reads [69]:

$$F_{\xi,\sigma,\mu}(x) = \exp\left\{-\left[1 + \xi((x-\mu)/\sigma)\right]^{-1/\xi}\right\}.$$

Extreme value fits. Fig. 5.5 clearly indicates a satisfactory agreement between the empirical distribution functions of $M_{n_j}(2^j)$ and the GEV distribution. Moreover, Fig. 5.6 (left plot) shows unambiguously that the estimated parameter ξ depends neither on the scale 2^j nor on the sample size N:

$$\xi_{j,N} \simeq \xi_0. \tag{5.17}$$

Simple algebra shows that the tail of the GEV probability density function is controlled by the exponent $1/\xi$. The estimated $1/\xi_0$ turns out to be very far from the exponent q_c^+ that would be expected under independence of the $T_A(a,t_k)$ and happens to be consistently close to q_*^+ (cf. Fig. 5.6, left plot). Similar observations can be obtained from the tails of the empirical distributions of $|T_A(a,t)|$.
Moreover, Fig. 5.6 (middle plots) clearly shows that the coefficients $\mu_{j,N}$ and $\sigma_{j,N}$ are

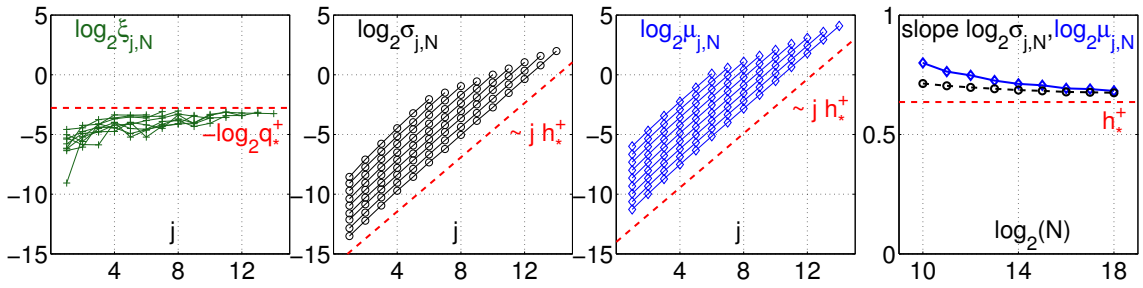


Figure 5.6: **Extreme value fits and multifractal properties.** $\log_2 \xi_{j,N}$ ('+') (left plot) $\log_2 \sigma_{j,N}$ ('$\circ$') and $\log_2 \mu_{j,N}$ ('$\diamond$') (middle plots) versus j, for various N. The horizontal dashed (red) line (left plot) corresponds to $-\log_2 q_*^+$, while the diagonal one (middle plot) has slope h_*^+ (intercept being arbitrary). Right plot: estimated slopes for $\log_2 \sigma_{j,N}$ ('$\circ$') and $\log_2 \mu_{j,N}$ ('$\diamond$') do not vary significantly with N and are close to h_*^+ (dashed, red line).

characterized by power law behaviors, with respect to the scales 2^j, where the multiplicative factors depend on N, while the power law exponents do not and turn out to be equal to h_*^+, for all N (cf. Fig. 5.6, right plot):

$$\mu_{j,N} \simeq \mu_{0,N} 2^{j h_*^+}, \quad \sigma_{j,N} \simeq \sigma_{0,N} 2^{j h_*^+}. \tag{5.18}$$

These findings (Eqs. (5.17) and (5.18)) are consistent with the analyses recently proposed in [131]. Combined together, the observations above yield:

$$\{M_{n_j}(2^j)\}_{j=j_1,\ldots,j_2} \stackrel{d}{\simeq} \{2^{j h_*^+}(\sigma_{0,N}\Lambda_{\xi_0}^j + \mu_{0,N})\}_{j=j_1,\ldots,j_2}, \tag{5.19}$$

where each $\Lambda_{\xi_0}^j$ is a random variable drawn from the same $F_{\xi_0,1,0}$ GEV distribution, which does not depend on j.

Linearization effect: Slope h_*^+. Combining the definition of $\hat{\zeta}(q)$ (Eq. (2.65)) with empirical results Eqs. (5.16) and (5.19) implies, as $q \to +\infty$,

$$\begin{aligned}
\hat{\zeta}(q) &= \sum w_j \log_2 S_N(2^j, q) \stackrel{d}{\simeq} -\sum w_j \log_2 n_j \\
&+ q h_*^+ \sum j w_j + q \sum w_j \log_2(\sigma_{0,N}\Lambda_\xi^j + \mu_{0,N}) \\
&\simeq 1 + q\left(h_*^+ + \sum w_j \log_2(\sigma_{0,N}\Lambda_\xi^j + \mu_{0,N})\right),
\end{aligned}$$

since $n_j \simeq N 2^{-j}$ yields $-\sum_j w_j \log_2 n_j \simeq 1$. In itself, it explains the linearization effect observed for each realization.
Moreover, taking the average over realizations yields:

$$\hat{\mathbb{E}}_{MC}\hat{\zeta}(q) \simeq 1 + q h_*^+ + q \sum w_j \, \hat{\mathbb{E}}_{MC} \log_2(c_{0,N}\Lambda^j + d_{0,N}).$$

Since $\hat{\mathbb{E}}_{MC} \log_2(\sigma_{0,N}\Lambda_\xi^j + \mu_{0,N})$ does not depend on j and $\sum w_j \equiv 0$, this explains the linearization effect observed as an average over realizations, cf. Eq. (5.13):

$$\hat{\mathbb{E}}_{MC}\hat{\zeta}(q) \simeq 1 + q h_*^+. \tag{5.20}$$

Linearization effect: Critical order q_*^+. On the one hand, the empirical results reported above suggest a power law tail behavior $x^{-q_*^+/q}$ for the variables $|T_A(a,t)^q|$, observed from a single realization, and therefore, that they behave as if they exhibit infinite mean when $q \gtrsim q_*^+$. This explains that the maximum $M_{n_j}(2^j)^q$ takes control of the sum $S_n(q,2^j)$ at q_*^+. On the other hand, we observed in Fig. 5.3 that $\zeta(q)$ evolves continuously and without discontinuity from $\lambda(q)$ at small qs, to $1+qh_*^+$ for large qs. This implies and explains the existence of a critical order q_*^+ and defines it as:

$$\lambda(q_*^+) = 1 + q_*^+ h_*^+.$$

Using the Legendre transform in Eq. (5.14), this can be rewritten in clear agreement with the definition Eq. (5.14) as well as with the conjecture in [108] (stated in Eq. (2.129)), as:

$$1 + q_*^+ (d\lambda/dq)_{q=q_*^+} - \lambda(q_*^+) = 0. \tag{5.21}$$

These two different arguments explain separately that the linearization effect starts to occur when $q \gtrsim q_*^+$.

5.2.4 Conclusions and perspectives

Multifractal properties and extreme values. Observations Eq. (5.19), indicating that $M_{n_j}(2^j) \simeq C_N 2^{jh_*^+}$, as $a = 2^j \to 0$, where C_N is a suitable random variable, are strikingly consistent with the multifractal paradigm. Indeed, recall that multifractal analysis associates with each time position t a Hölder exponent as $|T_A(a,t)| \simeq c(t)a^{h(t)}$, as $a \to 0$. Then, the largest increments (hence the maxima) are observed in the limit $a \to 0$ for the smallest h, that is where $A(t)$ is the most singular. By the definition Eq. (5.14), such smallest exponent is h_*^+.

Heavy tails, dependence and linearization effect. The analyses reported here show that the existence of the linearization effect is a combined consequence of two major properties of CPM: Their increments are heavy tailed and possess a specific dependence structure resulting from the multiplicative construction.

Perspectives. First, it is conjectured that the present analyses of the linearization effect holds for all multifractal processes and not only CPM or those resulting from multiplicative constructions (such as CMC or IDC, cf. Sections 2.7.1-a) and 2.7.1-c), respectively). Indeed, multifractal processes will in general gather the two key ingredients mentioned above: Heavy tails and a form of time dependence structure, which the multifractal spectrum characterizes in an indirect way. Second, a full and relevant multifractal analysis needs to be based on wavelet Leaders rather than on increments and involves both positive and negative qs [89, 92], see also Section 2.5.5). It is of interest to understand how these relations between multifractality, heavy tails, dependence, extreme values and linearization effect extend to this more accurate framework and accommodate the negative qs. Indeed, it is conjectured in [108] that for negative statistical orders q, which are practically accessible by wavelet Leader based estimation, there is a negative critical order $q_*^- = \inf_q\{1 + q\frac{d\lambda(q)}{dq} - \lambda(q) = 0\}$ (cf. Eq. (2.130)), $q_c^- < q_*^-$, below which $\zeta(q)$ and $\lambda(q)$ do not coincide. Preliminary results show that similar arguments apply, with minima $m_{n_j}(2^j)$ and q_*^-, h_*^- replacing maxima $M_{n_j}(2^j)$ and q_*^+, h_*^+, respectively. These two research directions need further investigation.

5.3 Data Quantization and Multifractal Analysis

Most stochastic models used to describe scaling in real-life data are continuous time and continuous valued processes. However, for most applications, the analyzed data are sampled in time. The impact of this sampling on the estimation of the multifractal parameters has been analyzed in various articles (e.g., [70, 126, 172]) and will not be considered here. In a number of situations, the data available for the analysis also present quantization in amplitude. This is very often the case in Image Processing where the necessarily limited sizes of images yield quantized boundary lines separating various regions in the image, as illustrated in Fig. 5.7. An informative example is provided by the analysis of crack propagations, where the data consist of boundary lines that split images into two binary regions. It is conjectured that the characteristics of the crack propagation can be inferred from the analysis of the scaling properties of these boundary lines (cf. [118, 154] for a thorough description of this application). Because it is often needed that a large number of such images are captured along time, this may in addition impose that sensors are used at poor resolution levels, hence resulting into the fact that the boundary lines are available for the analysis only through (possibly severely) quantized versions. The same limitation may also pertain for the high-speed acquisition of one dimensional signals, as illustrated in Fig. 5.8.

So far, the impact of this quantization on the statistical performance of the procedures aiming at estimating multifractal attributes, defined in Section 2.6, received little attention. Its characterization precisely constitutes the aim of the present section.

One of the main goals of this section consists in studying the robustness against quantization that the choice of a particular multiresolution quantity brings (or not) to multifractal analysis. To this end, we compare the performance of analysis procedures designed from different multiresolution quantities, namely first and second order increments – motivated by the fact that they have been and still are used in practice for multifractal analysis by many physicists – wavelet coefficients and wavelet Leaders. The work reported in this section has been obtained in collaboration with Stéphane G. Roux (ENS Lyon) and published in [181].

5.3.1 Quantization

In general, a quantizer can be defined as a set of intervals $\mathcal{D} = \{D_i, i \in I\}$ covering a space (e.g. the real line $\mathbb{R}$) together with a set of levels $\mathcal{Y} = \{y_i, i \in I\}$, so that the quantizer q is given by $q(x) = y_i \Leftrightarrow x \in D_i$. Quantization is often viewed as a decomposition of the signal $x = \tilde{x} - z$ into the quantized signal $\tilde{x}$ and quantization noise or quantization error $z = \tilde{x} - x$, i.e.:

$$\tilde{x} = x + z. \tag{5.22}$$

The quality of quantization is usually assessed by some distortion measure based on this difference signal, for instance the mean squared (quantization) error, or a signal to (quantization) noise ratio.

In contrast to Nyquist sampling, quantization is a non linear operation and thus much more intricate to study, since linear system theory does not apply directly. The theory of quantization has been developed roughly over the last 60 years. Most results come from information theory and are typically focussing on the study of (average) distortion versus rate – (average) number of bits per input sample – of quantization. Another line

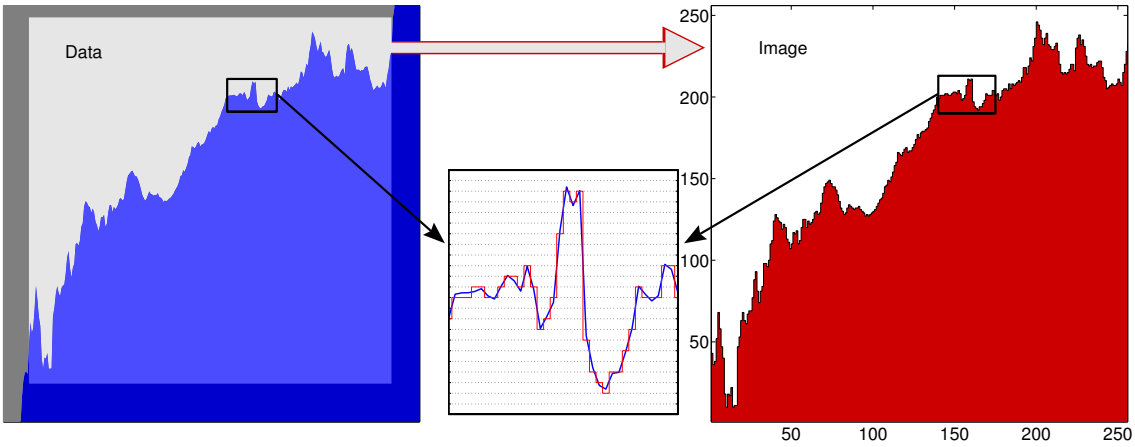


Figure 5.7: **Quantization of 2d data.** Illustration of quantization process in 2d data acquisition.

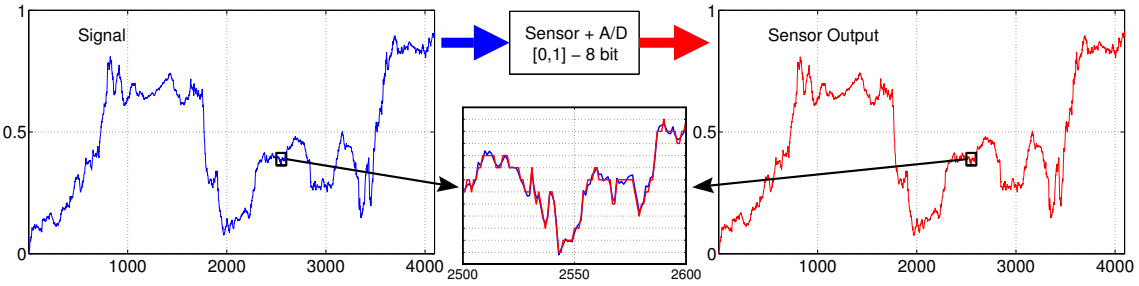


Figure 5.8: **Quantization of 1d data by sensor.** Illustration of quantization in 1d data acquisition by a sensor.

of research adopts a more statistical point of view and studies the statistical distributions of input and quantized signals. It establishes, for instance, quantization theorems – in analogy to sampling theorems – stating conditions on the band limit nature of the characteristic function of the input process for perfect reconstruction of the input process (or its moments). For more details on the theory of quantization, the reader is referred to e.g. [78, 184] and references therein.

The point of view we adopt in this section is mainly guided by the two applications we evoked in the introduction and has to be appropriate for giving answers to the following practically relevant issues: Given that the signal we want to analyze is quantized, what is the influence on the estimation of multifractal attributes? Which multiresolution quantities should be chosen? What can be done to limit the influence of quantization on estimation performance in multifractal analysis?

In this setting, we typically do not have any flexibility in choosing an optimal quantization scheme. For instance, an image of a crack will always give rise to a signal that is uniformly quantized, with quantization interval width fixed by the resolution of the camera (as illustrated in Fig. 5.7). Therefore, information theoretical considerations in terms of (optimal) rate and distortion do not seem very relevant. Also, as a consequence of the difficult statistical context posed by scale invariant processes (e.g. non stationarity), an analytical statistical point of view may not be very helpful in answering the questions stated above. Therefore, we will mainly resort to numerical simulations. We choose to concentrate on

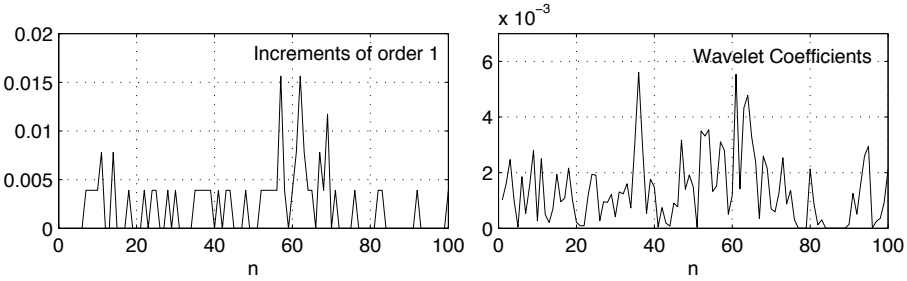


Figure 5.9: **Quantization and coefficients of multiresolution quantities.** Time series of $|T_X^{(I_1)}(2^j,t)|$ (left) and $|T_X^{(d)}(2^j,t)|$ (right) ($j=4$) for quantized ($b=7$) signals.

the study of the estimator for the log-cumulant c_p (Eq. (2.71)) for multifractal trajectories that are subject to quantization. The numerical results enable us to give a frequency-domain interpretation of the (non)robustness of certain multiresolution quantities against quantization, and to provide precise practical guidelines for multifractal analysis of signals subject to amplitude quantization.

5.3.2 Quantization and multifractal analysis: Numerical study

5.3.2-a) Quantization

We consider uniform quantization over the real axis, with an infinite number of levels and quantization interval width Δ. The quantized version $X^{\Delta}(n)$ of the original discrete-time signal $X(n)$ is therefore defined as:

$$X^{\Delta}(n) = [X(n)/\Delta] \cdot \Delta, \tag{5.23}$$

where $[\cdot]$ denotes the rounding operation. We measure the quantization level (in bit) by:

$$b = -\log_2 \Delta. \tag{5.24}$$

Note that this quantization level measure is absolute in the sense that it simply expresses the number of bits necessary to (fixed rate) code the quantized unit interval, and does not take into account a measure of the variability of the signal. This absolute measure is chosen for the following reason: The non stationary nature of scaling processes makes it difficult to give a meaningful definition of a measure taking into account characteristics of the signal for one particular scaling process, for instance a signal to quantization noise ratio. What is more, even if we could define such a measure (in average) for one process, it would not allow us to meaningfully compare quantization levels for two different processes since the processes will in general have different average time evolutions of their statistics (mean, variance, flatness etc.). Therefore, quantitative and qualitative comparisons are restricted to comparisons between different quantization levels for one given process with fixed parameters, for which the quantization level b is a sufficient measure. For this reason, we will as well present results for one multifractal process with fixed (multifractal) parameter setup only. Similar results have been obtained for other mono- or multifractal processes.

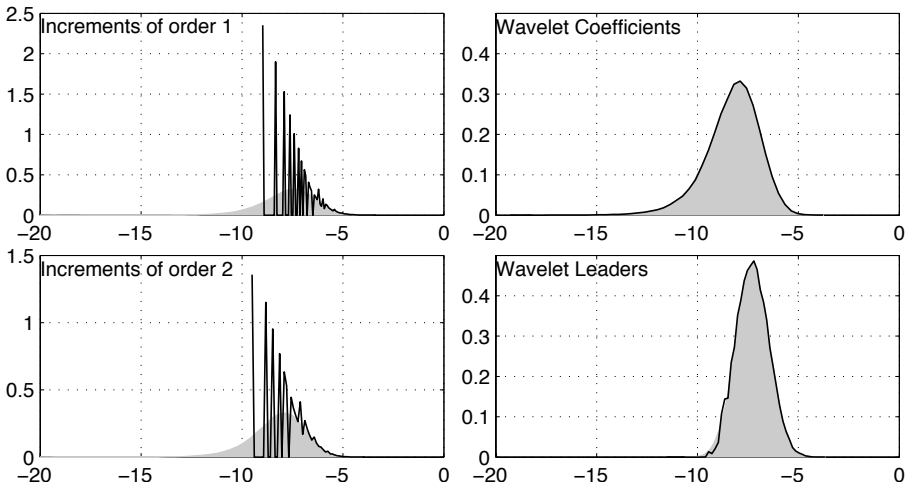


Figure 5.10: **Quantization and empirical distributions.** Empirical distribution of $\ln|T_X(2^j,\cdot)|$ ($j = 3$) for non quantized (gray area) and quantized, $b = 12$, (solid line) signals.

5.3.2-b) Multiresolution quantities

Our main goal is to study the relative robustness towards quantization of estimation procedures Eq. (2.71) based on different multiresolution quantities $T_X(a,k)$ of $X(k)$. We consider increments (cf. Eq. (2.13)) of orders 1 and 2 for dyadic analysis scales $a = 2^j$, discrete wavelet coefficients (Eq. (2.34)) and wavelet Leaders (Eq. (2.54)). For convenience, they are denoted as follows:

$$T_X^{(I_1)}(2^j,t) = \quad \delta_{2^j\tau_0}X(t) \quad = X(t+2^j\tau_0) - X(t), \tag{5.25}$$

$$T_X^{(I_2)}(2^j,t) = \quad \delta_{2^j\tau_0}(\delta_{2^j\tau_0}X(t)) \quad = X(t+2\cdot 2^j\tau_0) - 2X(t+2^j\tau_0) + X(t), \tag{5.26}$$

$$T_X^{(d)}(2^j,t) = \quad d_X(j,k) \quad = \langle\psi_{j,k}|X\rangle, \tag{5.27}$$

$$T_X^{(L)}(2^j,t) = \quad L_X(j,k) \quad = \sup_{\lambda'\subset 3\lambda_{j,k}} |d_{\lambda'}|, \tag{5.28}$$

where τ_0 stands for an (arbitrary) time unit which is, by consistency with the definition of wavelet coefficients, chosen as the inverse of the sampling frequency.

5.3.2-c) Thresholding

The estimation of log-cumulants Eq. (2.71) is based on the log of the multiresolution quantities. Therefore, it is necessary to remove coefficients that take the value zero. This thresholding operation is performed by considering only coefficients that satisfy $|T_X(a,k)| > \varepsilon$ in Eq. (2.71), where ε is the machine precision.

5.3.2-d) Numerical simulations

The impact of quantization on the estimation procedures, with multiresolution quantities Eqs. (5.25–5.28), is assessed by applying them to a large number N_{MC} of realizations of mrw at different quantization levels b and studying their relative performance. All results presented here are obtained using Daubechies wavelets with $N_\psi = 2$, using $N_{MC} = 1000$ realizations of size $N = 2^{14}$, with process parameters $(H, \beta) = (0.72, \sqrt{0.08})$, i.e., $(c_1, c_2) = (0.8, -0.08)$ (cf. Section 2.7.2-b)).
The estimation performance are assessed by the (root) mean squared error:

$$mse = \sqrt{\left(\hat{\mathbb{E}}_{MC}\hat{c}_p - c_p\right)^2 + \widehat{\mathrm{Var}}_{MC}\hat{c}_p},$$

where $\hat{\mathbb{E}}_{MC}$ and $\widehat{\mathrm{Var}}_{MC}$ denote the sample mean and sample variance over Monte Carlo realizations, respectively.

5.3.3 Quantization impacts

5.3.3-a) Distributions of $\ln|T_X(2^j,\cdot)|$

Let us first study the empirical distributions of $\ln|T_X(2^j,\cdot)|$. Fig. 5.10 compares the empirical distributions of $\ln|T_X(2^j,\cdot)|$ (for a given j) for non quantized and quantized at level b signals. We observe that while quantization does not have any visible impact on the distributions of $\ln|T_X^{(d)}(2^j,\cdot)|$ and $\ln|T_X^{(L)}(2^j,\cdot)|$, the distributions of $\ln|T_X^{(I_1)}(2^j,\cdot)|$ and $\ln|T_X^{(I_2)}(2^j,\cdot)|$, obtained from quantized data, are lattice and significantly different from the distribution obtained for the non quantized signal. This will, in turn, affect the estimation of the $\hat{C}(j,p)$.

5.3.3-b) $\hat{C}(j,p)$ as linear functions of j

The estimation procedure Eq. (2.71) is based on the fact that for scale invariant processes, the cumulants $C(j,p)$ of $\ln|T_X(2^j,\cdot)|$ (Eq. (2.72)) behave as linear functions of j for some range of scales 2^j (cf. Eq. (2.68)). Let us now investigate the impact of quantization on this central element of the estimation procedure. Fig. 5.11 compares the means over Monte Carlo realizations of $\hat{C}(j,1)$ and $\hat{C}(j,2)$ as functions of j, for non quantized data and for data quantized at different levels b. It yields a central observation: Quantization affects $\hat{C}(j,p)$ at fine scales first, and then at coarser and coarser scales as Δ increases (equivalently b decreases). Also, we observe that this impact is much more dramatic for increments than for wavelet coefficients and Leaders. Whereas for increments the influence of quantization propagates very fast up to the coarsest scale as Δ increases, it remains restricted to fine scales for wavelet coefficients and Leaders, leaving the coarser scales unchanged and usable to perform the linear regressions yielding $\hat{c}_p$. For wavelet coefficients and Leaders, a meaningful range of scales for linear regression can still be found for quantizations significantly below $b = 5$ for $\hat{c}_1$ and $b = 8$ for $\hat{c}_2$, whereas for increments, linear regression is meaningless already for $b = 7$ for $\hat{c}_1$ and $b = 12$ for $\hat{c}_2$.

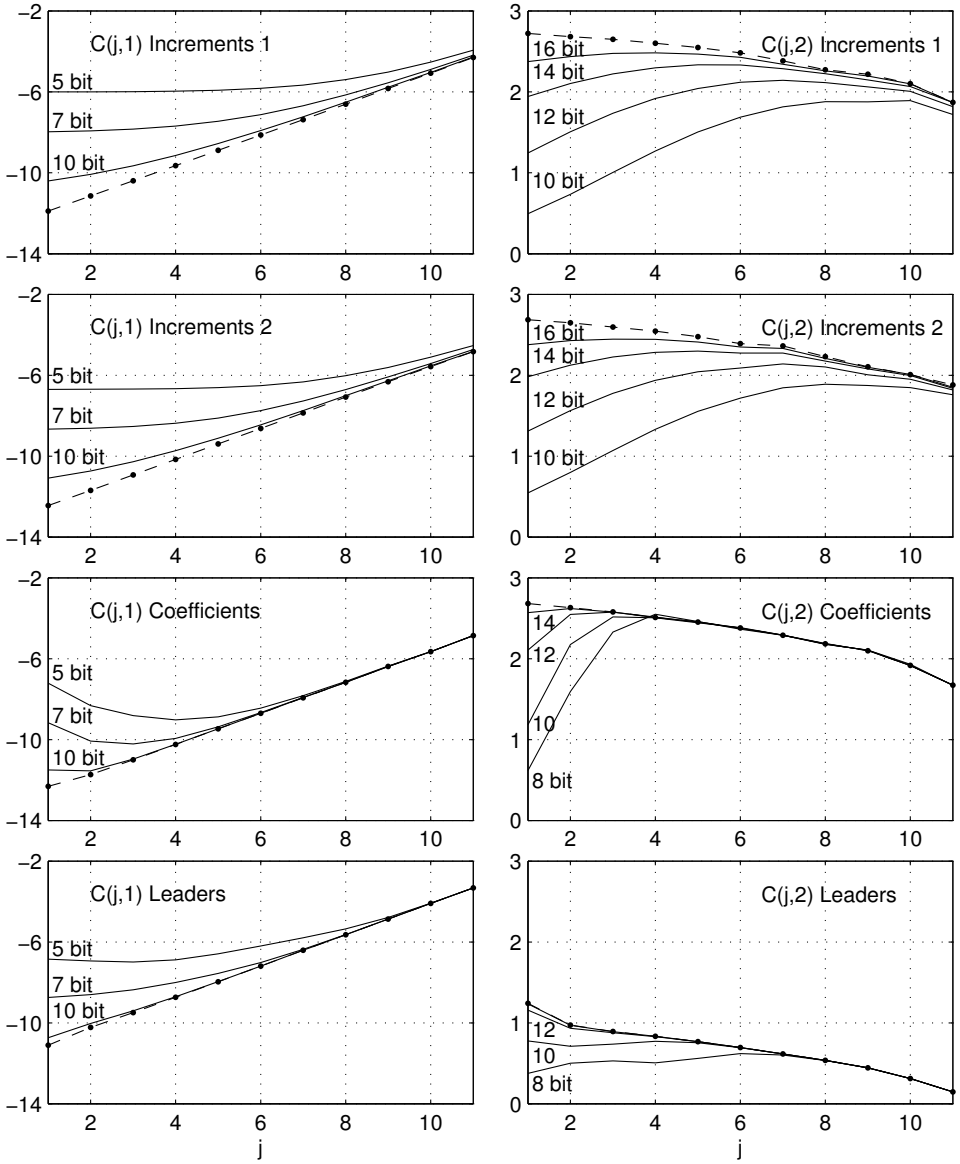


Figure 5.11: $\hat{C}(j,p)$ **as linear functions of** j**.** $\hat{C}(j,1)$ (left column) and $\hat{C}(j,2)$ (right column) vs. j for non quantized data (dashed dotted line) and data quantized at different levels b (solid lines).

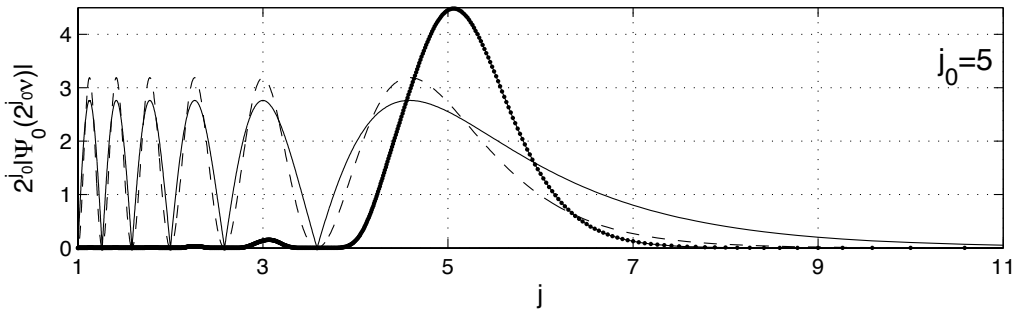


Figure 5.12: **Band pass filters.** Fourier transforms of the mother wavelets $|2^{j_0}\Psi_0(2^{j_0}\nu)|$ versus $j = -\log_2(\nu)$ for Daubechies ($N_\psi = 2$) (solid dotted line), Increments of first (solid line) and second (dashed line) orders.

Interpretation of altered scaling behavior. This can be understood as follows. The increments $T_X^{(I_1)}(2^j, t)$ and $T_X^{(I_2)}(2^j, t)$ can be read as wavelet coefficients obtained with specific mother-wavelets: $\psi_0(t) = \delta(t+\tau_0) - \delta(t)$ and $\psi_0(t) = -\delta(t+2\tau_0) + 2\delta(t+\tau_0) - \delta(t)$, respectively. Such ψ_0 possess respectively $N_\psi = 1$ and $N_\psi = 2$ vanishing moments and are commonly referred to as *poor man's* wavelets, because they act as band pass filters whose Fourier transforms $\Psi_0(\nu)$ are poorly localized in frequency, compared to those of standard mother wavelets, such as the Daubechies. This is illustrated in Fig. 5.12 where the Fourier transforms of the increments are compared to that of a Daubechies wavelet ($N_\psi = 2$). For simplicity the frequency axis is in octaves $j = -\log_2 \nu$. For small frequencies, the behavior of the Fourier transforms is controlled by N_ψ according to $|\Psi_0(\nu)| \sim C \cdot |\nu|^{N_\psi}, |\nu| \to 0$. For large frequencies, $|\Psi_0^{(d)}(\nu)|$ is characterized by a good frequency localization. In contrast, the filter responses of increments are given by $|\Psi_0^{(I_1)}(\nu)| \sim |\sin(\pi\nu)|$ and $|\Psi_0^{(I_2)}(\nu)| \sim \sin(\pi\nu)^2$ and show much poorer frequency localizations with important side lobes whose amplitudes do not decrease. This poor frequency localization turns out to have a significant impact on the robustness of the multifractal parameter estimation procedures against quantization.

Indeed, Eq. (5.22) evokes that quantization mimics noise superimposition to the original non quantized data x. Fig. 5.11 suggests that this noise z mostly contributes at fine scales, or equivalently, at high frequencies. Hence, any estimates involving such scales are poor whatever the chosen multiresolution quantity. However, the well-localized in frequency nature of the wavelet band pass filters significantly limits the contamination of larger scales by the noise. Therefore, restricting the linear regressions to larger scales yields satisfactory estimates. Conversely, the poor frequency localization of the increment band pass filters results in a significant pollution of the large scales by the fine scale noise. This implies that, to perform estimation, one has to restrict the regression range to much larger scales, if there are any left that are not polluted, which hence significantly degrades estimation performance.

5.3.3-c) Statistical performance

Increments. $T_X^{(I_1)}(2^j, \cdot)$ and $T_X^{(I_2)}(2^j, \cdot)$ take on fewer and fewer different discrete values as Δ increases, until eventually they only consists of 0 or Δ values. Then, since the $T_X(a, k)$ are thresholded at machine precision ε, $\hat{C}(j, 1) = \mathbb{E} \ln |T_X^{(I)}(2^j, \cdot)| \to \ln \Delta$ and

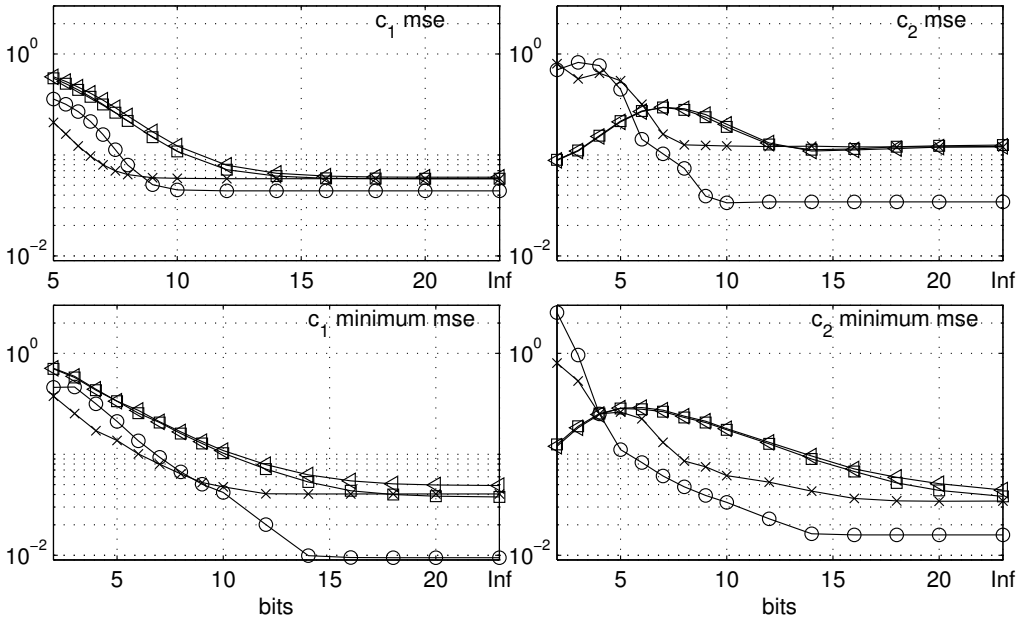


Figure 5.13: **Quantization and statistical performance.** mse of $\hat{c}_1$ (left column) and $\hat{c}_2$ (right column), obtained for a fixed regression range $j_1 = 5$, $j_2 = 11$ (top row), and for the optimal regression range j_1, j_2 (bottom row). The symbols ($\triangleleft$, $\square$, $\times$, $\circ$) correspond to (increments of order 1, increments of order 2, wavelet coefficients, wavelet Leaders).

$\hat{C}(j,2) = \mathrm{Var} \ln |T_X^{(I)}(2^j,\cdot)| \to 0$ (cf. Fig. 5.11, top row and second row). In turns, the final log-cumulant estimates become $\hat{c}_1 = 0$ and $\hat{c}_2 = 0$, no matter what values c_1 and c_2 actually take. Hence, the estimations based on $T_X^{(I_1)}$ and $T_X^{(I_2)}$ become meaningless for severely quantized signals. Therefore, their statistical performance for large Δ will not be discussed here any further. Ultimately, as Δ approaches infinity, a similar effect occurs for all T_X however, at significantly higher Δ than for increments.

Fixed regression range. Fig. 5.13 (top row) compares the mse of the estimations of c_1 and c_2, using a fixed regression range, at coarse scales $j_1 = 5$ and $j_2 = 11$. For non quantized signal (i.e., the number of bits tends to infinity), we observe that increment and wavelet coefficient based estimators achieve comparable performance, whereas the wavelet Leader based estimation is better, and significantly so for c_2, in agreement with results reported in Chapter 4. When the signal is quantized, the performance of the increment based procedures degrade dramatically and fast when b decreases, whereas the wavelet coefficient and Leader based estimations maintain their performance over an impressive range of coarse quantization levels: For c_1, the performance of the increment based procedures start degrading at $b = 15$, while that of wavelet coefficient and Leader based procedures do at $b = 9$, a factor of ≈ 60 in Δ, with a difference in mse of up to a factor 7. For c_2, the situation is similar: The performance of the increment based procedures start degrading as soon as $b = 13$, while that of coefficients and Leaders are maintained up to $b = 7$ and $b = 10$, respectively, a factor of ≈ 60 in Δ, with a difference in mse of up to a factor 10. Further, we note that the mse mainly reproduces standard deviation, apart from at severe quantization levels, where the bias becomes dominant.

Optimal regression range. In practice, the range of scales used to perform the linear regressions yielding the final estimates $\hat{c}_p$ is not fixed a priori but by visual inspection of

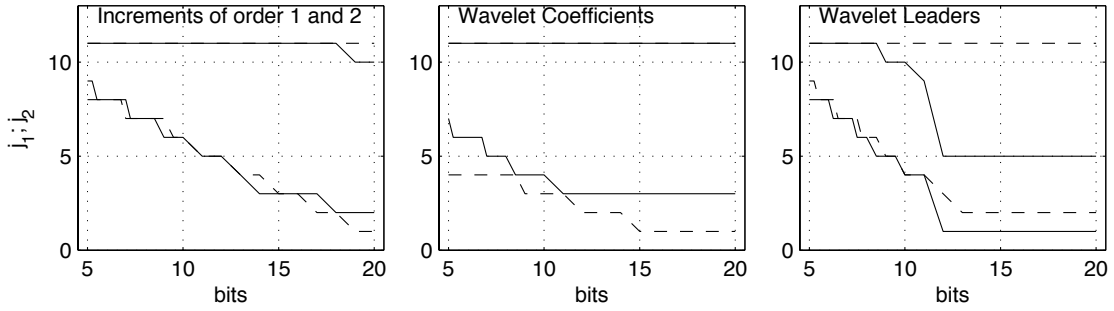


Figure 5.14: **Optimal regression range.** Optimal j_1 and j_2 to obtain minimal mse for $\hat{c}_1$ (solid lines) and $\hat{c}_2$ (dashed lines) vs. b.

$\hat{C}(j,p)$ vs. j in order to determine a region in which the scaling model is valid. Fig. 5.13 (bottom row) shows the mse obtained by choosing the regression range such that the mse of the estimation is minimal. Since the quantization affects the estimates $\hat{C}(j,p)$ starting at fine scale, we force j_1 and j_2 to be non-decreasing as Δ increases, in order to avoid meaningless estimates obtained at scales heavily affected by quantization. Comparing top and bottom rows in Fig. 5.13, we observe that, as expected, estimation can in general be improved by choosing an appropriate regression range. Whereas for increments, this improvement is small and confined to $b > 14$, relevant estimates of c_1 and c_2 are obtained still at $b = 5$ when using wavelet coefficients or Leaders, a significant improvement compared to the fixed regression range performance. We note further that whereas wavelet Leaders consistently outperform wavelet coefficients for sufficiently large b, their performance eventually degrades faster for very heavily quantized signals.
Finally, Fig. 5.14 shows the optimal regression ranges. As expected, increasing Δ forces j_1 to increase, restricting the estimation to coarser and coarser scales. This happens much earlier for increments than for wavelet coefficients and Leaders.
Conclusions. These analyses lead us to conclude that increments can not be used when the data are quantized, even for low quantization levels. We found that wavelet coefficients and Leaders are significantly more robust to quantization than increments of any order. By choosing an appropriate range of scales for regression, the effects of quantization on $\hat{c}_p$ can be circumvented even for coarse Δ levels when using wavelet coefficients and Leaders, whereas this is not the case for increments.

5.3.4 Conclusions and perspectives

We showed here that signal quantization can significantly impair multifractal analysis. Mostly, it pollutes the finest scales, hence implying a restriction towards the largest scales of the range of scales used in the linear regression involved in multifractal parameter estimation. However, we demonstrated that choosing mother wavelets with a good frequency resolution contains the noise pollution to as low as possible scales, hence limiting the necessary narrowing of the regression range and the estimation performance degradation. Conversely, the absence of localization of the frequency response of the increment based band pass filters results in a stronger narrowing of the regression range for a given quantization level and hence in poor performance. Therefore, wavelet coefficients and Leaders are to be preferred over increments of any orders to analyze quantized data. Also, we showed that wavelet Leaders consistently outperform wavelet coefficients for

non quantized data as well as for a large range of quantization levels. It is only for very heavily quantized signals that wavelet coefficients eventually become more robust than wavelet Leaders. These results provide the practitioners with a careful framework for real life data analysis, in situations where amplitude quantization occurs.

An automatic selection of the most relevant regression range of scales given a quantization level is an open and practically important issue. Also, the impact of quantization in image processing in situations where textures are described using multifractal models, but where the amplitudes are quantized, would constitute an important continuation of this work.

Part II

Bootstrap and Multifractal Analysis: Theory and Practice

Scaling analysis and multifractal analysis in practice mostly amount to measuring scaling exponents $\zeta(q)$, log-cumulants c_p and the multifractal spectrum $D(h)$ from a finite size observation. The theoretical and practical characterization of the corresponding *estimation* procedures have been the topic of Part I of this manuscript. In real life applications, however, practical interest lies as much in the *confidence* that can be granted to an estimate as in the estimate itself. Equivalently, *statistical tests* validating the precise multifractal properties of the data under analysis are a major practical concern. Surprisingly, and despite its increasing popularity, multifractal analysis suffers here from a significant difficulty: Little is known theoretically on the statistical performance of the estimation procedures commonly used in practice. When X is a Gaussian self similar process, these statistical performance can be studied, and asymptotic results have been obtained (for instance, in [4, 26, 52, 84, 169]). However, for the more general and useful case of MMC processes and for non Gaussian H-sssi processes, no theoretical statistical performance study is available. This is primarily a consequence of the fact that most stochastic multifractal processes practically used are defined from multiplicative martingales. The construction of such mathematical models is exceptionally intricate and gives rise to strongly dependent and heavy tailed (hence strongly non Gaussian) stochastic processes. The statistical performance of the analysis procedures hence turn out to be too difficult for analytic derivation (even asymptotically). Hence, in practice, no tools for assessing the confidence that can be accorded to the obtained estimates are available, although in real-life applications, this is as much of practical interest as the estimate itself.

Therefore, we propose here the use of *Bootstrap* as a solution to overcome such difficulties [64]. The use of bootstrap has never been reported before for multifractal analysis. Also, the proposed bootstrap estimation schemes constitute, to our knowledge, the only procedures on-hand that can actually be applied to finite size observations for practically assessing confidence in multifractal attribute estimates and affirming the multifractal nature of data with satisfactory statistical performance.

The goal of Chapters 6 and 7 in this part of the manuscript is the definition of a relevant *bootstrap resampling scheme* that fits the requirements and specific context of multifractal analysis. The key idea is the use of bootstrap in the *wavelet domain*, rather than in the domain in which the data live, hence passing – according to commonly accepted intuitions based on the decorrelation properties of the wavelet transform for fBm [4, 6, 77, 70, 166, 169] – from in general non stationary processes with intricate dependence structure to stationary series of wavelet coefficients that are expected to be only weakly correlated. The idea of bootstrap in the wavelet domain is not new and has been used for the generation of surrogate time series [13, 40, 41, 137, 151, 152] and spatial patterns [183]. The originality of the proposed bootstrap lies in the fact that it takes into account the particularities of multifractal analysis — hence concentrating on wavelet domain properties rather than surrogate data creation — and in the specific block construction that intuitively match key aspects of wavelet Leader based multifractal analysis.

This bootstrap resampling scheme can in turn be used for the design of confidence intervals for multifractal attributes, and of statistical tests aiming at rejecting hypothesis formulated on the multifractal properties of the data under analysis. Their definition and empirical validation are the goals of Chapter 8, and of Chapter 9 and 10, respectively. The practical relevance of these procedures is assessed by means of numerical simulations for several synthetic multifractal processes. The empirical results indicate satisfactory performance.

The sparseness of theoretical results for the statistical performance of multifractal at-

tribute estimation procedures not only causes the above mentioned practical limitations due to the lack of procedures for assessing confidence in estimates or significance of hypothesis on multifractal properties, but also prevents the theoretical validation of the bootstrap procedures proposed to overcome such difficulties: Some of the conditions for validity of bootstrap, related to finiteness of moments and to the structure of the estimator, can be verified relatively easily in the context of multifractal analysis and will be accounted for in Chapter 7. However, the theoretical results available at present do not enable the validation of the condition on the dependence of multiresolution coefficients involved in bootstrap resampling. Therefore, and further motivated by the empirical bootstrap performance results, the aim of Chapter 11 is to establish elements of theory for characterizing the *dependence structure* of multiresolution quantities for H-sssi and MMC processes. The analytic results are backed up with numerical simulations and suggest that the key role which the number of vanishing moments of the analyzing wavelet plays in multifractal analysis has to be seen in a new light: Whereas the (sufficiently large) number of vanishing moments is indeed responsible for weak correlation of wavelet coefficients, it is ineffective in controlling higher order dependence.

Chapter 6

Bootstrap and Resampling Techniques

Contents

This chapter states the basic facts about the bootstrap which we need in the following Chapter 7 for defining the bootstrap procedure for multifractal analysis. It sketches the general idea of the bootstrap principle, its manifestation and consistency for the i.i.d. case, and describes in more detail procedures for dependent data. Also, it synthesizes theoretical results on the consistency of block bootstrap estimation for dependent data, and theoretical and practical considerations for the choice of an optimal block size.

6.1 Bootstrap Principle

The bootstrap is a computer-intensive statistical technique that was first introduced by Efron nearly 30 years ago [64]. It has since found numerous applications in which it outperformed conventional approaches, or provided answers in problems too complex for conventional solutions to exist (cf. [83, 190, 191, 192, 194] for recent applications, and [187, 189, 195, 196] for an overview on the topic). It is recently gaining popularity due to continuously growing computer facilities.
Typically, the bootstrap is involved in inference problems where one would like to characterize the statistical properties of the estimator for a population parameter of the random process underlying the data, without making strong assumptions on the structure of the random process, as in the following situation: Let the sample $\mathcal{X}_N = \{x_1, \cdots, x_N\}$ be a finite size realization of the random process $\{X_n\}$, i.e. of a sequence of random variables (r.v.) $X_1, X_2, \cdots$ with unknown joint (population) distribution F. Suppose that the parameter of interest is the population parameter θ, which depends on the unknown population distribution F, $\theta = \theta(F)$. Also, suppose that an estimator for θ based on the sample $\mathcal{X}_N$, $\hat{\theta}_N = t(\mathcal{X}_N)$, has been fixed. Many inference problems are concerned with the statistical characterization of $\hat{\theta}_N$ with respect to θ (for instance, accuracy of $\hat{\theta}_N$ – bias, mean squared error, etc.) and hence demand for knowledge of the sampling distribution of $\hat{\theta}_N - \theta$. However, since the joint distribution of the sample $\mathcal{X}_N$ is unknown, the sampling distribution of $\hat{\theta}_N - \theta$ is unknown. The bootstrap approach to solving this type of problems is the following:

1. First, construct an estimate of the joint distribution F from the sample $\mathcal{X}_N$. In the simplest case, the estimate is the empirical distribution of the sample.

2. Second, create a *bootstrap sample* or *resample* $\mathcal{X}_N^*$ by suitably sampling from the estimate of the population. In the simplest case, this is an i.i.d. random sample from the estimate of the population. Then, approximate the relation between the population and the sample by the relation between the sample and the bootstrap sample $\mathcal{X}_N^*$.

In other words: Use the relation between sample and a suitably constructed resample – hence between $\hat{\theta}_N = t(\mathcal{X}_N)$ and $\hat{\theta}_N^* = t(\mathcal{X}_N^*)$ – to reproduce the relation between population and sample – hence between θ and $\hat{\theta}_N$ – assuming[1]:

$$\hat{\theta}_N - \theta \overset{d}{\approx} \hat{\theta}_N^* - \hat{\theta}_N.$$

[1]The above equation is of course only meaningful in the heuristic sense of the distribution of $\hat{\theta}_N^* - \hat{\theta}_N$ being "close" to the distribution of $\hat{\theta}_N - \theta$. Formal results for rescaled versions of these distributions can be obtained for specific situations. Typically, they measure "closeness" by asymptotic convergence of the distributions to the same limit distribution, and by the rate of convergence.

The estimate $\hat{\theta}_N = t(\mathcal{X}_N)$ depends only on the sample $\mathcal{X}_N$. Also, since $\mathcal{X}_N^*$ depends only on the sample $\mathcal{X}_N$, so does the *bootstrap estimate* $\hat{\theta}_N^* = t(\mathcal{X}_N^*)$. Therefore, $\hat{\theta}_N$ and $\hat{\theta}_N^*$ can always be calculated in practice.
This principle is common to all bootstrap methods and, more generally, resampling techniques. Different bootstrap methods and their particular properties arise from the exact definition of the central parts of the principle: How is the estimate for the population constructed from the sample? How are the bootstrap samples obtained from this estimate? The term *bootstrap resampling* commonly refers jointly to the construction of the population estimate and sampling from the estimate, hence to the construction of a bootstrap sample from the sample of observations. The next section gives an overview of the most important bootstrap methods and some of their properties.

6.2 Bootstrap Resampling

6.2.1 Bootstrap for i.i.d. data

The bootstrap principle is most intuitive and explicit in the case when the sample $\mathcal{X}_N = \{x_1, \cdots, x_N\}$ consists of realization of i.i.d. random variables with common marginal distribution P. The sample therefore has joint distribution $F_N = P^N$, and any population parameter of interest is a functional of the marginal distribution only, $\theta = \theta(P)$.
Let $\hat{\theta}_N = t(x_1, \cdots, x_N)$ be an estimator of θ, and suppose we are interested in the distribution of $\hat{\theta}_N - \theta$, which we call G_N. The distribution G_N is unknown since P and θ are unknown. The bootstrap approach for estimation of G_N in the i.i.d. case is:

1. First, construct an estimate $\hat{P}_N$ of P from the available data $\mathcal{X}_N = \{x_1, \cdots, x_N\}$. Commonly, $\hat{P}_N$ is chosen to be the empirical distribution function of the sample,

$$\hat{P}_N(y) = \frac{1}{N} \sum_{n=1}^{N} h(y - x_n),$$

   where $h(\cdot)$ denotes the Heaviside (step) function[2]. This gives rise to the *nonparametric i.i.d. bootstrap* and corresponds to Efron's original resampling plan [64]. Alternatively, if a plausible parametric model for P is available, $\hat{P}_N$ can be obtained by fitting a model distribution to the sample $\mathcal{X}_N$. This alternative choice of $\hat{P}_N$ gives rise to the *parametric bootstrap*.

2. Second, reproduce the relation between the population and the i.i.d. sample $\mathcal{X}_N$ by drawing an i.i.d. bootstrap sample $\mathcal{X}_N^* = \{x_1^*, \cdots, x_N^*\}$ from the estimated distribution $\hat{P}_N$:

$$\mathcal{X}_N^* = \{x_1^*, \cdots, x_N^*\} : \quad x_n^* \overset{i.i.d.}{\sim} \hat{P}_N.$$

   For the nonparametric i.i.d. bootstrap, this corresponds simply to random drawing, with replacement, samples x_n from the original sample $\mathcal{X}_N$.

The bootstrap version of $\hat{\theta}_N = t(x_1, \cdots, x_N)$ is $\hat{\theta}_N^* = t(x_1^*, \cdots, x_N^*)$. The conditional distribution $\tilde{G}_N$ of $\hat{\theta}_N^* - \hat{\theta}_N$ is usually approximated by its empirical distribution, which is obtained numerically by simulation, i.e., drawing a large number R of resamples $\mathcal{X}_N^{*(r)}$, $r =$

[2]The Heaviside function is defined as: $h(t) = 1$ for $t \geq 0$, and $h(t) = 0$ for $t < 0$

$1, \cdots, R$, and calculating bootstrap estimates $\hat{\theta}_N^{*(r)}$, $r = 1, \cdots, R$. Then:

$$G_N(y) \approx \tilde{G}_N(y) = \frac{1}{R} \sum_{r=1}^{R} h(y - (\hat{\theta}_N^{*(r)} - \hat{\theta}_N)).$$

It can be shown that under mild conditions on P and on the nature of $\hat{\theta}$, the conditional distribution $\tilde{G}_N$ of $\hat{\theta}_N^* - \hat{\theta}_N$ is a consistent estimator of the distribution G_N of $\hat{\theta}_N - \theta$. For instance, if the parameter of interest is the centered and scaled sample mean $T_N = \sqrt{N}(\bar{x}_N - \mu)/\sigma$ and its bootstrap version is $T_N^* = \sqrt{N}(\bar{x}_N^* - \bar{x}_N)/\widehat{\mathrm{Std}}_N X$, it can be shown [38, 161] that if the variance of the population exists, $\sigma^2 = \mathrm{Var} x < \infty$, then:

$$\sup_y |P^*(T_N^* \leq y) - P(T_N \leq y)| = o(1) \quad a.s. \quad \text{as} \quad N \to \infty.$$

Also, under some additional conditions (existence of third moment and non-degeneracy of the population distribution):

$$\sup_y |P^*(T_N^* \leq y) - P(T_N \leq y)| = O(N^{-1}(\log\log N)^{1/2}) \quad a.s. \quad \text{as} \quad N \to \infty,$$

and therefore, not only is the bootstrap approximation a consistent estimator, but it is more accurate than the standard normal approximation, which is of order $O(N^{-1/2})$ only.

Extensive accounts for the i.i.d. bootstrap can be found in, e.g., [25, 48, 56, 65, 68, 81, 119, 157].

6.2.2 Bootstrap for dependent data

In many practical situations, data display some form of dependence, and this is also in general the case for empirical quantities involved in estimation procedures in multifractal analysis. In such situations, the simple i.i.d. bootstrap is inappropriate and fails in general. To appreciate why, consider the seminal work of Singh [161], which investigates the specific case of the centered and scaled sample mean for m-dependent data. The work shows that the i.i.d. bootstrap distribution of the centered and scaled sample mean does converge to a normal distribution, but with a variance that is different from the variance of the distribution of the centered and scaled sample mean. The i.i.d. bootstrap resampling, consisting in drawing with replacement one observation at a time, entirely ignores the dependence structure of the data and hence the lag-covariance terms in the asymptotic variance of the (centered and rescaled) sample mean.

6.2.2-a) Model based bootstrap

The model based bootstrap is a simple way to extend the i.i.d. bootstrap to dependent data. It can be used when the data admits specific model assumptions, such as:

$$X_n = f(X_{n-1}, \cdots, X_{n-p}; \pi) + \varepsilon_n,$$

where π is a vector of parameters, and $\{\varepsilon_n\}$ is a sequence of i.i.d. random variables, independent of the random variables $\{X_1, \cdots, X_p\}$, with common distribution[3], or such

[3] For instance, this could be an $AR(p)$ process.

as (linear) regression models [56]. If such model assumptions are applicable, the i.i.d. bootstrap can be used by applying it to the i.i.d. innovations ε_n, rather than applying it to the dependent original sample. After estimation of the parameters $\hat{\pi}$ and the residuals $\hat{\varepsilon}_n$ from the estimated model, a bootstrap sample is obtained by i.i.d. bootstrap resampling from the residuals $\hat{\varepsilon}_n$ (cf. [67], [74] or [101] for more details):

$$X_n^* = f(X_{n-1}^*, \cdots, X_{n-p}^*; \hat{\pi}) + \varepsilon_n^*.$$

6.2.2-b) Block bootstrap

The model based bootstrap is only applicable in the specific cases where the data admit rather strong model assumptions. In many practical situations, such assumptions are not justified, or there is not enough structural knowledge on the data available. Alternatively, block bootstrap methods are applicable to dependent data in a nonparametric and model-free manner. The key idea is to resample blocks of consecutive samples instead of resampling one observation at a time as in the i.i.d. bootstrap. Consequently, the dependence structure of the original data is preserved within each block. Under certain assumptions, the block bootstrap asymptotically reproduces the dependence structure of data generated by a weakly dependent process.
Block bootstrap methods are popular for statistical inference for dependent data, since they do not require many structural assumptions for the data. There exist many variants of block bootstrap, for instance [42] and [79] (Nonoverlapping Block Bootstrap (NBB)), [102] and [115] (Moving Block Bootstrap (MBB)), [140] (Circular Block Bootstrap (CBB)) or [141] (Stationary Block Bootstrap (SBB)).
An extensive account for all of these methods would go beyond the scope of this synthetic overview (cf. [104] and references therein for more details). In what follows, we will therefore concentrate only on the most popular and efficient block bootstrap methods. [104] shows numerically and analytically that the MBB and CBB outperform the NBB and SBB in terms of mean squared error in many situations. Therefore, we concentrate in this section on the MBB and in particular the CBB method and describe them in more detail. Also, we provide a short synthesis of conditions for asymptotic consistency and related theoretical aspects. An extensive account for theoretical and practical properties of these methods, as well as their limitations, can be found in [104] and the references therein.

6.2.2-b)-1 Moving Block Bootstrap (MBB) Let $X_1, X_2, \cdots$ be a sequence of stationary random variables, and $\mathcal{X}_N = \{x_1, \cdots, x_N\}$ an observation of sample size N. Let the estimator of interest be of the form[4] $\hat{\theta}_N = t(F_N)$, where F_N is the empirical distribution function of $\mathcal{X}_N$, and $t(F_N)$ is a real-valued functional of F_N. Let $l \leq N$ be an integer, and suppose for simplicity that N is a multiple of l, $N/l = m \in \mathbb{N}$. Define the blocks $\mathcal{B}_i$ of l consecutive observations starting at x_i:

$$\mathcal{B}_i = \{x_i, \cdots, x_{i+l-1}\}, \quad 1 \leq i \leq N - l + 1.$$

Let $\{\mathcal{B}_1^*, \cdots, \mathcal{B}_m^*\}$ be a simple random sample of blocks, drawn independently with replacement from the collection of available blocks $\{\mathcal{B}_1, \cdots, \mathcal{B}_{N-l+1}\}$, and let the samples

[4] This class of estimator depends only on the one-dimensional (marginal) empirical distribution and includes estimators such as, for instance, the sample mean and M-estimators of location and scale.

in the blocks $\mathcal{B}_i^*$ be denoted by $\mathcal{B}_i^* = \{x^*_{(i-1)l+1}, \cdots, x^*_{il}\}$. Then, the moving block bootstrap sample is given by:

$$\mathcal{X}_N^* = \{\mathcal{B}_1^*, \cdots, \mathcal{B}_m^*\} = \{x_1^*, \cdots, x_N^*\}.$$

The bootstrap version of the estimator $\hat{\theta}_N$ is given by $\hat{\theta}_N^* = t(F_N^*)$, where F_N^* is the empirical density function of $\mathcal{X}_N^*$. Finally, the sampling distribution of $\hat{\theta}_N - \theta$ is approximated by the empirical distribution of $\hat{\theta}_N^* - \hat{\theta}_N$. It is obtained by drawing R bootstrap resamples $\mathcal{X}_N^{*(r)}$, $r = 1, \cdots, R$ and calculating R bootstrap estimates $\hat{\theta}_N^{*(r)}$, $r = 1, \cdots, R$.

6.2.2-b)-2 Circular Block Bootstrap (CBB) The MBB resampling suffers from an undesirable boundary effect: Whereas observations x_i, $l \leq i \leq N - l - 1$ at the center of the sample are in exactly l of the blocks, the $l - 1$ observations at the beginning and those at the end of the sample are in less than l of the blocks and are hence less likely to be within a bootstrap sample than those at the center. A simple solution to this boundary problem [140] is to periodically extend the sample $\mathcal{X}_N$ by the $l - 1$ first data points, giving the "circularized" sample:

$$\tilde{\mathcal{X}}_N = \{x_1, \cdots, x_N, x_1, \cdots, x_{l-1}\}.$$

The circular block bootstrap (CBB) is given by MBB on the circularized sample $\tilde{\mathcal{X}}_N$, and hence assigns the same probability mass to each of the observations x_i.

6.3 Theoretical and Practical Aspects of Block Bootstrap Methods

6.3.1 Consistency under the smooth function model

Theoretical accounts for the consistency and performance of block bootstrap methods exist for different specific cases, for instance in the seminal works [102] and [115] for the MBB. We choose to summarize a result, taken from [104], providing sufficient conditions for consistency for block bootstrap estimation for parameters and estimators such as defined for multifractal analysis.

Suppose that $\{Y_n\}_{n\in\mathbb{Z}}$ is a real-valued stationary random process in $\mathbb{R}^{d_1}$, $f : \mathbb{R}^{d_1} \to \mathbb{R}^{d_2}$ is a Borel measurable function, and suppose that the population parameter of interest and its estimator are of the form $\theta = H(\mathbb{E}f(Y_n))$ and $\hat{\theta}_N = H(N^{-1}\sum_{n=1}^N f(Y_n))$. Here, $H : \mathbb{R}^{d_2} \to \mathbb{R}^d$ is a smooth function in the neighborhood of $\mathbb{E}f(Y_n)$. Hence, both the population parameter θ and its estimator are smooth functions of population and sample means, respectively, of the transformed sequence $\{X_n = f(Y_n)\}_{n\in\mathbb{Z}}$. This model is commonly called the *smooth function model* (see [37] and [81]). For estimators belonging to this class, [104] (Theorem 4.1) gives the following *sufficient* conditions for the consistency of the block bootstrap distribution function estimator:

$$\sqrt{N}(\hat{\theta}_N^* - \hat{\theta}_N) = \sqrt{N}(H(\hat{\mathbb{E}}_N^* f(Y^*)) - H(\hat{\mathbb{E}}_N f(Y)))$$

of the sampling distribution of:

$$\sqrt{N}(\hat{\theta}_N - \theta) = \sqrt{N}(H(\hat{\mathbb{E}}_N f(Y)) - H(\mathbb{E}f(Y))) :$$

C0 The block size l is chosen such that l goes to infinity as the sample size N grows to infinity, but at a slower rate than N: $l \to \infty$, $l/N \to 0$ as $N \to \infty$.

C1 H is differentiable in a neighborhood $N_H = \{x \in \mathbb{R}^d : ||x - \mathbb{E}X_1|| < 2\eta\}$ of $\mathbb{E}X$ for some $\eta > 0$; $\sum_{|\alpha|=1} |D^\alpha H(\mathbb{E}X_1)| \neq 0$, and the first-order partial derivatives of H satisfy a Lipschitz condition of order $\kappa > 0$ on N_H.

C2 The moments of the transformed variables $\{X_n = f(Y_n)\}$ are finite up to some order $2 + \delta$, $\delta > 0$:

$$\mathbb{E}||X||^{2+\delta} < \infty. \tag{6.1}$$

C3 The dependence decays sufficiently fast:

$$\sum_{m=1}^{\infty} \alpha(m)^{\delta/(2+\delta)} < \infty, \tag{6.2}$$

where $\alpha(m)$ is the α-mixing coefficient at lag m for the time series of transformed variables $\{X_n = f(Y_n)\}$ (see e.g. [63]).

6.3.2 Block size: Theory and practice

Performance of block bootstrap estimators depend on the block size l, since the sampling distribution of an estimator depends on the joint distribution of $\{x_1, \cdots, x_N\}$. Hence, the block size l must grow to infinity with sample size N to eventually capture the dependence structure of the whole sequence $\{X_n\}_{n\geq 1}$. Also, consistency of block bootstrap estimators is usually linked to conditions on the block size of type **C0**, obligating l to grow at a slower rate than N.

In practice, having to choose an appropriate block size is one of the main drawbacks of block bootstrap methods, since little of practical use is known for this choice. From a theoretical point of view, there exist results on the optimal rate β, $l \sim N^\beta$, with $0 < \beta < 1$ typically depending on the type of estimator of interest. Also, for specific cases, expressions for the optimal (in terms of mean squared error) block size can be obtained (see [104] for an overview). Such results are, however, not very useful in most real-world cases: First, optimal rate results do not help in small sample situations since they only give the rate at which l should grow with N, but not the exact size for fixed N. Second, mean squared error expansions typically require much more structural knowledge of the process than is available in many practical situations. Some authors proposed nonparametric plug-in methods, based on bootstrap after bootstrap or subsampling, for estimation of a "good" block size [104]. Such methods are, however, computationally very expensive and thus often of limited use in practice. Therefore, in practice, the block size l is often chosen empirically, for instance by numerical inspection of the dependence properties of the data, or by analogy with estimation problems similar to the real-world problem, i.e. by resorting to synthetic model processes and simulation.

Chapter 7

Bootstrap Resampling and Estimation for Multifractal Analysis

Contents

Scale invariant processes are in general *non stationary* and possess intricate dependence structure and *power law type correlations*. In a statistical analysis and modeling perspective, these properties represent major difficulties. Indeed, any of these properties are decisive against the straightforward applicability of (standard) bootstrap methods: Stationarity is an essential requirement for (direct) application of any resampling technique, and the bootstrap fails for non stationary data. Power law type correlation is a significant difficulty for application of resampling techniques and is likely to violate the mixing conditions for bootstrap consistency. Although alternative techniques such as subsampling could theoretically cope with intricate dependence such as LRD (cf. e.g. [104] and references therein), they add extra practical difficulties due to additional renormalization constants, and are in general not second order correct[1]. Therefore, it is no good to attempt the construction of bootstrap samples from realizations of scale invariant processes.
In contrast, the *wavelet coefficients* of scaling processes with stationary increments are *stationary*[2]. The wavelet coefficients of fBm are only weakly correlated and expected to fulfill mixing conditions for bootstrap consistency [4, 6, 70, 166, 169]. For other scale invariant processes (non Gaussian H-sssi and MMC processes), practically no result was known at the time this chapter was started (see, a contrario, [77]). We therefore suppose that wavelet coefficients of such processes have similarly dependence as for fBm (as suggested for the covariance in [77]). Hence, the discrete wavelet coefficients of scaling processes are assumed to fulfill the prerequisites for bootstrap.
Therefore, we propose in this chapter the use of wavelet domain bootstrap for scaling and multifractal analysis. The contributions lie in the definition of two original block bootstrap resampling schemes: First, the definition of a wavelet domain time (1d) and space (2d) block bootstrap resampling, and second, the definition of time-scale (1d) and space-scale (2d) block bootstrap resampling in the time (space) scale plane of the discrete wavelet transform (DWT).
This work has been reported and used in [173, 174, 175, 179, 180, 182].

7.1 Multifractal Attribute Estimation and Bootstrap

7.1.1 Bootstrap in the wavelet domain

DWT domain bootstrap was first introduced in [137, 151]. The aim was the creation of surrogate data from time series and long range dependent time series. Percival [137] coined the term *wavestrapping* for this approach. Wavestrapping is based on the assumption that the difficult statistical context in the time domain can be circumvented by passing through the wavelet domain: The approach uses nonparametric i.i.d. bootstrap for the creation of bootstrap samples of the wavelet coefficients and the approximation coefficients, hence assuming independence of both the wavelet and the approximation coefficients. Then, bootstrap samples of the original time series are obtained by inverse wavelet transform. The use of wavestrapping has been reported in [40, 41, 183]. Modified versions have been adopted in [13] and [152], where a stationary bootstrap and a scale-by-scale block bootstrap are proposed for the creation of bootstrap resamples of

[1] In fact, their convergence is often worse than that of the Normal approximation.

[2] The restriction to processes with stationary increments can be further relaxed to processes with stationary N-th order increments, under condition that $N \leq N_\psi$, where N_ψ is the number of vanishing moments of the analyzing wavelet.

hierarchical processes and of inertial sensor noise time series, respectively. The initial goal – the creation of surrogate time series – remained the same.

For multifractal analysis purposes, wavestrapping needs to be modified to account for the following particularities: First, the key of multifractal analysis is that the (functional analysis) characterization of the process is based directly on its wavelet Leaders [89, 92]. Therefore, the creation of surrogate trajectories from bootstrap samples of wavelet coefficients by inverse wavelet transform is irrelevant. Also, it is – with the exception of certain special cases – unlikely that this would result in time series that reproduce the multifractal properties of the original process. What is more, estimations are based on quantities defined directly in the wavelet domain. Hence, bootstrap samples need to reproduce the statistical properties of the wavelet coefficients and Leaders. Therefore, the residual dependence of wavelet coefficients and the potential structural inter-scale dependence of Leaders should be taken into account, since they may have significant impact on the statistical properties of the estimators.

Therefore, we propose to base bootstrap estimation for multifractal attributes on DWT domain block bootstrap samples of wavelet coefficients and Leaders. The procedure will be defined in detail in Section 7.2.

7.1.2 Bootstrap and structure functions

Since the logarithm of the expectation of the structure functions $S(j,q)$ and the expectation of cumulants $C(j,p)$ theoretically display a linear dependence with scale j (cf. Eqs. (2.58) and (2.68)), model based bootstrap, i.e. bootstrap resampling of centered residuals of a linear model, could be applied. Such an approach would have the advantage of being computationally much less demanding than DWT domain bootstrap, since bootstrap samples would consist of only a few structure function residuals. It encounters, however, severe practical limitations and difficulties:

1. The information available in the sample of coefficients is reduced to only a few structure function values: Indeed, for a sample X of length N, there are only roughly $\log_2 N$ structure function and cumulant values and hence residuals. In practical multifractal analysis, the number of available (dyadic) scales exceeds very rarely the value 20 and is usually in the neighborhood of the value 10, but can be significantly smaller: For some applications, for instance for many biomedical signals and for images, typical sample sizes are of the order of $N = 128$, hence resulting in less than 6 available scales.

2. The structure functions and cumulants (and their linear model residuals) form heteroscedastic samples, since based on dyadic wavelet trees, for which the number of coefficients (involved in the calculation of structure functions) decreases with increasing scale j. The heteroscedasticity is (asymptotically) known only for Gaussian self similar processes [169], for which consequently bootstrap methods for heteroscedastic residuals (cf. [56]) could be used. In general, it is unknown. Although bootstrap procedures for unknown heteroscedasticity have been proposed [73, 120], this constitutes a major technical difficulty for model based bootstrap resampling of structure functions.

3. The definition of wavelet Leaders Eqs. (2.54) and (2.55) is likely to introduce additional dependence between wavelet Leaders at different scales and hence structure

function and cumulant linear model residuals. Together with heteroscedasticity and a small number of available scales, dependent residuals constitute a major practical difficulty for applying bootstrap to the linear regression model.

4. Finally, bootstrap resampling of the samples of structure functions and cumulants can not provide bootstrap estimations on the structure functions and cumulants, and can hence not be used for constructing, for instance, confidence limits for structure functions.

To conclude, theoretical and practical difficulties and limitations suggest that in the specific context of multifractal analysis, model based bootstrap for structure functions is impracticable.

7.2 Wavelet Domain Block Bootstrap Resampling and Estimation Procedures

In this section, we define the wavelet domain block bootstrap resampling and estimation procedures for multifractal analysis: First, we define how the bootstrap sample of wavelet Leaders $\mathcal{L}_X^* = \{L_X^*(j,k)\}$ ($\mathcal{L}_X^* = \{L_X^*(j,k_1,k_2)\}$) is obtained from the collection of Leaders $\mathcal{L}_X = \{L_X(j,k)\}$ ($\mathcal{L}_X = \{L_X(j,k_1,k_2)\}$) (Sections 7.2.1 and 7.2.2). Then, we describe how bootstrap estimates are obtained from the bootstrap sample of Leaders $\mathcal{L}_X^*$ (Section 7.2.3). For clarity reasons, we prefer to define wavelet domain block bootstrap resampling separately for the two practically most important cases, namely for 1d (signals) and for 2d (images) data. The definitions can be readily extended to the general n-dimensional data case without any theoretical or practical difficulties, as will become obvious from the definitions in 1d an 2d. The equivalent procedures for wavelet coefficient based estimation are obtained by replacing wavelet Leaders $\mathcal{L}_X = \{L_X(j,k)\}$ with wavelet coefficients $\mathcal{D}_X = \{d_X(j,k)\}$ in the definitions, mutatis mutandis.

The wavelet domain block bootstrap resampling for multifractal analysis should ideally ensure bootstrap samples with (asymptotically) the same dependence – both in *time* and in *scale* – as in the original sample of Leaders, by appropriate block construction. Since wavelet coefficients and Leaders at a given scale j form stationary processes, blocks can extend over a certain time length for capturing time dependence. However, coefficients or Leaders at different scales j and j' do not have the same marginal distributions, and consequently any model-free block resampling must not mix coefficients at different scales. The general solution for model-free block bootstrap in the wavelet domain is therefore given by blocks with certain time length that extend over a certain range of scales, and resampling that assures that blocks appear at the same range of scales in the bootstrap samples as they do in the original sample. The number of practically available scales in multifractal analysis being relatively small (typically, smaller than 20), we choose not to consider blocks extending over a fraction of the number of available of scales in this manuscript, and study only the two extreme cases: In the first one, which we call *time block* (T-B), blocks extend only in time direction and hence consist of coefficients from the same scale only. In the second one, which we term *time-scale block* (TS-B), blocks have a certain time length and extend over all available scales – they hence consist of all coefficients within a certain time interval.
Chronologically, since little is known on the inter-scale dependence of wavelet Leaders, the conceptually simpler and easier to implement T-B construction has been defined first

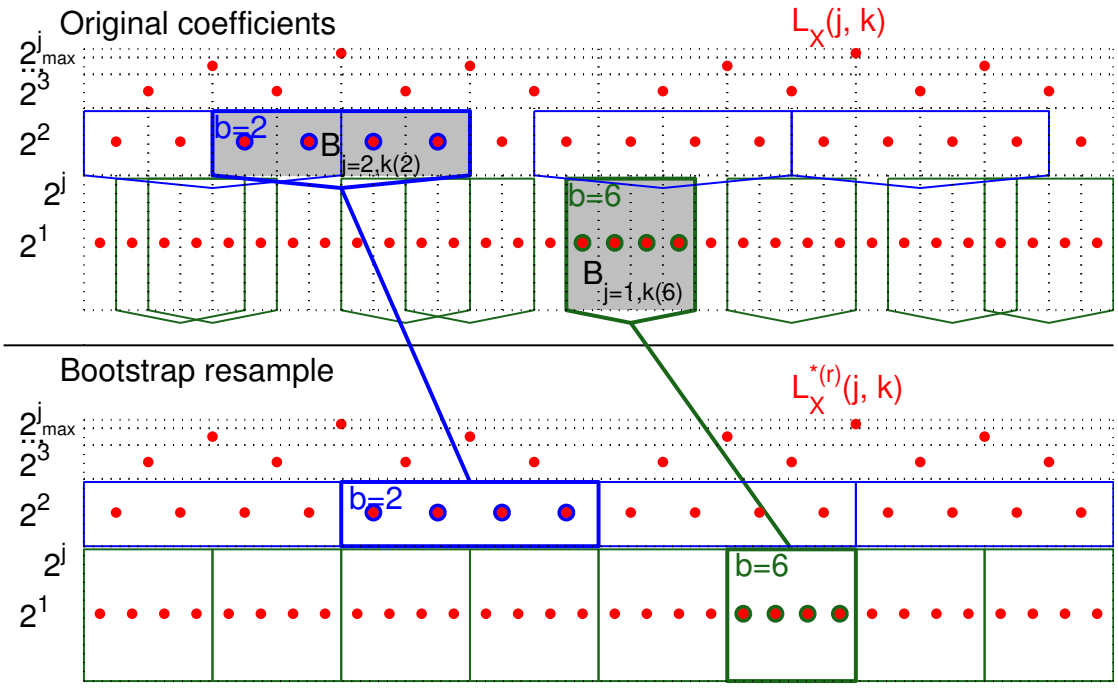


Figure 7.1: **Time block bootstrap.** Illustration of time block bootstrap resampling for $N = 2^6$ and $l = 4$. On top, the set of original coefficients ('•') in the time-scale plane. The coefficients marked by circles belong to the time blocks $\mathcal{B}_{j=2,k(2)}$ ('○') and $\mathcal{B}_{j=1,k(6)}$ ('○') at scales $j = 2$ and $j = 1$, respectively. The bootstrap resample $\{L^*_X(j,1), \cdots, L^*_X(j,n_j)\}$ at scale j consists of the coefficients that lie within the $m_j = \lceil \frac{n_j}{l} \rceil$ blocks $\mathcal{B}_{j,k(b)}$, $b = 1, \cdots, m_j$ drawn independently and with replacement from the collection of available blocks at scale j. The final bootstrap sample $\mathcal{L}^*_X = \{L^*_X(j,k)\}$ is composed of the bootstrap samples $\{L^*_X(j,1), \cdots, L^*_X(j,n_j)\}$ at all available scales j.

[173, 174, 179]. The practically more complex TS-B construction was defined afterwards from heuristic arguments based on the key properties of wavelet Leaders [182, 175].

7.2.1 Time block and space block bootstrap resampling

The time (or space) block bootstrap sample $\mathcal{L}^*_X = \{L^*_X(j,k)\}$ ($\mathcal{L}^*_X = \{L^*_X(j,k_1,k_2)\}$) is obtained by circular block bootstrap resampling of Leaders, at each scale j independently. Hence, the blocks are made up of a vector of l (a square of $l \times l$) consecutive (neighboring) circularized Leaders at a common scale j. Therefore, the time block (space block) bootstrap aims at constructing bootstrap samples that asymptotically reproduce residual correlation between Leaders of the same scale, while blinding out any potential inter-scale correlation.

7.2.1-a) Definition of time block bootstrap resampling: 1d

Let $\{L_X(j,1), \cdots, L_X(j,n_j)\}$ be the time series of the n_j Leaders at the scale j. Define $\{L_X(j,n_j+1), \cdots, L_X(j,n_j+l-1)\} = \{L_X(j,1), \cdots, L_X(j,l-1)\}$, and the circularized

sample of Leaders $\{L_X(j,1),\cdots,L_X(j,n_j+l-1)\}$. Let

$$\mathcal{B}_{j,k}=\{L_X(j,k),\cdots,L_X(j,k+l-1)\},\ 1\le k\le n_j$$

be the block of l consecutive Leaders, starting at $L_X(j,k)$. Suppose, for convenience[3], that $n_j/l=m_j\in N$. Then, the bootstrap sample $\{L_X^*(j,1),\cdots,L_X^*(j,n_j)\}$ at scale j is composed of the collection of $m_j\cdot l=n_j$ Leaders within a random sample of blocks $\{\mathcal{B}_{j,1}^*,\cdots,\mathcal{B}_{j,m_j}^*\}$, drawn independently and with replacement from the collection of available blocks $\{\mathcal{B}_{j,1},\cdots,\mathcal{B}_{j,n_j}\}$:

$$\{L_X^*(j,1),\cdots,L_X^*(j,n_j)\}=\{\mathcal{B}_{j,1}^*,\cdots,\mathcal{B}_{j,m_j}^*\}.$$

The final bootstrap sample $\mathcal{L}_X^*=\{L_X^*(j,k)\}$ is the union of the bootstrap samples $\{L_X^*(j,1),\cdots,L_X^*(j,n_j)\}$ for all available scales j:

$$\mathcal{L}_X^*=\{\{L_X^*(j=1,\cdot)\};\cdots;\{L_X^*(j=j_{max},\cdot)\}\}.$$

The time block bootstrap resampling procedure is illustrated in Fig. 7.1.

7.2.1-b) Definition of space block bootstrap resampling: 2d

Let $\{L_X(j,k_1,k_2)\}_{k_1,k_2=1}^{n_{1,j},n_{2,j}}$ be the collection of the $n_{1,j}\cdot n_{2,j}$ Leaders at the scale j, and consider the circularized collection $\{L_X(j,k_1 \mod n_{1,j},k_2 \mod n_{2,j})\}_{k_1,k_2=1}^{n_{1,j}+l-1,n_{2,j}+l-1}$. Let

$$\mathcal{B}_{j,k_1,k_2}=\{L_X(j,k_1',k_2')\}_{k_1'=k_1,k_2'=k_2}^{k_1+l-1,k_2+l-1},\ 1\le k_1\le n_{1,j},\ 1\le k_2\le n_{2,j}$$

be the square of $l\cdot l$ neighboring circularized Leaders. Suppose, for simplicity[4], that $n_{1,j}\cdot n_{2,j}/l^2=m_j\in N$. Then, the bootstrap sample $\{L_X^*(j,\cdot,\cdot)\}$ at scale j is composed of the collection of $l^2\cdot m_j=n_{1,j}\cdot n_{2,j}$ Leaders within a random sample of blocks $\{\mathcal{B}_{j,1,1}^*,\cdots,\mathcal{B}_{j,m_j,m_j}^*\}$, drawn independently and with replacement from the collection of available blocks $\{\mathcal{B}_{j,1,1},\cdots,\mathcal{B}_{j,n_{1,j},n_{2,j}}\}$:

$$\{L_X^*(j,1,1),\cdots,L_X^*(j,n_{1,j},n_{2,j})\}=\{\mathcal{B}_{j,1,1}^*,\cdots,\mathcal{B}_{j,m_j,m_j}^*\}.$$

The final bootstrap sample is the union of the bootstrap samples for all available scales j:

$$\mathcal{L}_X^*=\{\{L_X^*(j=1,\cdot,\cdot)\};\cdots;\{L_X^*(j=j_{max},\cdot,\cdot)\}\}.$$

7.2.2 Time-scale block and space-scale block bootstrap resampling

In contrast to the time blocks, which "extend" only over one scale and hence produce bootstrap samples for which any potential inter-scale dependence is whitened out, the time-scale (or space-scale) blocks extend over all available scales. Hence, they preserve inter-scale correlations for coefficients within a certain time neighborhood, determined by the block size. Such a block construction is expected to be particularly well adapted

[3] If $n_j/l\notin\mathbb{N}$, m_j is defined as $m_j=\lceil n_j/l\rceil$.

[4] If $n_{1,j}\cdot n_{2,j}/l^2\notin\mathbb{N}$, m_j is defined as $m_j=\lceil n_{1,j}\cdot n_{2,j}/l^2\rceil$.

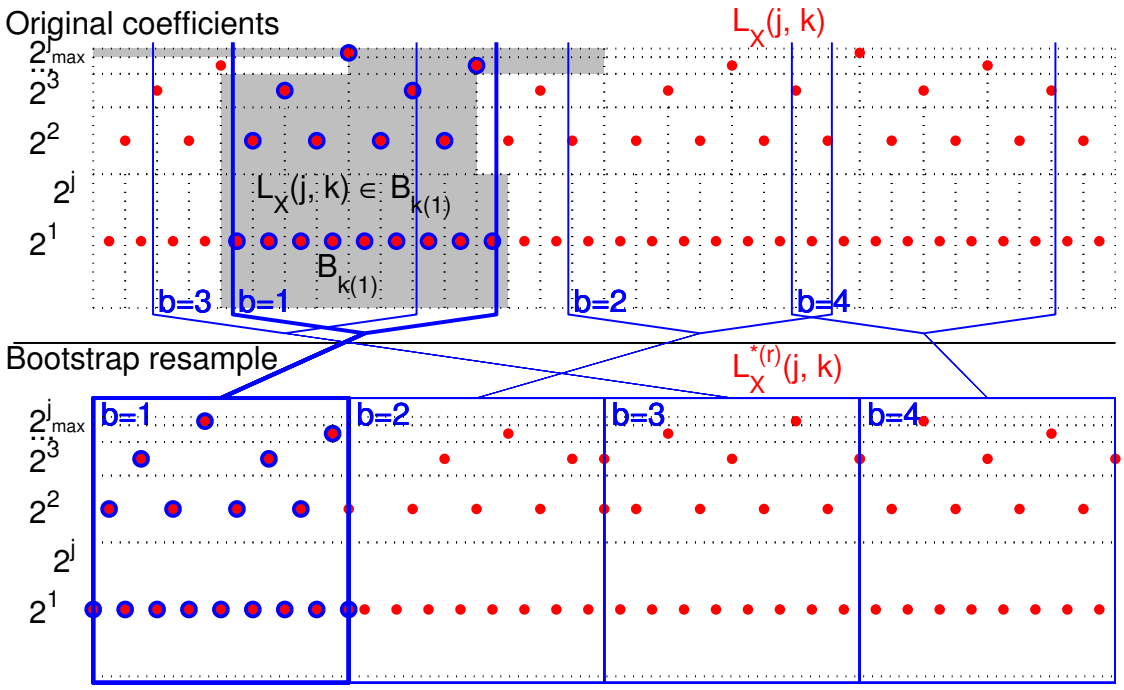


Figure 7.2: **Time-scale block bootstrap for signals.** Illustration of time-scale block bootstrap resampling for $N = 2^6$ and $l = 2^3$. On top, the set of original coefficients ('•') in the time-scale plane. The coefficients marked by circles ('○') belong to the time-scale block $\mathcal{B}_{k(1)}$. The bootstrap resample $L_X^{*(r)}$ (bottom) consists of the coefficients that lie within the $m = \lceil \frac{N}{2l} \rceil = 4$ time-scale blocks $\mathcal{B}_{k(b)}$, $b = 1, \cdots, m$, drawn independently and with replacement from all available overlapping circular time-scale blocks.

for wavelet Leaders: First, the definition of wavelet Leaders heuristically enlarges dependence for different time positions k and k' and in particular for different scales j and j' (cf. Sections 3.1 and 3.2), hence appropriateness of block constructions for capturing inter-scale dependence. Second, time-scale blocks ensures that bootstrapped Leaders retain one of their key properties for multifractal analysis, namely that they are hierarchical quantities and hence non-decreasing with increasing scale. An example for such a block for two-dimensional observations is illustrated in Fig. 7.3.

7.2.2-a) Definition of time-scale block bootstrap resampling: 1d

Let k denote the time index for the discrete time observation $X[k] = X(kT_s)$ at sampling period T_s and let k' be the time indices of the wavelet Leaders L_X of X at scale j'. Then, the time-scale block of Leaders $\mathcal{B}_k$ is defined as the collection of time-circularized Leaders in the stripe of time length $2l$ extending over all scales in the time-scale plane, centered at time k :

$$\mathcal{B}_k = \{L_X(j', k' \mod n_{j'}) : |k - k'2^{j'}| \leq l,\ 1 \leq j' \leq j_{max}\},\ 1 \leq k \leq N.$$

The bootstrap sample $\mathcal{L}_X^*$ is composed of the concatenation in time of the Leaders $L_X^*(j, k)$ within a random sample of $m = \lceil \frac{N}{2l} \rceil$ blocks of Leaders $\mathcal{B}^*$, drawn independently

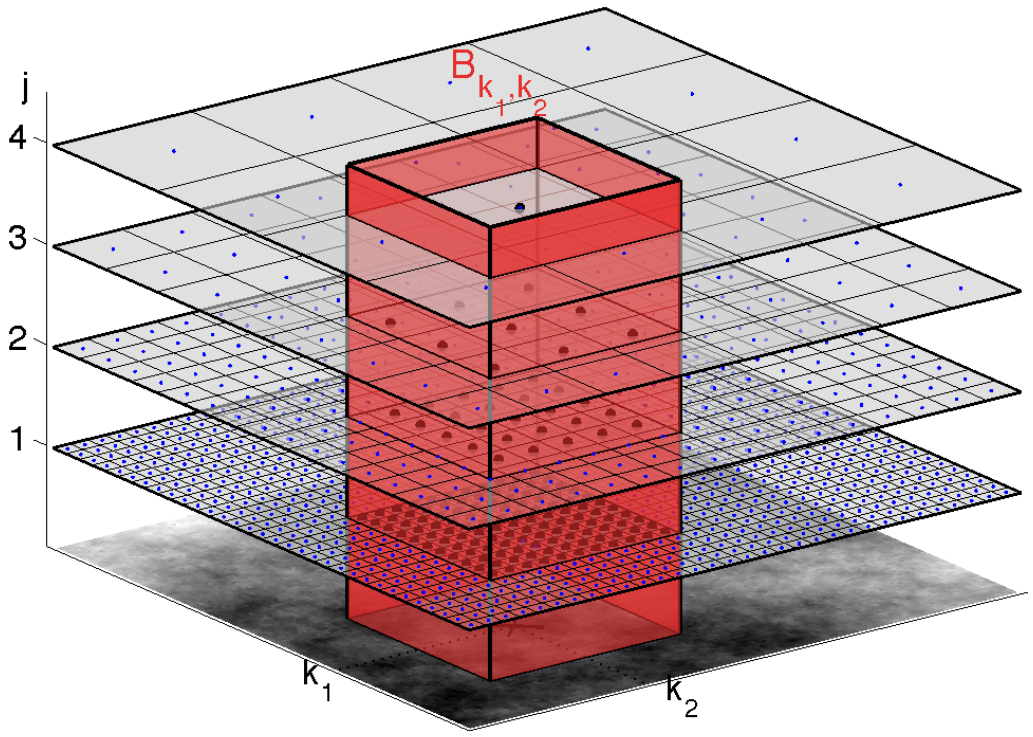


Figure 7.3: **Illustration of space-scale block construction.** The space-scale block of Leaders $\mathcal{B}_{k_1,k_2}$ at position (k_1, k_2) consists of the collection of Leaders (fat black dots '•') that are within a box with square base of base length of $2l$ pixel, centered at (k_1, k_2) and extending over all scales (red volume).

and with replacement from the available blocks $\{\mathcal{B}_1, \cdots, \mathcal{B}_N\}$:

$$\mathcal{L}_X^* = \{\mathcal{B}_1^*, \cdots, \mathcal{B}_m^*\}.$$

The definition of a time-scale block and of a bootstrap sample is illustrated in Fig. 7.2.

7.2.2-b) Definition of space-scale block bootstrap resampling: 2d

Suppose the image is of size $N_1 \times N_2$. The space-scale blocks of Leaders are defined as $3D$ boxes with a square base of side length of $2l$ pixels and extending over all scales. The Leaders lying within such a box form one space-scale block. The blocks are constructed overlapping and from space-circularized Leaders, i.e. from $L_X^c(j, k_1, k_2) = L_X(j, k_1 \bmod n_{1,j},\ k_2 \bmod n_{2,j})$. More precisely, the collection of Leaders $L_X(j', k_1', k_2')$ that form a space-scale block $\mathcal{B}_{k_1,k_2}$, of $2l \times 2l$ pixels square base and located at position (k_1, k_2), is given by:

$$\mathcal{B}_{k_1,k_2} = \{L_X(j', k_1' \bmod n_{1,j'}, k_2' \bmod n_{2,j'}) : \\ |k_1 - k_1' 2^{j'}| \leq l,\ |k_2 - k_2' 2^{j'}| \leq l,\ 1 \leq j' \leq j_{max}\}. \quad (7.1)$$

This definition of a space-scale block is illustrated in Fig. 7.3.

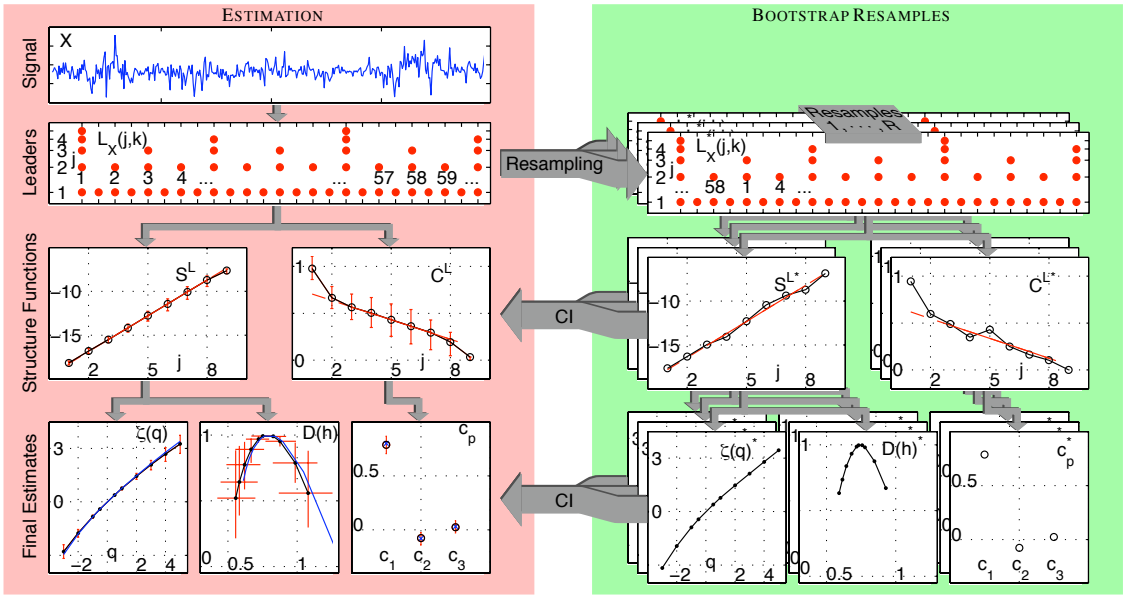


Figure 7.4: **Estimation and bootstrap estimation scheme.** From the data, the discrete wavelet coefficients and wavelet Leaders are computed (top left). From the wavelet Leaders, structure functions $S^L(j,q)$ and $\hat{C}^L(j,p)$ are calculated (middle left) and the corresponding estimates for the multifractal parameters $\theta \in \{\zeta(q), D(q), h(q), c_p\}$ are obtained (bottom left). From the sample of wavelet Leaders $\mathcal{L}_X$, a large number R of bootstrap samples of wavelet Leaders $\mathcal{L}_X^{*(r)}$, $r = 1, \cdots, R$ are generated by T-B or TS-B bootstrap: Each bootstrap sample is obtained by drawing at random, with replacement, time (scale) blocks of Leaders from the set of available overlapping, circularized blocks of Leaders (top right). From these R bootstrap samples, bootstrap structure function estimations $S^L(j,q)^*$ and $\hat{C}^L(j,p)^*$ are obtained (middle right). From the bootstrap structure function estimations $S^L(j,q)^*$ and $\hat{C}^L(j,p)^*$, bootstrap multifractal attribute estimates are calculated (bottom right). The empirical distributions of bootstrap estimates are used for e.g. confidences interval construction (center and bottom left, in red). (Results are obtained as means over 1000 realizations of CPM-MF-fBm of sample size (estimations in black, theoretical values in blue). Bootstrap percentile confidence limits (see Chapter 8) are estimated from $R = 399$ bootstrap samples per realization).

The bootstrap sample $\mathcal{L}_X^*$ is composed of the Leaders $L_X^*(j,k_1,k_2)$ within a random sample of $m = \lceil \frac{\sqrt{N_1 \cdot N_2}}{2l} \rceil$ blocks of Leaders $\mathcal{B}^*$, drawn independently and with replacement from the available space-scale blocks $\{\mathcal{B}_{1,1}, \cdots, \mathcal{B}_{N_1,N_2}\}$:

$$\mathcal{L}_X^* = \{\mathcal{B}_{1,1}^*; \cdots; \mathcal{B}_{m,m}^*\}.$$

The Leaders $L_X^*(j,k_1,k_2)$ are concatenated in space, such that each resampled Leader remains located at its original scale j.

7.2.3 Bootstrap estimation and bootstrap inference

Bootstrap inference for quantities such as confidence intervals, statistical tests, bias or mean squared error estimates etc., characterizing statistical properties of the estimates

and structure functions $\hat{\theta} \in \{\hat{\zeta}(q), \hat{c}_p, \hat{D}(q), \hat{h}(q), S(j,q), \hat{C}(j,p)\}$, are based on the empirical distributions of the bootstrap estimates, $\hat{\theta}^{*(r)} \in \{\hat{\zeta}(q)^{*(r)}, \hat{c}_p^{*(r)}, \hat{D}(q)^{*(r)}, \hat{h}(q)^{*(r)}, S(j,q)^{*(r)}, \hat{C}(j,p)^{*(r)}\}$, $r = 1, \cdots, R$. The bootstrap estimates $\hat{\theta}^{*(r)}$ are obtained by applying the estimation procedures defined in Section 2.6 to R bootstrap samples of Leaders $\mathcal{L}_X^{*(r)}$, $r = 1, \cdots, R$.
The final multifractal attribute estimation and bootstrap estimation procedure can hence be summarized as follows:

1. First, calculate the Leaders $\mathcal{L}_X$ from the observation X, and apply the estimation procedures defined in Section 2.6 to obtain structure function and multifractal attribute estimates $\hat{\theta} \in \{\hat{\zeta}(q), \hat{c}_p, \hat{D}(q), \hat{h}(q), S(j,q), \hat{C}(j,p)\}$.

2. Second, create a large number R of bootstrap samples of Leaders $\mathcal{L}_X^{*(r)}$, $r = 1, \cdots, R$, following the bootstrap resampling procedures described above (Sections 7.2.1 and 7.2.2).

3. Third, apply the estimation procedures defined in Section 2.6 to each of these bootstrap samples of Leaders $\mathcal{L}_X^{*(r)}$ to obtain a collection of structure function and multifractal attribute bootstrap estimates $\hat{\theta}^{*(r)}$, $r = 1, \cdots, R$.

4. Finally, use the relation between estimates and the empirical distribution of the bootstrap estimates:
$$\hat{\theta}^* - \hat{\theta}$$
to approximate the relation between the parameters and the sampling distribution of the estimates:
$$\hat{\theta} - \theta$$
and for bootstrap inference, i.e. for the construction of confidence intervals or statistical tests, or for the estimation of bias or mean squared error etc..

The outline of this estimation, bootstrap resampling and bootstrap estimation procedure is illustrated in Fig. 7.4 for a 1d signal. The precise definition of confidence intervals and statistical tests based on this bootstrap approximation of the sampling distribution of the structure function and multifractal attribute estimators will be considered in Chapters 8, 9 and 10.

7.2.4 Double bootstrap

Bootstrap can be used to estimate the accuracy of the bootstrap estimate $\hat{\theta}^*$ itself. This "bootstrapping the bootstrap", i.e. application of bootstrap to the bootstrap estimation itself, is commonly referred to as *double bootstrap* (see e.g. [35, 36, 56, 66]).
The double bootstrap estimates $\hat{\theta}^{**}$ of the bootstrap estimate $\hat{\theta}^{*(r)}$ are obtained by applying the bootstrap procedure[5] in Section 7.2.3 to the bootstrap sample $\mathcal{L}_X^{*(b)}$: First, apply the resampling procedures in Section 7.2.1 or Section 7.2.2 S times to the bootstrap sample of Leaders $\mathcal{L}_X^{*(r)}$ to obtain S double bootstrap samples of Leaders $\mathcal{L}_X^{**(r,s)}$, $s = 1, \cdots, S$. Second, apply the estimation procedures of Section 2.6 to each of these double bootstrap samples of Leaders $\mathcal{L}_X^{**(r,s)}$ to obtain a collection of structure function and

[5] The first step can obviously be skipped, since it corresponds to the second step of the initial bootstrap estimation and hence is already calculated.

multifractal attribute double bootstrap estimates $\hat{\theta}^{**(r,s)}$, $s = 1, \cdots, S$.
One example in which double bootstrap is used are pivoted (or studentized) statistics (e.g. [35, 36]):

$$S_\theta = \frac{\hat{\theta} - \theta}{\hat{\sigma}^*_\theta}, \tag{7.2}$$

where $\hat{\sigma}^*_\theta = \widehat{\mathrm{Std}^*}\hat{\theta}^*$ is a bootstrap estimation of the standard deviation of $\hat{\theta}$. The idea behind this normalization is the elimination of free parameters in the statistics: For instance, if $\hat{\theta}$ is Gaussian, S_θ has no free parameter. The bootstrap version of this statistic is given by:

$$S^{*(r)}_\theta = \frac{\hat{\theta}^{*(r)} - \hat{\theta}}{\hat{\sigma}^{**(r)}_\theta}, \; r = 1, \cdots, R, \tag{7.3}$$

where $\hat{\sigma}^{**(r)}_\theta = \widehat{\mathrm{Std}^{**}}\hat{\theta}^{**(r,\cdot)}$ is the double bootstrap standard deviation estimate of $\hat{\theta}^{*(r)}$.
The double bootstrap can also be used for coverage error correction for bootstrap confidence intervals or size error correction for statistical tests (cf. adjusted bootstrap confidence intervals and test acceptance regions in Sections 8.1.2 and 9.1.3, respectively).

7.3 Statistical Aspects of Bootstrap Resampling for Multifractal Analysis

This section discusses the conditions for consistency of the circular block bootstrap (cf. Section 6.3.1) for the time block and time-scale bootstrap sampling and estimation that we defined in the previous section: Applicability of the smooth function model conditions **C1** to structure function, cumulant, and multifractal attribute estimation; Condition of finite moments **C2**; Theoretical and practical choice of block size. The study of the dependence conditions **C3** is postponed to Chapter 11.

7.3.1 C1: Smooth function model

7.3.1-a) Structure functions

It is easy to see that the logarithm of structure functions, $Y(j,q) = \log_2\left(\mathbb{E}\left[L_X(j,\cdot)^q\right]\right)$, falls into the smooth function model class by identifying $f(y) = y^q$ and $H(x) = \log(x)$, hence $Y(j,q) = H(\mathbb{E}f(L_X(j,\cdot)))$ and $\hat{Y}(j,q) = H(\frac{1}{n_j}\sum_{k=1}^{n_j} f(L_X(j,k)))$ (cf. Section 6.3.1). Also, they satisfy the additional conditions **C1**: The (n-th order) derivatives of $H(x)$ exist on $\mathbb{R}^+$, $\frac{d^n}{dx^n}H(x) = (n-1)! \cdot (-1)^{(n-1)} \cdot x^{-n}$. Since for the scaling processes considered here, $0 < \mathbb{E}L_X(j,\cdot)^q$ at least for some range of q centered around $q = 0$, the derivatives of $H(x)$ do exist around $\mathbb{E}x = \mathbb{E}f(y) = \mathbb{E}L_X(j,\cdot)^q$.

7.3.1-b) Cumulants

The cumulants $\hat{C}(j,p)$ (cf. Eq. (2.72)) of $Y = \ln L_X(j,\cdot)$ are smooth functions of the sample moments of Y and therefore also belong to the smooth function model class. Let us denote them by C^Y_p for convenience and let us, for instance, express the first three of

them in terms of the raw moments of Y by identifying $f_k(Y) = Y^k$:

$$\begin{aligned} C_1^Y &= \mathbb{E}f_1(Y) \\ C_2^Y &= \mathbb{E}f_2(Y) - (\mathbb{E}f_1(Y))^2 \\ C_3^Y &= 2(\mathbb{E}f_1(Y))^3 - 3\mathbb{E}f_1(Y)\mathbb{E}f_2(Y) + \mathbb{E}f_3(Y). \end{aligned}$$

Hence, the functions $H(\cdot)$ for the first three cumulants are, by setting $z_k = \mathbb{E}f_k(Y)$:

$$\begin{aligned} H_{C_1^Y}(z_1) &= z_1 \\ H_{C_2^Y}(z_1, z_2) &= z_2 - z_1^2 \\ H_{C_3^Y}(z_1, z_2, z_3) &= 2z_1^3 - 3z_1z_2 + z_3. \end{aligned}$$

Also, the $H_{C_p^Y}$ satisfy the conditions **C1**, since they are twice differentiable w.r.t. z_i, and since the partial derivatives w.r.t. z_p are in general non-zero (cf. Eq. (2.72)):

$$\frac{d}{dz_p}H_{C_p^Y}(z_1, \cdots, z_p) = \frac{d}{dz_p}z_p - \sum_{k=1}^{p-1}\binom{p-1}{k-1}H_{C_k^Y}(z_1, \cdots, z_k)z_{p-k} = \frac{d}{dz_p}z_p = 1.$$

7.3.1-c) Linear fits: Multifractal attribute estimations

Finally, notice that the multifractal attribute estimation for $\zeta(q)$ and c_p (Eqs. (2.65) and (2.71)) are within the smooth function model and satisfy **C1**, since they are deterministic linear combinations of the smooth functions of means $Y(j,q) = \log_2 S(j,q)$ and $\hat{C}(j,p)$.

Conclusions. To conclude, (log of) structure functions, cumulants and multifractal attributes are within the smooth function model and satisfy the additional conditions **C1**.

7.3.2 C2: Finiteness of moments

7.3.2-a) Structure functions

The moment condition, $\mathbb{E}||X||^{2+\delta} < \infty$, for structure functions $S^L(j,q)$ translates to the finite moment condition for the wavelet Leaders:

$$\mathbb{E}L_X(j,\cdot)^{q(2+\delta)} < \infty,\ 0 < \delta. \tag{7.4}$$

Self-similar processes with stationary increments. For fv H-sssi processes, the moments of all orders exist. Hence, if $\delta < \infty$, the condition Eq. (7.4) is fulfilled for:

$$-\infty < q < \infty.$$

Multifractal multiplicative cascade processes. For MMC processes, the range for which moments of Leaders theoretically exist is finite, $-\infty < q_c^- < 0 < q_c^+ < +\infty$ (cf. Section 2.7.1-d)). Therefore, the condition Eq. (7.4) is only fulfilled for the range of statistical orders:

$$\frac{q_c^-}{2+\delta} < q < \frac{q_c^+}{2+\delta}.$$

MMC processes and linearization effect. Notice that in practice, the wavelet Leaders of MMC processes behave as if their moments would exist only in a narrower range of

statistical orders $q \in [q_*^-, q_*^+] \subset [q_c^-, q_c^+]$ (linearization effect, Section 2.8, Eqs. (2.129) and (2.130), and Section 5.2). This phenomenon remains a controversial theoretical issue and has in practice been subject mainly to empirical studies. Therefore, it is at present not clear in how far it has impact on the asymptotic consistency of the block bootstrap for structure functions, i.e., if the range of statistical orders q for fulfilling the finite moment condition **C2** has to be narrowed to:

$$\frac{q_c^-}{2+\delta} < \frac{q_*^-}{2+\delta} < q < \frac{q_*^+}{2+\delta} < \frac{q_c^+}{2+\delta}.$$

Empirical elements for answering this question are stated in Section 8.3.6.

7.3.2-b) Cumulants

For cumulants $C(j,p)$, the finite moment condition **C2**, $\mathbb{E}||X||^{2+\delta} < \infty$ translates to a finite moment condition for the log of the wavelet Leaders:

$$\mathbb{E}[(\ln L_X)^{p(2+\delta)}] < \infty,\ 0 < \delta. \tag{7.5}$$

From Eq. (2.70), it follows that equivalently, the derivative $d^{p(2+\delta)}/dq^{p(2+\delta)}\zeta(q)$ of $\zeta(q)$ have to exist in the neighborhood of $q = 0$. Therefore, for most practically relevant processes and for those considered in this manuscript, this condition is fulfilled[6].

7.3.2-c) Linear fits: Multifractal attribute estimations

Finally, the moment conditions for structure functions $S(j,q)$ and cumulants $C(j,p)$ directly translate to scaling exponents $\zeta(q)$ and log-cumulants c_p, consisting of linear combinations of $S(j,q)$ and $C(j,p)$, respectively.

Conclusions. To conclude, (the logarithm of) structure functions, cumulants and multifractal attributes satisfy the moment condition **C2** within some neighborhood of $q = 0$ and for some $p \geq 1$ for most practically relevant process classes.

7.3.3 C3: Dependence structure

The dependence condition **C3** is not fulfilled for LRD processes [103]. As stated above, this is precisely the underlying motivation for placing estimation and bootstrap procedures in the wavelet domain. No results on the dependence structure of wavelet coefficients for MMC processes were known when the work presented in this chapter was started. Therefore, we study the dependence structure of wavelet coefficients and its implications for bootstrap in Chapter 11.

7.3.4 Block size

The choice of the block size is critical for the performance of block bootstrap methods. For certain simple estimators and for situations where sufficient structural knowledge on the data is available, mean squared error expansions can be obtained and exploited for

[6] Note that this is not valid in general for all processes and excludes, for instance, log-stable cascades [155] and α-stable motion [153] (see Section 2.6.2). Such processes are not considered in this manuscript.

choosing the block size. For asymptotic consistency of the block bootstrap the block size l has to go to infinity as the sample size N goes to infinity, slower than N, at a rate N^β, $0 < \beta < 1$ (Condition **C0**). The optimal rates β can differ for different types of estimators. (cf. Section 6.3.2 and references therein).

The block size choice for block bootstrap estimation for multifractal analysis is confronted with the following difficulties:

- Too little is known theoretically on the statistics and dependence of wavelet Leaders for choosing a (close to mean squared error) optimal block size (cf. Section 3.2). Also, in practice, such theoretical results would still involve knowledge on the precise process properties (for instance, H-sssi or MMC process, and parameter values), which are in general unavailable in most practical situations.

- Practical multifractal analysis demands for estimates involving many statistical orders q. Also, it involves different types of estimates, such as (moment based) structure functions, cumulants, and linear fits. Theoretically, the optimal block size is different for all of these estimates[7].

- Finally, for the time-scale (space-scale) block bootstrap, it is not immediately clear how the time length (space occupation) of the blocks relates to an optimal block size, since blocks extend as well over scales and do not contain the same number of discrete coefficients at different scales.

Also, (bootstrap) estimation in multifractal analysis meets the following particularities that have to be considered for block size choice in practice:

- The sample size in multifractal analysis practically spans only roughly three orders of magnitude, since a relevant multifractal analysis of signals (images) smaller than roughly $N \approx 2^9$ ($N \times N \approx 2^9 \times 2^9$) is difficult due to the very limited number of scales available for linear regressions (due to border effects, $j_{max} \approx 5$, depending on N_ψ chosen), and since signals (images) larger than $N \approx 2^{22}$ ($N \times N \approx 2^{11} \times 2^{11}$) are rarely encountered in practical multifractal analysis. Therefore, whereas it remains important from a theoretical point of view, choosing a rate $\beta < 1$ for block size evolution with sample size N that is optimal is not a crucial issue in practical multifractal analysis, and it is preferable from a practical point of view to concentrate on the small sample properties.

- The practical usability of the bootstrap estimation procedures for multifractal analysis would be rather limited due to computation time if for each estimate a different block size was used, making a different bootstrap resampling for each estimate necessary. Since the number of estimates for practical multifractal characterization can easily attain the order of 100 or more, depending on how many different values of q have to be considered, it is usually preferable to reuse the same bootstrap samples and hence block size for all estimates.

For these reasons, the block size for block bootstrap estimation for multifractal analysis is fixed empirically by numerical simulation on synthetic H-sssi and MMC processes, and

[7] If, for instance, we consider the in practice relatively small number of 10 values for q, this sums up to a total of thirty estimates $\zeta(q)$, $D(q)$, $h(q)$ plus the structure functions, and for each of them a (different) optimal block size would have to be chosen.

based on heuristic arguments. Also, the same block size is used for any of the multifractal attribute estimates. The block size choice is described in the next paragraphs. It gives good empirical results for structure functions and multifractal attribute estimates for Gaussian H-sssi processes, and for multifractal attribute estimates for non Gaussian H-sssi and MMC processes. Numerical results for confidence intervals and hypothesis tests will be given in Chapter 8 and Chapters 9 and 10, respectively. Some numerical results, similar to those underlying the empirical block size choice stated in the next paragraphs, will be presented in Chapter 11 for validating their robustness, in the context of related issues.

7.3.4-a) Time block and space block bootstrap

Numerical simulations for fv H-sssi and MMC processes show that the residual correlation of wavelet coefficients at a given scale j is approximately of the order of magnitude of the time support of the wavelet chosen for analysis. Hence, heuristically, blocks for small samples should be roughly the size of the time (space) support of the analysing wavelet, which is $2 \cdot N_\psi$. The block size of the time block and space block bootstrap is set to:

$$l = 2 \cdot N_\psi. \tag{7.6}$$

For scales at which the number of coefficients is practically too small for block bootstrap, $n_j < 2l$, the block size is divided by 2 until $n_j \geq 2l$.

7.3.4-b) Time-scale block and space-scale block bootstrap

The philosophy of the time-scale (space-scale) blocks is to create bootstrap samples that reproduce the across-scales behavior of wavelet Leaders (cf. Sections 2.5 and 7.2.2). Hence, the time (space) support of the time (space) scale blocks should be sufficiently large to ensure that it extends over all available scales or over all scales involved in linear regressions for multifractal attribute estimation: $l > 2^{j_2}$. Numerical simulations show that a good empirical choice is:

$$l \approx \lceil C \cdot N^\beta \rceil \tag{7.7}$$

with $\beta = \frac{2}{3}$ and $C \approx 0.6$, which corresponds to $l = 2^6$ for $N = 2^{10}$ and $l = 2^{10}$ for $N = 2^{16}$.

7.4 Conclusions and Perspectives

In this chapter, we defined wavelet domain time block bootstrap resampling and time-scale block bootstrap resampling (and the equivalent procedures in 2d). From a conception point of view, the TS-B bootstrap heuristically has an edge over the T-B bootstrap, since the blocks are constructed in such a way that they capture time-scale dependence within a certain time neighborhood, and over all scales. Therefore, it seems particularly well adapted for wavelet Leader based estimation. It is, however, practically more difficult to implement than the T-B bootstrap, especially for higher dimensional data, due to the border effects of the wavelet transform and related issues[8].
We have also discussed the possibility of a more general block construction, consisting of

[8] It is, for instance, not immediately obvious how to ensure a n out of n bootstrap procedure for coarse scales where only few coefficients are left.

blocks with certain time length spanning only over a certain range of scales. Though potentially interesting, such constructions add additional practical difficulties as compared to the TS-B and are not considered in this manuscript.

Chapter 8

Bootstrap Confidence Intervals

Contents

In a large number of applications, bootstrap has been used for *confidence interval* design. Bootstrap confidence intervals were found useful in situations and for data of diverse nature, typically too involved for analytic constructions. This has, for instance, been reported in [56, 190, 191, 192, 195, 196]. For multifractal analysis, no confidence intervals are available in practice. This is mainly due to the involved statistical properties of multifractal processes, which are largely responsible for the lack of statistical performance results for parameter estimation procedures. The only exception is the specific case of Gaussian H-sssi processes, where asymptotic constructions for the self-similarity parameter have been obtained. This absence of confidence intervals constitutes a major limitation for the practical multifractal analysis of real-world data and in applications, where they are often appreciated as being at least as important as the estimates themselves. It is hence natural to consider the nonparametric bootstrap resampling and estimation procedures for multifractal analysis that we proposed in the previous Chapter 7 for the construction of confidence intervals for multifractal attributes. Therefore, the aim of the present chapter is to define bootstrap confidence intervals for multifractal attributes, and to assess their practical reliability, precision and applicability.
The work presented in this chapter has been reported in [173, 179, 180, 182].

8.1 Bootstrap Confidence Intervals for Multifractal Analysis

A level (or coverage) $(1-\alpha)$ *confidence interval* (CI) $[\hat{\theta}_l, \hat{\theta}_r]$ for the real-valued population parameter $\theta \in \{\zeta(q), D(q), h(q), c_p\}$ is constructed by finding values $\hat{\theta}_l, \hat{\theta}_r$ for which:

$$\mathsf{Pr}\left(\hat{\theta}_l \leq \theta \leq \hat{\theta}_r\right) = 1-\alpha. \tag{8.1}$$

Hence, confidence interval construction implies knowledge of the distribution of $\hat{\theta}-\theta$ in order to be able to find the values $\hat{\theta}_l, \hat{\theta}_r$ that verify Eq. (8.1). This distribution of $\hat{\theta}-\theta$ is, however, unknown. Therefore, we approximate it by the distribution of the bootstrap version $\hat{\theta}^*-\hat{\theta}$, which is obtained using the bootstrap estimation procedure defined in Section 7.2.3, based on the multifractal analysis specific block bootstrap resampling defined in Section 7.2.1 and in Section 7.2.2. This bootstrap distribution approximation is then used for the construction of bootstrap confidence intervals.
It is clear from Eq. (8.1) that there exists in general an infinity of level $(1-\alpha)$ confidence interval constructions if the distribution of $\hat{\theta}$ is continuous, and there is a large number of possible nonparametric bootstrap CI definitions proposed in the literature [56, 61, 80, 195]. They can be cast into two classes: First, simple bootstrap CI are directly based on the distribution of $\hat{\theta}^*-\hat{\theta}$. They are conceptually simple and easy to implement and computationally inexpensive since they are based on one single bootstrap layer and hence also only on a "first order" bootstrap distribution approximation, which may in practice result in coverage error [56, 195]. The second class, called double bootstrap CI, uses a second layer of bootstrap estimation in order to define more accurate bootstrap distribution approximations, or to correct for coverage error in simple bootstrap CI. Commonly used definitions involve pivoting or studentizing of the bootstrap distribution, bootstrap estimation of a variance stabilizing transform, or double bootstrap based bias correction of the simple bootstrap CI limits (cf. e.g. [35, 56, 66, 195]). Double bootstrap CI therefore have the potential of being more precise than simple bootstrap CI. However, they also substantially increase computational load and conceptual complexity

due to the additional "bootstrapping the bootstrap" layer.
Among the large number of classical nonparametric bootstrap CI definitions proposed in the literature, we select three simple bootstrap CI (Normal, basic and percentile CI) and three double bootstrap CI (studentized, adjusted basic and adjusted percentile CI) as being representative for each respective class. These CI are defined in the remainder of this section.

8.1.1 Simple bootstrap confidence intervals

8.1.1-a) Normal (asymptotic) bootstrap confidence interval

The simple $(1-\alpha)$ Normal (or asymptotic) bootstrap confidence interval is based on the assumption that $\hat{\theta}$ has an (approximately) Normal distribution centered around θ. It uses the bootstrap standard deviation estimates $\hat{\sigma}^*_\theta = \widehat{\mathrm{Std}}^*\hat{\theta}^*$ and is defined as:

$$\mathrm{CI}^{nor}_{\theta,(1-\alpha)} = \left[\hat{\theta} - q_{(\frac{\alpha}{2})}\hat{\sigma}^*, \hat{\theta} + q_{(\frac{\alpha}{2})}\hat{\sigma}^*\right], \tag{8.2}$$

where q_α is the α quantile of the standard Normal distribution.

8.1.1-b) Basic bootstrap confidence interval

The $(1-\alpha)$ basic bootstrap CI is derived from the definition of the equi-tailed $(1-\alpha)$ limits for $\hat{\theta}-\theta$,

$$\mathrm{Pr}(\hat{\theta}-\theta \leq t_{\frac{\alpha}{2}}) = \mathrm{Pr}(\hat{\theta}-\theta \geq t_{1-\frac{\alpha}{2}}) = \frac{\alpha}{2}, \tag{8.3}$$

where t_α is the α-quantile of $\hat{\theta}-\theta$. Rewriting Eq. (8.3) leads to:

$$\hat{\theta} - t_{1-\frac{\alpha}{2}} \leq \theta \leq \hat{\theta} - t_{\frac{\alpha}{2}}, \tag{8.4}$$

and, by plugging in the bootstrap estimates[1] $\hat{t}_\alpha = \hat{\theta}^*_\alpha - \hat{\theta}$, this becomes:

$$\mathrm{CI}^{bas}_{\theta,(1-\alpha)} = \left[2\hat{\theta} - \hat{\theta}^*_{1-\frac{\alpha}{2}}, 2\hat{\theta} - \hat{\theta}^*_{\frac{\alpha}{2}}\right], \tag{8.5}$$

where $\hat{\theta}^*_\alpha$ is the α quantile of the empirical distribution of the bootstrap estimates $\hat{\theta}^*$.

8.1.1-c) Percentile bootstrap confidence interval

The $(1-\alpha)$ percentile bootstrap confidence interval is defined directly from the α quantiles $\hat{\theta}^*_\alpha$ of the empirical distribution of the bootstrap estimates $\hat{\theta}^*$:

$$\mathrm{CI}^{per}_{\theta,(1-\alpha)} = \left[\hat{\theta}^*_{\frac{\alpha}{2}}, \hat{\theta}^*_{1-\frac{\alpha}{2}}\right]. \tag{8.6}$$

It can be derived from Eq. (8.3) by assuming that there exists a (unknown) variable transformation for which the transformed distribution of $\hat{\theta}-\theta$ is symmetric, without explicit recourse to this transformation [56].

[1] Note that Eq. (8.2) is obtained from Eq. (8.4) similarly, by plugging in the Normal assumption bootstrap estimates, i.e., the quantiles of $\mathcal{N}(0,\hat{\sigma}^*)$, for $\hat{t}_\alpha$.

8.1.2 Double bootstrap confidence intervals

8.1.2-a) Studentized bootstrap confidence interval

The studentized bootstrap CI is a basic bootstrap CI for the studentized random variable $S_\theta = \frac{\hat{\theta}-\theta}{\hat{\sigma}^*}$ (Eq. (7.2)), where $\hat{\sigma}^* = \widehat{\text{Std}}^*\hat{\theta}^*$ is the bootstrap estimate of the standard deviation of $\hat{\theta}$. Hence, it is based on the probability implication of Eq. (8.3) for the random variable S_θ rather than $\hat{\theta}-\theta$. The studentized bootstrap CI is defined as:

$$\text{CI}^{stu}_{\theta,(1-\alpha)} = \left[\hat{\theta} - \hat{\sigma}^* S^*_{\theta,(1-\frac{\alpha}{2})}, \hat{\theta} - \hat{\sigma}^* S^*_{\theta,(\frac{\alpha}{2})}\right], \tag{8.7}$$

where $\hat{\sigma}^* S^*_{\theta,(\alpha)}$ is the α quantile of the empirical distribution of the bootstrap version of S_θ, $S^*_\theta = \frac{\hat{\theta}^*-\hat{\theta}}{\hat{\sigma}^{**}}$ (Eq. (7.3)). Studentizing attempts to produce a random variable S_θ whose distribution no longer depends on unknowns (or at least depends on less unknowns), through (empirical) normalization by the standard deviation. If accurate standard deviation estimations are available, it is expected that the distribution of this random variable is better reproduced by the distribution of its bootstrap version [56].

8.1.2-b) Adjusted basic bootstrap confidence interval

The basic bootstrap confidence limits Eq. (8.5) are in general not exact in the sense they do not exactly reproduce the probability implication Eq. (8.3), since the distribution of $\hat{\theta}^* - \hat{\theta}$ is only an approximation of the distribution of $\hat{\theta} - \theta$. For instance, for the upper basic limit:

$$\text{Pr}(\hat{\theta}^*_{\frac{\alpha}{2}} \leq 2\hat{\theta} - \theta) = \frac{\alpha'}{2} \neq \frac{\alpha}{2} \tag{8.8}$$

in practice, and similarly for the lower limit. The idea behind the adjusted bootstrap confidence interval is to use a second layer of bootstrap to "adjust" the upper (respectively, lower) limit such that it approximately fulfills the probability implication Eq. (8.3), i.e. to find a $\gamma(\alpha)$ such that:

$$\text{Pr}(\hat{\theta}^*_{\gamma(\frac{\alpha}{2})} \leq 2\hat{\theta} - \theta) = \frac{\alpha}{2}. \tag{8.9}$$

This is achieved by using the double bootstrap version:

$$p(r) = \text{Pr}^*(\hat{\theta}^{**(r,\cdot)} \leq 2\hat{\theta}^{*(r)} - \hat{\theta}),\ r = 1, \cdots, R, \tag{8.10}$$

and setting:

$$\gamma\left(\frac{\alpha}{2}\right) = p_{\frac{\alpha}{2}}. \tag{8.11}$$

where p_α is the α quantile of the empirical distribution of p [56]. Equivalently, for the lower adjusted basic limit this gives:

$$\gamma\left(1 - \frac{\alpha}{2}\right) = p_{1-\frac{\alpha}{2}}, \tag{8.12}$$

and hence:

$$\text{CI}^{adjb}_{\theta,(1-\alpha)} = \left[2\hat{\theta} - \hat{\theta}^*_{\gamma(1-\frac{\alpha}{2})}, 2\hat{\theta} - \hat{\theta}^*_{\gamma(\frac{\alpha}{2})}\right]. \tag{8.13}$$

8.1.2-c) Adjusted percentile bootstrap confidence interval

The adjusted percentile confidence interval uses double bootstrap correction of the upper and lower limits of the percentile bootstrap confidence interval Eq. (8.6), in the same spirit as the adjusted basic confidence interval Eq. (8.13) does for the basic bootstrap confidence limits Eq. (8.5), by double bootstrap estimation of γ:

$$\mathsf{Pr}(\theta \leq \hat{\theta}^*_{\gamma(\frac{\alpha}{2})}) = \mathsf{Pr}(\hat{\theta}^*_{\gamma(1-\frac{\alpha}{2})} \geq \theta) = \alpha. \tag{8.14}$$

using the double bootstrap version:

$$p(r) = \mathsf{Pr}^*(\hat{\theta}^{**(r,\cdot)} \geq \hat{\theta}),\ r = 1, \cdots, R. \tag{8.15}$$

Then,

$$\gamma(\alpha) = p_\alpha,$$

where p_α is the α quantile of the empirical distribution of p, and the adjusted percentile bootstrap confidence interval is given by [56]:

$$\mathsf{CI}^{adjp}_{\theta,(1-\alpha)} = \left[\hat{\theta}^*_{\gamma(\frac{\alpha}{2})}, \hat{\theta}^*_{\gamma(1-\frac{\alpha}{2})}\right]. \tag{8.16}$$

8.2 Validation of Bootstrap Confidence Intervals

The quality of the nonparametric bootstrap confidence intervals for multifractal attributes defined in Section 8.1 may depend on many factors. In order to check whether the CI work, to compare them and to study how they are affected by parameters such as sample size, process nature and parameters, the attribute and the CI definition, we need to define how we assess "quality" for CI. The analytical verification of bootstrap confidence intervals in terms of e.g. consistency of bootstrap quantile estimation is not feasible for multifractal analysis, since it would require knowledge of the distributions of both $\hat{\theta} - \theta$ and $\hat{\theta}^* - \hat{\theta}$, which are not accessible. Also, such (asymptotic) results would give little information only on practical CI quality for typical real-world sample sizes. Therefore, we validate the reliability of the confidence intervals defined in Section 8.1 by numerical simulations. To this end, we apply the estimation and bootstrap estimation procedures to a large number N_{MC} of realizations of H-sssi and MMC processes with known and a priori controlled multifractal attributes $\theta \in \{\zeta(q), D(q), h(q), c_p\}$, and use the bootstrap CI calculated for each realization to assess their statistical properties.

8.2.1 Empirical coverage of confidence intervals

The defining and most important property of a confidence interval is given by Eq. (8.1), stating that a $(1-\alpha)$ confidence interval $\mathsf{CI}_{\theta,(1-\alpha)}$ for a parameter θ should contain θ with probability $1-\alpha$:

$$\mathsf{Pr}\left(\theta \in \mathsf{CI}_{\theta,(1-\alpha)}\right) = 1 - \alpha. \tag{8.17}$$

Hence, it is natural to verify this key property for the estimated confidence intervals. Therefore, we evaluate the reliability of the bootstrap confidence limits $\mathsf{CI}_{\theta,(1-\alpha)}$ defined in Section 8.1 by their *empirical coverage*, which is defined as:

$$\mathcal{C}_{\theta,(1-\alpha)} = \widehat{\mathbb{E}}_{N_{MC}} \mathbb{I}\left\{\theta \in \mathsf{CI}_{\theta,(1-\alpha)}\right\}, \tag{8.18}$$

where $\mathbb{I}\{\cdot\}$ is the indicator function of the event $\{\cdot\}$, and $\widehat{\mathbb{E}}_{N_{MC}}$ stands for mean over Monte Carlo realizations. Consequently, the empirical coverage consists of the empirical probability that the parameter θ lies in the estimated confidence region and hence ideally equals $1-\alpha$. It can be verified by numerical simulations and can therefore serve as a practical measure for the reliability of the confidence interval.

Since it is likely that the estimation for $\theta \in \{\zeta(q), D(q), h(q), c_p\}$ are biased (see Chapter 4, e.g. Tab. 4.5 and 4.6), we also define the *bias corrected empirical coverage*:

$$\mathcal{C}^{bc}_{\theta,(1-\alpha)} = \widehat{\mathbb{E}}_{N_{MC}} \mathbb{I}\left\{\theta + \left(\widehat{\mathbb{E}}_{N_{MC}}\hat{\theta} - \theta\right) \in \mathsf{CI}_{\theta,(1-\alpha)}\right\}, \tag{8.19}$$

The bias corrected coverage allows to assess the performance of the confidence limits without the contribution of a potential bias[2] in the estimation of θ.

8.2.2 Further quality measures for confidence intervals

In addition to the empirical coverage of CI, one can define complementary statistical characterizations for investigation of the tail probabilities, CI size, or asymmetry. For instance, the empirical coverage of the lower and the upper limits of the CI, i.e., the empirical counterparts of $\Pr\left(\hat{\theta}_l \geq \theta\right)$ and $\Pr\left(\hat{\theta}_r \leq \theta\right)$, should ideally be the same and both equal to $\frac{\alpha}{2}$. Also, the size of CI, $L_{CI_\theta} = \hat{\theta}_r - \hat{\theta}_l$, can be useful to compare different CI if they have the same (empirical) coverage: The one with the smaller size is in general preferable, since it further restricts the interval within which the parameter θ lies with probability equal to the (empirical) coverage.

We have numerically investigated such quality measures for bootstrap confidence limits for multifractal attributes. However, no clear conclusions could be obtained from the results. Therefore, we will not further consider them in this manuscript.

8.3 Performance of Bootstrap Confidence Intervals

The goal of this section is to study the practical relevance of bootstrap CI estimation for multifractal attributes, and to check whether these CI exhibit satisfactory performance. Also, it investigates some multifractal analysis specific aspects related to bootstrap CI estimation and aims at providing elements of answers to the following questions:

- Does bootstrap perform better than Gaussian expansions (Section 8.3.1)?
- Which of the bootstrap CI in Section 8.1 is to be preferred for multifractal analysis (Section 8.3.2)?
- Is bootstrap CI estimation more difficult for wavelet Leaders than for wavelet coefficients (Section 8.3.3)?
- How relevant are bootstrap CI for H-sssi and MMC processes, respectively? (Section 8.3.4)?

[2] We note that results not reported here show that the difference between $\mathcal{C}_{\theta,(1-\alpha)}$ and $\mathcal{C}^{bc}_{\theta,(1-\alpha)}$ is in general small. This is also clear from results reported in Chapter 4, demonstrating that the statistical performance of estimation procedures for multifractal analysis is mostly determined by standard deviation/variability of the estimates, whereas the bias remains relatively small.

- Which of the bootstrap resampling procedures – T-B (Section 7.2.1) or TS-B (Section 7.2.2) – is to be preferred, and in which situation (Section 8.3.5)?

- Does the precise nature of the multifractal attribute have an impact on bootstrap CI performance (Section 8.3.6)?

- Does sample size play a critical role for bootstrap CI relevance (Section 8.3.7)?

To this end, Tab. 8.1, 8.2 and 8.3 summarize empirical coverage results for 1d signals. The presentation of results in three different tables aims at helping to disassociate the different questions as far as possible. Although effort has been put into disassociating the different topics in the discussion, this is obviously not always completely achievable, accounting for some minor redundancy or overlap in the results and discussions.

Results for 2d signals (images) are presented in Tab. 8.4 and discussed separately in Section 8.3.8.

8.3.1 Gaussian expansion vs. bootstrap for wavelet coefficients

For wavelet coefficients of fBm, analytic expressions for the (asymptotic) distributions of structure functions $S(j,q)$ and the scaling exponent estimates $\hat{\zeta}(q)$, $q>0$ can be obtained [4, 70] and in turn used for CI construction. The goal of this section is to investigate the relevance of these asymptotic CI – together with those of bootstrap CI – for H-sssi and MMC processes, and to compare their relative performance.

8.3.1-a) Gaussian expansion for wavelet coefficients of fBm

The variance $\sigma_G^2(j,q)$, $q>0$, of the wavelet coefficient structure functions $S^d(j,q)$ can be expressed as [4]:

$$\sigma_G^2(j,q) = \mathsf{Var}\log_2 S(j,q) = \frac{(\log_2 e)^2}{n_j}\left(\sqrt{\pi}\Gamma\left(q+\frac{1}{2}\right)/\Gamma\left(\frac{q+1}{2}\right)^2 - 1\right), \qquad (8.20)$$

by using a standard approximation formula for change of variable $Y=f(X)$, $\sigma_Y^2 \simeq |f'(\mathbb{E}X)|^2\sigma_X^2$, and assuming that wavelet coefficients at a given scale are uncorrelated and Gaussian. Further assuming that wavelet coefficients at different scales are uncorrelated, the variance $V(q)$ of the wavelet coefficient based estimates of $\zeta(q)$ for fBm is given by:

$$V(q) \equiv \sum_{j=j_1}^{j_2} w_j^2 \sigma_G^2(j,q), \qquad (8.21)$$

and does not demand for estimation of any quantity. The variance $V(q)$ can in turn be used for the construction of an (asymptotic) Gaussian $(1-\alpha)$ confidence interval for $\zeta(q)$, which is defined as:

$$\mathsf{CI}^G_{\zeta(q),(1-\alpha)} = \left(\hat{\zeta}(q) + q_{(\frac{\alpha}{2})}\sqrt{V(q)}, \hat{\zeta}(q) + q_{(1-\frac{\alpha}{2})}\sqrt{V(q)}\right), \qquad (8.22)$$

where q_α is the α quantile of the standard normal distribution $\mathcal{N}(0,1)$.

$\mathcal{C}_\theta$	$N=2^9$				$N=2^{12}$				$N=2^{15}$			
$\alpha=0.05$	fBm		mrw		fBm		mrw		fBm		mrw	
θ	per	G	per	G	per	G	per	G	per	G	per	G
$\zeta(1)$	87.8	**91.6**	**91.5**	85.5	88.2	**89.2**	**89.7**	71.0	**89.9**	88.2	**72.5**	30.0
$\zeta(2)$	85.1	**91.4**	**90.8**	81.1	87.4	**88.9**	**89.6**	59.0	**88.1**	86.8	**82.2**	35.2
$\zeta(3)$	83.9	**91.6**	**91.0**	78.3	84.6	**89.4**	**90.5**	53.7	86.7	**87.7**	**88.9**	32.7
$\zeta(4)$	82.3	**92.1**	**90.4**	77.2	83.4	**90.0**	**91.9**	50.5	86.3	**89.2**	**92.6**	28.4
$\zeta(5)$	82.0	**92.7**	**90.5**	76.8	82.3	**90.9**	**92.7**	48.6	84.6	**90.4**	**92.1**	26.2

Table 8.1: **Gaussian expansion and bootstrap confidence interval coverage for wavelet coefficients.** Empirical coverages (in $\%$) of $100\cdot(1-\alpha)=95\%$ Gaussian (G) and percentile bootstrap (per) confidence intervals for wavelet coefficient based estimation and fBm ($H=0.8$) and mrw ($(H,\beta^2)=(0.75,0.08)$) of sample size $N\in\{2^9,2^{12},2^{15}\}$, obtained for $N_{MC}=3000$ realizations. Estimation is performed using Daubechies wavelets with $N_\psi=3$ and weighted (w^1) linear regressions over the scales $j\in[3,\log_2(N)-4]$. The bootstrap CI are estimated from $R=200$ T-B bootstrap samples. For each estimate and sample size, empirical coverages closest to the target coverage of 95% are marked in bold

8.3.1-b) Gaussian expansion vs. bootstrap confidence interval

Tab. 8.1 compares the performance of the Gaussian (G) confidence intervals Eq. (8.22) for $\zeta(q)$ and wavelet coefficient based estimation to the performance of their percentile bootstrap counterparts (per) Eq. (8.6). It summarizes empirical coverages Eq. (8.18) obtained for $N_{MC}=3000$ realizations of fBm ($H=0.8$) and mrw ($(H,\beta^2)=(0.75,0.08)$) of sample size $N\in\{2^9,2^{12},2^{15}\}$. The percentile bootstrap CI are obtained by T-B bootstrap (Section 7.2.1), using $R=200$ bootstrap samples per realization. The nominal target coverage is $100\cdot(1-\alpha)=95\%$.

Tab. 8.1 shows that for (Gaussian) fBm, the analytic Gaussian confidence intervals systematically have empirical coverages closer to the target coverage than their bootstrap counterparts, hence outperforming them. Nonetheless, the performance of the bootstrap confidence intervals are close to those of Gaussian CI, in particular for large sample size N and small statistical orders q.
However, for (non Gaussian and multifractal) mrw, the analytic Gaussian CI fail rather drastically, with coverages of the order of only 30% for large sample size ($N=2^{15}$). Clearly, some of the assumptions on which Eq. (8.21) is based – Gaussian coefficients, decorrelated in time and decorrelated between scales – are violated for mrw[3]. In contrast, the bootstrap CI estimation for mrw has satisfactory empirical coverage of $85-90\%$ in average, comparable to bootstrap CI empirical coverage for fBm.

Conclusions. These results lead us to the following conclusions: First, in the case of wavelet coefficient based estimation of $\zeta(q)$, $q>0$ for fBm, analytic Gaussian confidence limits are slightly more reliable and easier to calculate than bootstrap confidence limits and should hence be preferred. Second, for non Gaussian or multifractal processes, the

[3]For small sample size $N=2^9$, the situation appears to be less dramatic for the Gaussian CI, with coverages of $77-86\%$ as compared to $90-92\%$ for the bootstrap CI. This is due to $V(q)$ being close to the true variance of $\hat{\zeta}(q)$ by coincidence for this sample size and process parameter choice, despite violated assumptions for $V(q)$.

Gaussian expansions fail dramatically. Third, the bootstrap confidence intervals for coefficient based estimation are highly relevant for both fBm and, in particular, (non Gaussian and multifractal) mrw. This is, to the best of our knowledge, the first nonparametric CI procedure with satisfactory performance for non Gaussian and multifractal processes.

8.3.2 Comparison of different bootstrap confidence interval methods

The aim of this section is to validate, first, whether simple or double bootstrap CI estimations are to be preferred for multifractal attributes, and second, which of the precise definitions of simple (double) bootstrap CI are most relevant for multifractal attributes.
Tab. 8.2 compares the performance of the simple and double bootstrap confidence intervals defined in Section 8.1. It summarizes bias corrected empirical coverages Eq. (8.19) obtained for $N_{MC} = 1000$ realizations of fBm ($H = 0.8$) and mrw ($(H, \beta^2) = (0.72, 0.08)$) of sample sizes $N = 2^{12}$, for both wavelet coefficient and Leader based estimation. The bootstrap confidence intervals are obtained by T-B bootstrap (Section 7.2.1), with $R = 599$ primary bootstrap samples per realization, and $S = 50$ double bootstrap samples per primary bootstrap sample. The nominal target coverage is $100 \cdot (1 - \alpha) = 95\%$. For each estimate, empirical coverages closest to target coverage are marked in bold in Tab. 8.2.

The following statements are valid for both wavelet coefficient and for wavelet Leader based estimations:
First, Tab. 8.2 shows that for fBm, double bootstrap CI are slightly more relevant than simple bootstrap CI, having average empirical coverages of 92% (89%) for double bootstrap, as opposed to 89% (82%) for simple bootstrap confidence intervals for wavelet coefficient (Leader) based estimation. For mrw, the difference in empirical coverage between simple and double bootstrap CI is not significant, all being in average close to the target value of 95%.
Second, whereas for fBm, there is no significant difference in performance between the simple bootstrap CI nor, bas and per, for mrw, the percentile bootstrap CI per has overall coverage slightly closer to the target than nor and bas
Third, for fBm, the adjusted percentile bootstrap CI $adjp$ has overall coverage slightly closer to the target than stu and $adjb$, whereas for mrw, there is overall no significant difference in empirical coverage of the double bootstrap CI.

Conclusions. We conclude that for multifractal analysis, double bootstrap confidence intervals are in practice not worth the severe increase of computational load by a factor S, since they are not more relevant than single bootstrap CI for multifractal processes and only slightly so for fBm. Also, among the simple bootstrap CI considered, the conceptually simple percentile bootstrap CI shows overall best empirical coverage and is hence to be preferred. Therefore, we will in the sequel of this chapter only consider the percentile confidence interval.

8.3.3 Bootstrap for wavelet coefficients versus wavelet Leaders

As compared to wavelet coefficients, wavelet Leaders may introduce two additional potential difficulties with respect to bootstrap estimation (cf. results and discussions in Chapter 3 and Section 7.2): First, their marginal distributions are non Gaussian, and second, they may introduce additional dependence in time, and between scales. Hence, bootstrap CI

$\mathcal{C}_\theta^{bc}$ - $N = 2^{12}$ - $\alpha = 0.05$ - T-B												
fBm	d_X						L_X					
θ	nor	bas	per	stu	$adjb$	$adjp$	nor	bas	per	stu	$adjb$	$adjp$
$\zeta(2)$	88.5	86.3	87.3	89.0	90.5	**92.5**	78.6	79.5	78.2	85.1	88.0	**90.1**
$D(2)$	91.1	90.4	91.5	90.2	**94.3**	93.4	89.7	83.7	88.6	89.1	**93.9**	92.4
$h(2)$	88.4	85.2	87.0	88.1	90.1	**92.5**	79.0	78.9	81.1	85.6	87.0	**91.5**
c_1	88.6	89.3	88.3	91.5	92.9	**93.9**	78.2	77.9	77.3	83.4	90.0	**92.6**
c_2	90.3	92.3	89.7	91.1	93.3	**93.8**	86.8	83.4	85.5	88.0	**91.7**	91.1
mrw	d_X						L_X					
θ	nor	bas	per	stu	$adjb$	$adjp$	nor	bas	per	stu	$adjb$	$adjp$
$\zeta(2)$	96.0	93.1	96.4	**95.1**	96.1	97.9	**93.7**	92.4	93.1	96.7	97.0	98.4
$D(2)$	88.8	78.7	**93.3**	84.3	89.4	85.5	84.1	76.5	**90.3**	86.6	89.9	87.5
$h(2)$	91.1	84.0	**94.4**	89.7	93.4	92.7	87.9	83.3	91.0	91.1	**93.6**	93.4
c_1	97.3	97.0	96.1	98.2	98.3	99.5	97.4	97.4	**96.3**	99.2	99.1	99.6
c_2	94.6	93.8	93.4	92.5	93.9	**95.6**	95.2	91.5	94.5	**95.1**	95.5	96.9

Table 8.2: **Simple and double bootstrap confidence interval coverage.** Bias corrected empirical coverage $\mathcal{C}_\theta^{bc}$ (in $\%$) of $100 \cdot (1-\alpha) = 95\%$ simple (nor, bas, per) and double (stu, $adjb$, $adjp$) bootstrap confidence intervals for wavelet coefficient (d_X, left) and Leader (L_X, right) based confidence interval estimation for fBm (top, $H = 0.8$) and mrw (bottom, $(H, \beta^2) = (0.72, 0.08)$) of sample size $N = 2^{12}$. Results are obtained using $N_{MC} = 1000$ realizations, Daubechies wavelets with $N_\psi = 3$, and weighted (w^1) linear regressions over the scales $j \in [3, 8]$. The bootstrap CI are estimated from $R = 599$ T-B bootstrap samples per realization, and $S = 50$ double bootstrap samples per single bootstrap sample. For each estimate, empirical coverages closest to the target coverage of 95% are marked in bold.

estimation might be more difficult for wavelet Leaders in some situations. The investigation and clarification of this point is the precise aim of this section.

Tab. 8.3 summarizes bias corrected empirical coverages Eq. (8.19) for both wavelet coefficient and Leader based estimation of percentile CI Eq. (8.6), for both T-B and TS-B bootstrap resampling (cf. Section 7.2.1 and 7.2.2). Results are obtained using $N_{MC} = 1000$ realizations of fBm ($H = 0.8$) and mrw ($(H, \beta^2) = (0.72, 0.08)$) of sample size $N \in \{2^{12}, 2^{15}\}$, with $R = 199$ bootstrap samples per realization. The nominal target coverage is $100 \cdot (1-\alpha) = 90\%$. Additional results can be found in Tab. 8.2.

Tab. 8.3 indicates that for fBm and $N = 2^{12}$, bootstrap percentile CI estimation is slightly but systematically better for wavelet coefficients than for wavelet Leaders, regardless of which resampling procedure (T-B or TS-B) is employed: Empirical coverages of bootstrap confidence intervals for wavelet Leader based estimation is consistently further below target coverage than empirical coverages of CI for coefficient based estimation. An interpretation of this result can be given by noting that for fBm, wavelet coefficients are Gaussians, whereas wavelet Leaders are not and hence present extra difficulties for bootstrap estimation. For fBm and sample size $N = 2^{15}$, the situation is less clear, and there is no systematic difference in confidence interval relevance between wavelet coefficient and Leader based estimation. Also, for mrw and any of the sample sizes considered, bootstrap CI (T-B or TS-B) for wavelet coefficient and Leader based estimation do overall have comparable empirical coverages.

These observations are in agreement with results for fBm and mrw and $N = 2^{12}$ in Tab. 8.2 for T-B bootstrap resampling based estimation of different simple and double bootstrap confidence intervals: Whereas for fBm, (simple and double) bootstrap confidence intervals are overall slightly more relevant for wavelet coefficient based estimation than for wavelet Leader based estimation, there is no such difference for mrw, for which empirical coverages of bootstrap CI for wavelet coefficient based estimation are very close to those for wavelet Leader based estimation.

Conclusions. The results in Tab. 8.3 and Tab. 8.2 lead us to the conclusion that for fBm and small sample sizes, bootstrap confidence interval estimation is slightly more difficult for wavelet Leaders than for wavelet coefficients. For MMC processes and for larger sample sizes, there is no significant difference in CI performance between wavelet coefficient and Leader based estimation. This can be interpreted as a consequence of dependence structures and marginal distributions that are more complicated for wavelet Leaders than for wavelet coefficients for fBm, but similar for mrw.

8.3.4 Bootstrap for H-sssi vs. MMC processes

Due to different marginal and dependence properties for H-sssi and and MMC processes, bootstrap CI estimation might be more difficult or relevant for the one or the other process class. The discussion of such topics is the precise goal of this section.

For wavelet coefficient based estimation, Tab. 8.3 indicates that empirical coverages are comparable for the H-sssi process fBm and the MMC process mrw, with overall performance being slightly better for mrw than for fBm for small sample size ($N = 2^{12}$). This is in agreement with Tab. 8.1 and Tab. 8.2, where empirical coverages of T-B bootstrap confidence intervals for wavelet coefficient based estimation are overall slightly closer to target coverage for mrw than for fBm ($N = 2^{12}$).
Results in Tab. 8.3 show (in agreement with Tab. 8.2) that for wavelet Leader based estimation, bootstrap confidence intervals are more reliable for multifractal mrw than for monofractal fBm: empirical coverages are systematically closer to the target value of 90% for mrw than for fBm.
We conclude that the nature of the process under analysis – H-sssi or MMC – has an impact on the quality of bootstrap CI estimation: empirical coverages of bootstrap CI are slightly closer to nominal value for mrw than for fBm for wavelet coefficient based estimation, and more significantly so for wavelet Leader based estimation. We propose an explanation and interpretation of this observation in Chapter 11.

8.3.5 T-B vs. TS-B bootstrap

The T-B and TS-B (Sections 7.2.1 and 7.2.2, respectively) bootstrap resampling procedures for multifractal analysis use substantially different block constructions, potentially resulting in different bootstrap estimation performance for multifractal analysis due to different dependence and structural assumptions. The aim of this section is to investigate the impact of choosing T-B or TS-B bootstrap resampling on the quality of bootstrap CI estimation.

Wavelet coefficient based estimation. Tab. 8.3 indicates that for fBm, TS-B confidence intervals are overall more relevant than T-B CI: Empirical coverages of TS-B CI are

systematically closer to target coverage of 90% than those of the T-B CI. The difference in empirical coverage between T-B and TS-B is small ($1-7\%$) for small sample size, and more pronounced for large samples ($4-10\%$). Also, the empirical coverages of the T-B CI fluctuate significantly with the multifractal attribute considered ($74-90\%$ for $N=2^{12}$ and $76-96\%$ for $N=2^{15}$), whereas those of the TS-B CI remain relatively stable ($80-83\%$ for $N=2^{12}$ and $82-87\%$ for $N=2^{15}$).
For mrw, these differences between T-B and TS-B CI empirical coverages are even more pronounced: $2-9\%$ ($N=2^{12}$) and $1-10\%$ ($N=2^{15}$). Also, the empirical coverages of the T-B CI fluctuate rather dramatically with the multifractal attribute considered ($76-99\%$ for $N=2^{12}$ and $82-100\%$ for $N=2^{15}$), whereas those of the TS-B CI remain relatively constant ($81-91\%$ for $N=2^{12}$ and $85-92\%$ for $N=2^{15}$).

Wavelet Leader based estimation. Tab. 8.3 shows that for both fBm and mrw, TS-B confidence intervals for wavelet Leader based estimation are more relevant than T-B CI: Empirical coverages of TS-B CI are systematically closer to target coverage of 90% than those of the T-B CI. The difference in empirical coverage between T-B and TS-B is relatively small for small sample size ($2-8\%$ for fBm and $3-12\%$ for mrw), and more significant for large samples ($4-14\%$ for fBm and $6-12\%$ for mrw). Also, the empirical coverages of the T-B CI fluctuate significantly with the multifractal attribute considered, whereas those of the TS-B CI remain relatively constant.

Conclusions. We conclude that for both wavelet coefficient and Leader based estimation of bootstrap confidence intervals for multifractal attributes, TS-B bootstrap is to be preferred over T-B bootstrap: TS-B CI do achieve better overall empirical coverage, and are more robust with respect to both multifractal attribute choice and sample size, in particular for MMC processes.

8.3.6 Bootstrap performance and multifractal attributes

Overall, we observe that there is only little difference in CI empirical coverages between different multifractal attributes (cf. Tab. 8.3). However, the results in Tab. 8.3 illustrate the following important points:
First, bootstrap confidence intervals remain relevant for scaling exponents of negative statistical order $\zeta(q<0)$, with empirical coverages similar to those of CI for other multifractal attributes.
Second, empirical coverages of the wavelet coefficient based bootstrap CI for c_3 are relatively poor as compared to CI for other multifractal attributes. This may be due to the fact that coefficient based estimation $\hat{c}_3^d$ of c_3 has very poor statistical performance (see Chapter 4, e.g. Tab. 4.2 and 4.4).
Third, for mrw and TS-B bootstrap resampling, empirical coverages of bootstrap CI for $\zeta(q)$ are dropping as q increases and approaches the critical value $q=q_*^+=5$. Nevertheless, empirical coverages remain satisfactory for $\zeta(q=q_*^+=5)$ ($83-87\%$ for target coverage 90%). This constitutes important evidence in favor of a hypothesis that it is not the critical statistical order q_*^+ encountered through the linearization effect (cf. Sections 2.8 and 5.2), but the statistical order $q_c^+>q_*^+$ – for which, by definition of the process, moments become infinite – that plays the central role for bootstrap consistency (cf. Section 7.3.2-a)).

fBm	$N = 2^{12}$				$N = 2^{15}$			
	d_X		L_X		d_X		L_X	
θ	T	TS	T	TS	T	TS	T	TS
$\zeta(-2)$	–	–	66.0	72.4	–	–	71.0	83.8
$\zeta(1)$	78.4	81.6	69.6	75.6	80.6	87.4	71.6	85.2
$\zeta(2)$	76.6	80.6	69.0	74.0	79.4	85.8	71.8	83.8
$\zeta(3)$	75.4	79.4	69.0	73.6	78.2	83.8	73.6	84.0
$\zeta(5)$	74.0	78.4	70.4	72.6	75.8	82.2	76.6	84.4
$D(-2)$	–	–	75.6	80.8	–	–	80.6	84.2
$D(2)$	89.8	87.4	89.8	86.8	95.6	86.4	95.0	85.0
$h(-2)$	–	–	67.4	73.8	–	–	73.0	83.6
$h(2)$	75.0	79.8	71.6	74.0	77.6	84.0	76.4	85.2
c_1	76.2	83.2	66.8	75.2	80.2	87.0	71.4	85.4
c_2	79.0	83.4	77.2	79.8	79.0	85.2	82.8	87.6
c_3	78.2	82.6	83.4	88.6	78.6	82.4	88.4	91.0
mrw	$N = 2^{12}$				$N = 2^{15}$			
	d_X		L_X		d_X		L_X	
θ	T	TS	T	TS	T	TS	T	TS
$\zeta(-2)$	–	–	96.8	84.6	–	–	99.4	87.0
$\zeta(1)$	98.8	89.8	98.0	88.4	100.0	91.2	100.0	91.6
$\zeta(2)$	97.0	88.8	97.2	88.4	99.6	92.4	100.0	91.0
$\zeta(3)$	95.2	88.2	94.6	86.6	98.2	90.0	98.8	88.4
$\zeta(5)$	89.2	87.0	91.2	83.4	94.8	85.4	95.0	86.4
$D(-2)$	–	–	88.6	84.8	–	–	93.0	81.6
$D(2)$	89.8	87.4	89.8	86.8	95.6	86.4	95.0	85.0
$h(-2)$	–	–	90.0	80.0	–	–	96.8	84.8
$h(2)$	90.8	87.0	92.4	86.2	96.6	88.8	97.8	89.2
c_1	98.4	91.0	99.0	89.8	99.6	89.4	100.0	92.0
c_2	82.6	84.6	93.4	89.0	93.0	86.8	99.6	93.4
c_3	76.4	81.2	93.6	90.2	82.4	83.8	95.8	88.4

Table 8.3: **T-B and TS-B bootstrap confidence interval coverage.** Bias corrected empirical coverage $\mathcal{C}_\theta^{bc}$ (in %) of $100{\cdot}(1{-}\alpha) = 90\%$ percentile (per) bootstrap confidence intervals for wavelet coefficient (d_X, left columns) and Leader (L_X, right columns) based confidence interval estimation for fBm (top, $H = 0.8$) and mrw (bottom, $(H, \beta^2) = (0.72, 0.08)$) of sample sizes $N = 2^{12}$ (left) and $N = 2^{15}$ (right). Results are obtained using $N_{MC} = 1000$ ($N = 2^{12}$) and $N_{MC} = 500$ ($N = 2^{15}$) realizations, Daubechies wavelets with $N_\psi = 3$, and weighted (w^1) linear regressions over the scales $j \in [3, \log_2(N) - 4]$. The bootstrap CI are estimated from $R = 199$ bootstrap samples per realization.

8.3.7 Bootstrap performance and sample size

Tab. 8.3 shows that for T-B bootstrap based CI, empirical coverages systematically increase when the sample size is increased. Whereas for fBm larger sample size results in better empirical coverages, for mrw, this results in CI that are too conservative (i.e. too large, resulting in empirical coverages close to 100%), hence better CI coverage for small sample size, in contradiction to what would be expected if CI estimation was consistent.
In contrast, empirical coverages of TS-B bootstrap based CI do not necessarily increase with increasing sample size, but rather converges to target coverage: Whereas for fBm (d_X and L_X) and mrw (L_X) this also results in an increase in empirical coverages – since CI are too small for $N = 2^{12}$ – empirical coverages remain overall constant for mrw (d_X), for which empirical coverages closely reproduce the target coverage of 90% already for $N = 2^{12}$.

Conclusions. We conclude that whereas the T-B bootstrap based estimation of bootstrap CI has a trend to overestimate the variability of multifractal attribute estimates when the sample size N increases, resulting in too conservative confidence intervals for large N, the TS-B bootstrap based estimation produces CI with empirical coverages converging to target coverage as N increases.

8.3.8 Bootstrap performance for multifractal analysis of images

The aim of this section is to validate bootstrap multifractal attribute confidence interval estimation for 2d signals (images), hence revisiting some of the points discussed above for 1d signals, leading in general to similar conclusions, while leaving other ones out: Since it is practically difficult to significantly vary the sample size for 2d multifractal analysis, we do not consider the influence of sample size here. Also, we do not revisit Gaussian expansion results, since results are rather drastic for 1d signals and there is no reason that this would be different for images. Also, we concentrate on the percentile bootstrap CI only.

Tab. 8.4 summarizes bias corrected empirical coverages Eq. (8.19) of percentile confidence intervals Eq. (8.6) obtained by the T-B and TS-B bootstrap resampling (Section 7.2.1 and 7.2.2, respectively) for both wavelet coefficient and Leader based estimation. Results are obtained using $N_{MC} = 500$ 2d realizations of fBm ($H = 0.7$), fractionally integrated (by $\alpha = 0.5$) $Q_r(t)$ of CMC-LN ($m = 0.0125$, i.e. $(c_1, c_2) = (0.513, -0.025)$) and CMC-LP ($\beta = 0.8395, \gamma = 0.4195$, i.e. $(c_1, c_2, c_3) = (0.538, -0.080, 0.014)$) of sample sizes $N \times N = 2^{10} \times 2^{10}$, and using $R = 99$ bootstrap samples per realization. The nominal coverage is $100 \cdot (1 - \alpha) = 90\%$.

Wavelet coefficients vs. wavelet Leaders. First, we observe that for fBm, both T-B and TS-B based confidence intervals are more relevant for coefficients than for Leader based estimation, with on average 5% (T-B) and 10% (TS-B) smaller coverage errors for d_X than for L_X. Second, for multifractal CMC, TS-B CI have better empirical coverages for coefficient than for Leader based estimation, with coverage errors smaller by approximately 10% on average, whereas for T-B based CI, there is no significant difference.
We conclude that – as for (1d) signals – bootstrap CI estimation is slightly more difficult for Leaders than for coefficients, in particular for (Gaussian and monofractal) fBm.

	fBm				CMC LN				CMC LP			
	d_X		L_X		d_X		L_X		d_X		L_X	
	T	TS	T	TS	T	TS	T	TS	T	TS	T	TS
$\zeta(-2)$	$-$	$-$	73.8	79.4	$-$	$-$	89.0	79.6	$-$	$-$	86.4	80.8
$\zeta(-1)$	$-$	$-$	75.0	79.8	$-$	$-$	87.6	79.4	$-$	$-$	89.8	82.0
$\zeta(1)$	80.8	91.8	74.0	79.0	94.6	88.4	88.2	78.6	97.8	89.0	95.2	82.6
$\zeta(2)$	80.6	88.8	73.8	78.6	94.0	86.4	87.6	77.6	98.4	89.2	95.2	81.2
c_1	81.0	94.4	74.2	78.8	94.2	88.2	88.4	77.6	97.2	88.2	94.8	81.0
c_2	83.6	87.8	81.2	82.0	86.0	89.2	85.6	80.8	87.8	89.8	89.6	81.2
c_3	$-$	$-$	$-$	$-$	85.8	88.2	85.4	82.2	84.4	88.2	85.0	83.0

Table 8.4: **T-B and TS-B bootstrap confidence interval coverage for images.** Bias corrected empirical coverage $\mathcal{C}_\theta^{bc}$ (in %) of $100 \cdot (1-\alpha) = 90\%$ percentile (per) bootstrap confidence intervals for wavelet coefficient (d_X, left columns) and Leader (L_X, right columns) based confidence interval estimation for 2d fBm (left, $H = 0.8$), (fractionally integrated by $\alpha = 0.5$ $Q_r(t)$ of) CMC-LN (center, $(c_1, c_2) = (0.513, -0.025)$) and CMC-LP (right, $(c_1, c_2, c_3) = (0.538, -0.080, 0.014)$) of sample sizes $N \times N = 2^{10} \times 2^{10}$. Results are obtained using $N_{MC} = 500$ realizations, Daubechies wavelets with $N_\psi = 2$, and weighted (w^1) linear regressions over the scales $j \in [3, 7]$. The bootstrap CI are estimated from $R = 99$ bootstrap samples per realization.

H-sssi vs. MMC processes. Tab. 8.4 shows that for T-B bootstrap CI estimation, performance is slightly better for (multifractal) CMC than for fBm, with coverage errors smaller by approximately 5% (d_X) and 10% (L_X). For TS-B bootstrap CI estimation, the difference is less significant. Also, there is no significant difference in CI coverages between log-Normal and log-Poisson CMC and hence, a change in marginal distribution of the multipliers of the MMC processes (LN vs. LP) does not significantly impact CI performance.
We conclude that, as for (1d) signals, the bootstrap CI for (2d) images are slightly more relevant for MMC processes than for H-sssi processes.

T-B vs. TS-B bootstrap. The results in Tab. 8.4 indicate that the difference between T-B and TS-B bootstrap based CI coverages remain overall small. Nevertheless, for wavelet coefficient based estimation and both fBm and CMC, TS-B bootstrap based CI display better empirical coverages than T-B based ones, with coverage errors of approximately $2-4\%$ (TS-B) and $2-9\%$ (T-B). The situation is similar for wavelet Leader based estimation for fBm, whereas for CMC, TS-B bootstrap based CI are slightly less relevant than T-B bootstrap based CI. Also, while TS-B based CI empirical coverages remain constant regardless of which precise multifractal attribute is considered, T-B CI coverages display larger variation with multifractal attributes.
We conclude that similarly to (1d) signals, TS-B bootstrap CI perform overall slightly better than T-B based CI for multifractal analysis of (2d) images.

Multifractal attributes. Finally, confidence interval estimations for negative statistical orders q, $\zeta(q < 0)$, are highly relevant and produce empirical coverages comparable to the overall coverages for other multifractal attributes. Also, for 2d signals, there is not significant difference in CI performance among the multifractal attributes considered.

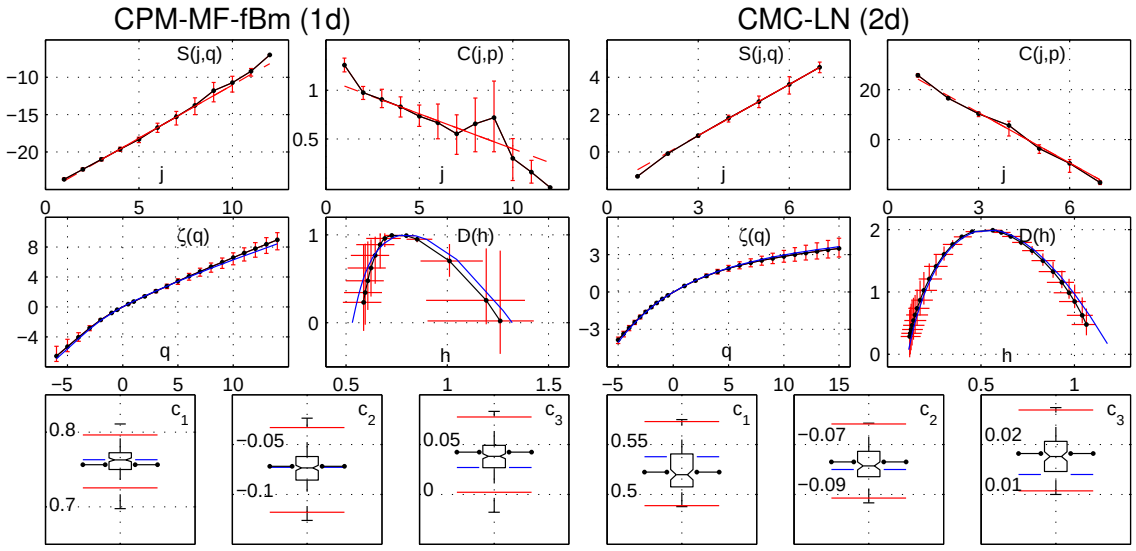


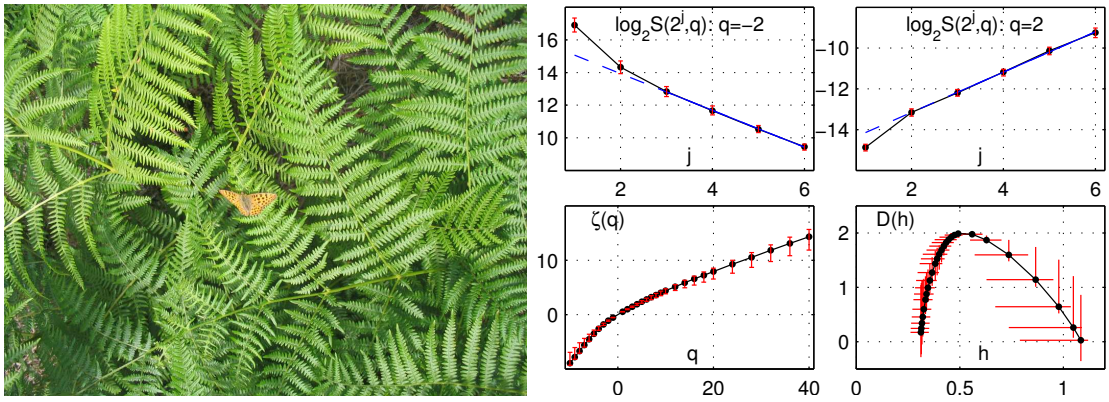
Figure 8.1: **Bootstrap confidence intervals for one single realization of CPM-MF-fBm and CMC-LN.** Estimation for one single realization of CPM-MF-fBm (left, $N = 2^{15}$, $(c_1, c_2, c_3) = (0.8, -0.08, 0.0311)$) and 2d CMC-LN (right, $N \times N = 2^{10} \times 2^{10}$, $(c_1, c_2, c_3) = (0.538, -0.08, 0.014)$, obtained from $Q_r(t)$ by fractional integration of order $\alpha = 0.5$): Estimates (solid black, '•') of structure functions $S(j, q = 2)$ (top left), cumulants $C(j, p = 2)$ (top right), scaling exponents $\zeta(q)$ (center left), multifractal spectrum $D(h)$ (center right) and log cumulants c_p (bottom) based on wavelet Leaders. The corresponding theoretical values are drawn in solid blue. The percentile bootstrap confidence intervals (in red) are obtained using $R = 399$ T-B bootstrap samples, and for $\alpha = 0.05$. The boxplots (bottom line) are based on the bootstrap estimates $\hat{c}_p^*$ and show the lower and upper quartile, median and support of their empirical distributions. Confidence intervals exclude zero values, illustrating that the proposed procedures do discriminate zero from non-zero values for those parameters, a non trivial result for c_3 (cf. discussion in Chapter 4).

8.3.9 Conclusions: Overall performance of bootstrap CI estimation

Difference in empirical coverages between process classes, confidence interval constructions, coefficients and Leaders, multifractal attributes, sample sizes and block bootstrap resampling being relatively small, let us conclude with an overall appreciation of bootstrap CI estimation performance for multifractal attributes. The results reported in Tab. 8.1 to 8.4 illustrate that, for the processes and sample sizes considered, bootstrap CI estimation has satisfactory performance, for both (1d) signals and (2d) images: Coverage errors remain moderate and mostly constrained to $5 - 10\%$ for both wavelet Leader and coefficient based estimation, and for both H-sssi and MMC processes. Nonetheless, in some situations, coverage errors can be significantly larger, for instance, for Leader and T-B based confidence intervals for fBm when sample size is small.

8.4 Conclusions and Perspectives

In this section, we defined nonparametric bootstrap confidence intervals for multifractal attributes, based on T-B and TS-B bootstrap resampling of wavelet coefficients and



$\hat{c}_1$	$\hat{c}_2$	$\hat{c}_3$
0.519	-0.032	0.015
$[0.481, 0.554]$	$[-0.049, -0.016]$	$[0.003, 0.026]$

Figure 8.2: **Analysis of real-world image.** Real world image (left) and estimates obtained from its gray level version: Structure functions (top right) for $q = -2$ and $q = 2$, scaling exponents $\hat{\zeta}(q)$ and spectrum $D(h)$ (bottom right), log-cumulant estimates (table). Estimates are obtained using fractional integration of order $\alpha = 0.5$ to ensure minimum regularity, Daubechies wavelet ($N_\psi = 3$) and ordinary (w^0) fits over the scales $2^3 \leq 2^j \leq 2^6$, for which structure functions $\log_2 S(2^j, q)$ are approximately linear functions of j. Percentile confidence intervals (red vertical and horizontal bars) are obtained with TS-B bootstrap and $R = 99$.

Leaders, and validated and compared their respective performance by means of numerical simulations based on synthetic 1d and 2d H-sssi and MMC processes. The proposed nonparametric bootstrap confidence interval estimation procedures are – to the best of our knowledge and apart from one singular special case (wavelet coefficient based estimation of $\zeta(q)$, $q > 0$, for fBm) – the only such procedures for multifractal attributes available at present. In particular, they are the only procedures available that produce confidence intervals with satisfactory performance for non Gaussian or MMC scaling processes, and that are relevant for wavelet Leader based estimation.
The procedures are readily applicable to single finite size realizations of 1d or 2d processes. This is illustrated in Fig. 8.1 for one 1d realization of CPM-MF-fBm (left), and one 2d realization of CMC-LN (right).

The results presented in this section show that the computationally demanding double bootstrap confidence limit methods are not worth the extra effort for multifractal analysis, achieving only marginal performance gains with respect to conceptually simple and computationally inexpensive simple bootstrap CI methods. In particular, the results lead us to the conclusion that the percentile bootstrap confidence limit is a good compromise between overall performance for different multifractal attributes, and computational cost. Also, TS-B bootstrap produces overall better results than T-B bootstrap. Yet, conceptually higher complexity as well as implementation (for instance, consideration of border effects across scale) and memory cost issues make it less clear whether TS-B bootstrap is to be preferred over T-B in practice. Finally, the results reported here also confirm the empirical choice of block size proposed in Section 7.3.4.
We conclude that for both 1d and 2d, for (Gaussian monofractal) H-sssi and for (non

Gaussian multifractal) MMC processes, and for both wavelet Leader and coefficient based estimation, bootstrap confidence intervals should in practice be based on TS-B bootstrap, and on simple bootstrap CI constructions.

Analysis of a real-world image. We illustrate the performance of the wavelet Leader based estimation procedures on (the gray level version of) a real-world image of size 1024×768, taken by the author of this manuscript with a standard digital camera. It consists of natural fern leaves (cf. Fig. 8.2, left) in reference to the synthetic fern image used as a paradigm for the concept of fractal. Fig. 8.2 (right) shows structure functions (top row), scaling exponents (bottom left) and the spectrum (bottom right) estimated from this image with the 2d WLMF, and the corresponding TS-B bootstrap confidence intervals. Results on log-cumulants are summarized in the table below the figure. Results demonstrate that the method can be readily applied to real-world images for their complete multifractal characterization.

Further issues. Finally, the results indicate some unexpected behaviors of bootstrap CI estimation performance that demand for further investigation: First, T-B bootstrap CI do not seem to converge to correct size and empirical coverage as sample size increases, but rather tend to become too conservative.
Second, bootstrap CI estimation performance is overall better for (multifractal and non Gaussian) mrw and CMC than for (monofractal and Gaussian) processes. This is rather surprising, in particular for wavelet coefficient based estimation, since wavelet coefficients for fBm are known to be Gaussian and only weakly correlated, whereas they are necessarily non Gaussian for MMC processes. Therefore, switching from fBm to an MMC process should theoretically hinder (bootstrap) estimation, in contrast to what numerical simulations indicate.
Third, investigation of bootstrap estimations for structure functions (reported in Chapter 11) shows that the variability of structure functions is dramatically underestimated by the bootstrap for MMC processes, though bootstrap estimation quality remains correct for multifractal attributes.
These issues lead us to study theoretically and numerically the dependence structure of wavelet coefficients and Leaders for MMC processes. These results are summarized in Chapter 11, together with bootstrap relevant discussions.

Chapter 9

Bootstrap Hypothesis Tests

Contents

A key practical issue in multifractal analysis consists in deciding which scaling model better fits the data. In much of the literature related to practical analysis of scaling (or multifractal) properties of empirical data, practitioners are trying to assess whether the data they analyze are better described by self-similar with stationary increments (H-sssi) processes, or by multifractal processes. Such a systematic concern corresponds to both theoretical and practical preoccupations:
First, while H-sssi processes are deeply tied to additive random walks, the most prominent and practically useful examples of multifractal processes are multifractal multiplicative cascade (MMC) processes and based on recursive multiplicative constructions. Thus, one can naturally imagine that the (physical or biological) mechanisms underlying the data are likely to be significantly different for data that can be associated with additive or with multiplicative structures.
Second, while the scaling properties of H-sssi processes are entirely described by a single parameter H, referred to as the self-similarity parameter, those of multifractal processes involve a whole collection of such parameters. Since the goal of practitioners consists in relating these scaling parameters to *physical* parameters controlling the data or the experiment, there is a need to decide on the actual number of independent parameters (multifractal attributes) to be involved.
Third, while practitioners often prefer MMC models whose increased number of parameters brings extra degrees of freedom and hence better fits of the data, this additional flexibility is achieved at the price of significantly increased practical difficulties. Indeed, while H-sssi processes correspond to a well-defined fully parametric class of stochastic additive processes mostly depending on a single (self-similarity) parameter, MMC processes fall into a much broader and less well-defined class of processes involving a larger number of parameters, for which the derivation of estimation or hypothesis test procedures and the assessment of their statistical performance is much more delicate.
These arguments (addition vs. multiplication, single vs. many parameters, significant extra technical and practical difficulties in estimation and modeling) underline the need for practical procedures helping in deciding whether MMC processes are needed to model data, or whether the data are satisfactorily well modeled by self-similar ones.
Along the same line, and for similar reasons, practitioners often need to validate the precise multifractal nature and properties of the data under analysis, i.e, to decide whether a *simple* multifractal model satisfactorily matches the data or if a *more elaborate* model is to be involved.
Yet, this issue received no systematic or detailed study, and no statistical procedure validating the precise multifractal nature and properties of the data under analysis is available.

Therefore, the goal of the present chapter is precisely to define statistical tests aiming at deciding whether data under analysis belong to a certain precise process class or not. The hypothesis tests rely on the combination of the three key ingredients *log-cumulants*, *wavelet Leaders* and *nonparametric bootstrap*: We use nonparametric block bootstrap resampling of wavelet Leaders and coefficients as robust means for obtaining approximate null distributions of test statistics for hypothesis tests on the log-cumulants c_p, and end up with a robust and powerful practical test procedure for the analysis of a single and finite length observation of empirical data.

The chapter is organized as follows. The statistical tests are described in Section 9.1, which gives a precise statement of the problem we consider (Section 9.1.1), gathers ba-

sics and notations for hypothesis tests (Section 9.1.2) and defines the nonparametric bootstrap tests, together with six different bootstrap acceptance regions for the tests that we analyze and compare (Section 9.1.3). In order to assess the statistical performance (significances, p-values and powers) of the proposed bootstrap tests, large sets of numerical simulations have been performed. The corresponding methodology, the simulation set up, as well as the multifractal processes used to conduct the numerical simulations, are presented in Section 9.2. The results are presented in Section 9.3, where we discuss test performance for discrimination of H-sssi and MMC models (Section 9.3.1), of different MMC models (Section 9.3.3), and test performance for non Gaussian H-sssi processes (Section 9.3.4). Further developments, conclusions and perspectives are found in Section 9.4.
The work presented in this chapter has been reported in [174, 175, 179].

9.1 Testing Statistical Hypothesis on log Cumulants

9.1.1 Formulation of the test problem

As pointed out in the introduction, a crucial practical issue in applications lies in deciding whether the data should be described by *monofractal* processes or with *multifractal* models, and also, which precise multifractal model should be used[1]. Whereas for (monofractal) H-sssi processes, the scaling exponents take the specific form $\zeta(q) = qH$, for most multifractal processes of interest, $\zeta(q)$ is a non linear (concave) function of q. Therefore, we use this property, without loss of generality, as a definition for multifractal processes. Note that such a definition is not strictly correct (cf. [89, 145] and Section 2.4.4 for a mathematically correct definition) but nevertheless sufficient for most applications[2].
Based on this practical or operational definition, discriminating between the two classes of models mostly amounts to deciding on the linearity of $\zeta(q)$ with respect to q. This is where the use of the log-cumulants is of central interest since for H-sssi processes $\zeta(q) = qH$ and hence $\forall p \geq 2 : \; c_p \equiv 0$, while for most multifractal processes of practical interest $c_2 \neq 0$. Thus, monofractal (finite variance H-sssi) is defined as $\forall p \geq 2, c_p \equiv 0$, and accordingly, $c_2 \neq 0$ defines multifractal. Also, this enables to discriminate between simple ($c_p \equiv 0, p \geq 3$) or more elaborate multifractal models. Hence, the log-cumulants c_1, c_2, c_3, provide us with central attributes to quantify the multifractal nature of the data.

This is why we want to test $c_p = c_{p,0}$ against the two sided alternative $c_p \neq c_{p,0}$. Eventually, the specific case $c_{2,0} = 0$ is seen as the test of mono- versus multi-fractal, and the case $c_{3,0} = 0$ as a test of simple (log-Normal) against more complicated (log-Poisson or higher order) multifractal model.

9.1.2 Statistical tests

The tests are constructed for the *null hypothesis*:

$$H_0 : \quad c_p = c_{p,0}, \tag{9.1}$$

[1] We restrict our discussion to processes with finite variance, cf. Sections 2.2.3 and 2.8 (last paragraph), and to processes for which the cumulant expansion Eq. (2.70) is valid up to at least $p = 3$.

[2] This excludes for instance Lévy stable self-similar processes [88] or random wavelet series [20] which are of theoretical interest but remain difficult to use in applications.

postulating that the true value of c_p is $c_{p,0}$, against the double sided alternative $c_p \neq c_{p,0}$. They are based on the basic and on the studentized *test statistics*:

$$\hat{t}_B = \hat{c}_p - c_{p,0}, \quad \hat{t}_S = \frac{\hat{c}_p - c_{p,0}}{\hat{\sigma}^*}, \tag{9.2}$$

where $\hat{\sigma}^*$ stands for the bootstrap estimate of the standard deviation of $\hat{c}_p$. The studentized test statistic attempts to make the random variable $\hat{t}_S$ pivotal, i.e., to remove the unknown parameter σ from the distribution of $\hat{c}_p$. Large values of the observed test statistic $\hat{t}$ indicate evidence against H_0, as is common convention for statistical tests [113].

Definition. A significance $(1-\alpha)$ test rejects H_0 when the probability of observing $\hat{t}$ under the null hypothesis is smaller than α. More precisely, the test d_α is:

$$d_\alpha = \begin{cases} 1 & \text{if } \hat{t} \notin \mathcal{T}_{(1-\alpha)} \\ 0 & \text{otherwise.} \end{cases} \tag{9.3}$$

where $\mathcal{T}_{(1-\alpha)}$ is the *acceptance region* of the test. The acceptance region $\mathcal{T}_{(1-\alpha)}$ is defined by:

$$\mathsf{Pr}\{t \in \mathcal{T}_{(1-\alpha)} | P_t^{H_0}\} = 1 - \alpha, \tag{9.4}$$

based on the distribution of the test statistic under H_0, called the *null distribution*:

$$P_t^{H_0}(\tau) = \mathsf{Pr}(\hat{t} < \tau | H_0). \tag{9.5}$$

Significance and p-value. The quantity α is called the level or *significance* of the test and equals the error rate in rejecting H_0.
The critical value of α for which the observed test statistic $\hat{t}$ would be regarded as just decisive against H_0 is called the *p-value* or significance p of $\hat{t}$,

$$p = (\alpha | d_\alpha = 0,\, d_{\alpha+\varepsilon} = 1), \quad \varepsilon > 0. \tag{9.6}$$

Under H_0, the p-value ideally has a uniform distribution on $[0, 1]$, yielding its interpretation as an error rate: If $\hat{t}$ were regarded as just decisive against H_0, then this is equivalent to a procedure that rejects H_0 with error rate p [56]. It therefore quantifies the plausibility of rejecting H_0 having observed $\hat{t}$.
Power. When performing a test, one may commit two types of errors: First, one may reject H_0 when it is true (error of the first kind), which ideally happens with probability equal to the significance α. Second, one may accept $H_0 : c_p = c_{p,0}$ when it is false (error of the second kind) and a certain alternative $c_{p,A}$ is true. The probability of rejecting H_0 when it is false and a certain alternative $c_{p,A}$ is true is called the *power* β of the test against this specific alternative.
A good test should have both low probability of rejecting H_0 when it is true (i.e., small α), and high probability when H_0 is false (i.e. large β), which are antagonistic goals. It is common practice to preset the significance α and then select a test with power β as large as possible.
For more details on hypothesis tests, the reader is referred to e.g. [113].

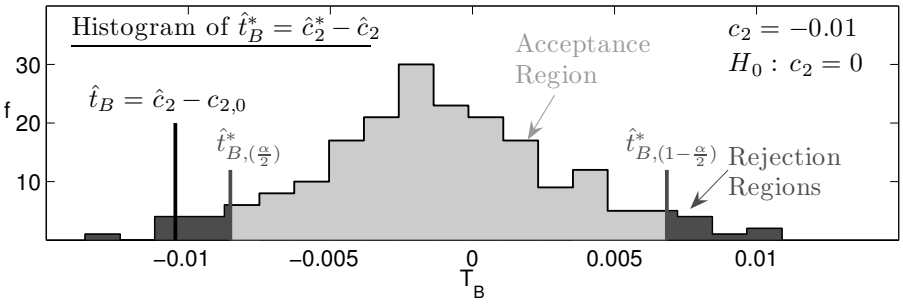


Figure 9.1: **Illustration of bootstrap hypothesis test.** Estimated null distribution and $(1-\alpha)$ basic bootstrap test for the second log-cumulant of mrw (single realization): The hypothesis H_0 is rejected if the observed value $\hat{t}$ of the test statistic is within the rejection regions, i.e., outside the interval $\widehat{\mathcal{I}}^{bas}_{(1-\alpha)} = [\hat{t}^*_{\frac{\alpha}{2}}, \hat{t}^*_{1-\frac{\alpha}{2}}]$.

9.1.3 Nonparametric bootstrap tests

The definition of the test Eq. (9.3) is based on the distribution of the test statistic under H_0 Eq. (9.5), which is unknown in most situations and in practice often approximated by parametric models. Due to the lack of plausible and robust models for H-sssi and multifractal processes, we use the nonparametric block bootstrap resampling and estimation defined in Chapter 7 to obtain estimates of the null distribution of $\hat{t}$ and construct *nonparametric bootstrap tests* [56, 82, 188, 193, 194, 195, 196].

9.1.3-a) Bootstrap null distribution estimation

The nonparametric bootstrap estimate of the null distribution Eq. (9.5) is given by the empirical distributions $\hat{P}_t^{H_0}$ of the bootstrap versions $\hat{t}^*_B$ and $\hat{t}^*_S$ of the test statistics $\hat{t}_B$ and $\hat{t}_S$ [56, 82, 195, 196], defined as:

$$\hat{t}^*_B = \hat{c}^*_p - \hat{c}_p, \tag{9.7}$$

$$\hat{t}^*_S = \frac{\hat{c}^*_p - \hat{c}_p}{\hat{\sigma}^{**}}, \tag{9.8}$$

$$\hat{P}_t^{H_0}(\tau) = \frac{1 + \sum_{r=1}^{R} h\left(\tau - \hat{t}^{*(r)}\right)}{R+1}, \tag{9.9}$$

where $h(\cdot)$ is the Heaviside (step) function. The fact that $c_{p,0}$ is replaced with $\hat{c}_p$ in the bootstrap versions $\hat{t}^*_B$ and $\hat{t}^*_S$ of the test statistics $\hat{t}_B$ and $\hat{t}_S$ ensures that their empirical distributions $\hat{P}_t^{H_0}$ approximately satisfy H_0 [82].
The null distribution estimate Eq. (9.9) can now be used for determining approximate limits of the acceptance regions $\widehat{\mathcal{I}}_{(1-\alpha)}$ Eq. (9.4), and in turn performing the test Eq. (9.3) and estimating the p-value Eq. (9.6). A typical bootstrap distribution, obtained for c_2 (for mrw) and the basic bootstrap test procedure (whose acceptance region is defined in Eq. (9.11) below) are illustrated in Fig. 9.1.

9.1.3-b) Bootstrap test acceptance regions

As we consider tests against double-sided alternatives, double-sided acceptance regions $\mathcal{I}_{(1-\alpha)}$ are used in the present work, that is, acceptance regions with finite lower and up-

per limits. There exists a large number of nonparametric bootstrap tests in the literature, producing different acceptance regions (cf. [56] for an overview). We analyze 6 different significance α bootstrap tests, including simple, computationally inexpensive, and more sophisticated, computationally more demanding double bootstrap methods: The *asymptotic* bootstrap test uses simple symmetric acceptance regions, employing only the bootstrap standard deviation estimations. The *basic* and *percentile* tests employ quantiles of the empirical bootstrap distributions Eq. (9.9) of $\hat{t}_B^*$. The *studentized* test uses quantiles of the empirical distributions of the pivoted test statistic $\hat{t}_S^*$. The *adjusted basic* and *adjusted percentile* tests use the double bootstrap estimations to correct for a bias in the limits of the acceptance regions of the basic and percentile tests. The three latter methods are potentially more efficient, however at the price of a costly double bootstrap layer.

Asymptotic (Normal) bootstrap test. Assuming $\hat{t}_B$ (Eq. (9.2)) to be approximately Normal under H_0, the bootstrap standard deviation estimate $\hat{\sigma}^*$ is used to construct the equi-tailed and symmetric acceptance region:

$$\widehat{\mathcal{T}}_{(1-\alpha)}^{nor} = [q_{(\frac{\alpha}{2})}\hat{\sigma}^* + c_{p,0}, -q_{(\frac{\alpha}{2})}\hat{\sigma}^* + c_{p,0}], \tag{9.10}$$

where q_α is the α quantile of the standard Normal distribution.

Basic bootstrap test. The bootstrap distribution Eq. (9.9) of $\hat{t}_B^*$ is used directly to define the equi-tailed acceptance region:

$$\widehat{\mathcal{T}}_{(1-\alpha)}^{bas} = [\hat{t}_{B,(\frac{\alpha}{2})}^*, \hat{t}_{B,(1-\frac{\alpha}{2})}^*], \tag{9.11}$$

where $\hat{t}_{B,(\alpha)}^*$ is the empirical α-quantile of Eq. (9.9) for $\hat{t}_B^*$.

Percentile bootstrap test. The percentile test is constructed by inversion of a percentile confidence interval $[\hat{c}_{p,(\frac{\alpha}{2})}^*, \hat{c}_{p,(1-\frac{\alpha}{2})}^*]$ (Eq. (8.6)) for the parameter c_p, and has acceptance region:

$$\widehat{\mathcal{T}}_{(1-\alpha)}^{per} = [-\hat{t}_{(1-\frac{\alpha}{2})}^*, -\hat{t}_{(\frac{\alpha}{2})}^*]. \tag{9.12}$$

Studentized bootstrap test. The studentized test is a basic bootstrap test for the pivoted test statistic $\hat{t}_S$ (Eq. (9.2)). The method thus demands a double bootstrap for calculating the standard deviation estimates $\hat{\sigma}^{**}$ and has acceptance region:

$$\widehat{\mathcal{T}}_{(1-\alpha)}^{stu} = [\hat{t}_{S,(\frac{\alpha}{2})}^*, \hat{t}_{S,(1-\frac{\alpha}{2})}^*]. \tag{9.13}$$

Adjusted p-value for basic bootstrap test. If the usual error rate interpretation of p is to be valid, the p-value must be uniformly distributed on $[0, 1]$ under H_0. This is, however, not guaranteed for composite null hypotheses, neither for approximate null models as in Eq. (9.9). The adjusted p-value method aims at estimating an improved p-value that is more nearly uniformly distributed than the unadjusted one. It treats p as the observed test statistic and estimates its distribution by resampling under the null model [56]. The double-sided adjusted p-value is given by:

$$p_{adj}^{bas} = 2\min\left(\mathsf{Pr}\{p^{bas*} \leq p^{bas}|\hat{F}_\theta\}, \mathsf{Pr}\{p^{bas*} > p^{bas}|\hat{F}_\theta\}\right). \tag{9.14}$$

Here, p^{bas} is the p-value of the basic bootstrap test, and the p^{bas*} are its bootstrap resamples, obtained by double bootstrap. The acceptance region is given by:

$$\widehat{\mathcal{T}}_{(1-\alpha)}^{adjbas} = \{t : p_{adj}^{bas} \geq \alpha\}. \tag{9.15}$$

$$d_\alpha = \begin{cases} 1 & \text{if } \hat{t} \notin \widehat{\mathcal{T}}_{(1-\alpha)} \\ 0 & \text{otherwise.} \end{cases} \qquad p = (\alpha | d_\alpha = 0,\, d_{\alpha+\varepsilon} = 1)\,,\ \varepsilon > 0$$

Normal:	$\widehat{\mathcal{T}}^{nor}_{(1-\alpha)} = [q_{(\frac{\alpha}{2})}\hat{\sigma}^* + c_{p,0}, -q_{(\frac{\alpha}{2})}\hat{\sigma}^* + c_{p,0}]$
Basic:	$\widehat{\mathcal{T}}^{bas}_{(1-\alpha)} = [\hat{t}^*_{B,(\frac{\alpha}{2})}, \hat{t}^*_{B,(1-\frac{\alpha}{2})}]$
Percentile:	$\widehat{\mathcal{T}}^{per}_{(1-\alpha)} = [-\hat{t}^*_{(1-\frac{\alpha}{2})}, -\hat{t}^*_{(\frac{\alpha}{2})}]$
Studentized:	$\widehat{\mathcal{T}}^{stu}_{(1-\alpha)} = [\hat{t}^*_{S,(\frac{\alpha}{2})}, \hat{t}^*_{S,(1-\frac{\alpha}{2})}]$
Adjusted Basic:	$\widehat{\mathcal{T}}^{adjbas}_{(1-\alpha)} = \{t : p^{bas}_{adj} \geq \alpha\}$
Adjusted Percentile:	$\widehat{\mathcal{T}}^{adjper}_{(1-\alpha)} = \{t : p^{per}_{adj} \geq \alpha\}$

Table 9.1: **Overview of significance α bootstrap tests.** Overview of significance α bootstrap tests and their corresponding acceptance regions Eqs. (9.10–9.13), (9.15) and (9.16).

Adjusted p-value for percentile bootstrap test. The adjusted p-value for the percentile bootstrap test is given by Eq. (9.14) by replacing p^{bas} and p^{bas*} with the p-value of the percentile bootstrap test p^{per} and its corresponding bootstrap resamples p^{per*}, respectively. It has acceptance region:

$$\widehat{\mathcal{T}}^{adjper}_{(1-\alpha)} = \{t : p^{per}_{adj} \geq \alpha\}. \tag{9.16}$$

The bootstrap tests, p-values and acceptance regions considered in this work are summarized in Tab. 9.1.

Hypothesis tests versus confidence intervals. There is a form of duality between hypothesis tests for parameters and confidence intervals for these parameters in the sense that – for a prescribed level α – a confidence region includes parameters that are not rejected by an appropriate hypothesis test [56]. In this sense, the percentile test acceptance region $\widehat{\mathcal{T}}^{per}_{(1-\alpha)}$ is obtained by inversion of a percentile confidence limit Eq. (8.6) for c_p, and the studentized test acceptance region $\widehat{\mathcal{T}}^{stu}_{(1-\alpha)}$ by inversion of a studentized confidence region Eq. (8.7) for c_p, and the tests reject H_0 if $c_{p,0}$ is outside the corresponding confidence regions.

9.1.3-c) Advanced comments on hypothesis tests

For sake of completeness, let us further note that α, p and β are only properly defined if the null hypothesis is *simple*, i.e. when the null distribution $P_t^{H_0}$ is completely specified by H_0. In most parametric and all nonparametric situations, H_0 is *composite* and whole families of distributions satisfy H_0. Then, the quantities α, p and β can not be well defined [113]. Therefore, the nonparametric bootstrap test constitutes an approximation, consist-

ing in the use of one distribution out of the family of distributions satisfying H_0 for test construction.

9.2 Validation of bootstrap tests on log-cumulants

9.2.1 Numerical simulation study

We evaluate the statistical performance of the proposed bootstrap test procedures by applying them to a large number N_{MC} of realizations of four different synthetic stochastic processes with a priori known and controlled multifractal properties and log-cumulant values c_p. For each realization, each test procedure defined above provides us with 2 outputs: The decision $\hat{d}_\alpha$, and the p-value of the observed test statistic. From averages over realizations, we evaluate the actual significances, p-values and powers of the tests and compare them with the theoretical targets, and against each other. The aim of this numerical study is to address the following issues: Do the bootstrap test procedures described above exhibit satisfactory statistical performance? Should one prefer wavelet coefficients or wavelet Leaders for testing mono- vs. multi-fractal, and for testing different multifractal models? What precise design of the acceptance region Eqs. (9.10–9.13), (9.15) and (9.16) yields the best statistical performance?

9.2.2 Performance under H_0: Actual significance and p-value

A first set of experiments is run to evaluate the actual significances and p-values of the procedures. For that, we test the hypothesis $H_0 : c_p = c_{p,0}$ when this hypothesis is true. We obtain N_{MC} estimates $\hat{d}_\alpha$ and $\hat{p}$ for each of the proposed tests and nominal α.
The actual significances $\hat{\alpha}_{MC}$ of the tests are estimated as ($\hat{\mathbb{E}}_{MC}$ denotes the average over Monte Carlo realizations):

$$\hat{\alpha}_{MC} = \hat{\mathbb{E}}_{MC}\{\hat{d}_\alpha | c_p \equiv c_{p,0}\} \tag{9.17}$$

and should ideally equal the nominal significance α.
The actual p-values $\hat{p}$ should ideally be uniformly distributed on $[0;1]$, if the null distribution estimation is appropriate and sufficiently accurate. Therefore, the average actual p-values $\hat{p}_{MC}$ of the tests, estimated as:

$$\hat{p}_{MC} = \hat{\mathbb{E}}_{MC}\{\hat{p} | c_p \equiv c_{p,0}\}, \tag{9.18}$$

should be close to $1/2$, and the empirical distributions of $\hat{p}$ should be uniform.

Note that due to the duality of confidence interval and test acceptance region constructions for bas, per, nor and stu, the actual significances $\hat{\alpha}_{MC}$ for tests $H_0 : c_{p,0} = c_p$ can be read as empirical coverages of the corresponding bootstrap confidence intervals as:

$$\mathcal{C}_{c_p,(1-\alpha)} = 100 \cdot (1 - \hat{\alpha}_{MC})\%. \tag{9.19}$$

For the same reason, any of the empirical coverage results for $\theta = c_p$ in Chapter 8 can be read as actual significances of the equivalent tests for $H_0 : c_{p,0} = c_p$:

$$\hat{\alpha}_{MC} = 1 - \frac{1}{100}\mathcal{C}_{c_p,(1-\alpha)}, \tag{9.20}$$

and therefore, also the corresponding conclusions in Chapter 8 apply to the equivalent bootstrap hypothesis tests under H_0.

c_2 fBm	d_X - $N = 2^{12}$					d_X - $N = 2^{15}$				
α nominal	**0.05**	**0.10**	**0.15**	**0.20**	**0.25**	**0.05**	**0.10**	**0.15**	**0.20**	**0.25**
nor	0.10	0.18	0.23	0.29	0.35	0.07	0.13	0.18	0.24	0.29
bas	0.08	0.16	0.23	0.30	0.34	0.07	0.13	0.18	0.23	0.29
per	0.12	0.19	0.26	0.31	0.37	0.08	**0.12**	0.18	0.24	0.30
stu	0.08	0.16	0.23	0.29	0.34	**0.06**	**0.12**	**0.17**	**0.22**	**0.27**
adjb	**0.07**	**0.13**	**0.20**	**0.24**	**0.29**	0.10	0.15	0.19	0.26	0.30
adjp	0.11	0.17	0.22	0.26	0.30	0.10	0.15	0.19	0.25	0.31
c_2 fBm	L_X - $N = 2^{12}$					L_X - $N = 2^{15}$				
α nominal	**0.05**	**0.10**	**0.15**	**0.20**	**0.25**	**0.05**	**0.10**	**0.15**	**0.20**	**0.25**
nor	0.13	0.20	0.27	0.32	0.38	0.18	0.27	0.34	0.40	0.46
bas	0.17	0.22	0.29	0.34	0.40	0.24	0.34	0.41	0.48	0.54
per	0.14	0.21	0.28	0.34	0.40	**0.12**	0.21	0.28	0.34	0.39
stu	**0.12**	0.19	0.26	0.30	0.34	0.20	0.30	0.38	0.44	0.50
adjb	0.14	0.19	0.22	0.28	0.34	0.30	0.33	0.40	0.46	0.51
adjb	**0.12**	**0.15**	**0.19**	**0.25**	**0.30**	0.16	**0.17**	**0.24**	**0.29**	**0.34**

Table 9.2: **Actual significances of bootstrap tests on c_2: fBm.** Actual significances $\hat{\alpha}_{MC}$ of bootstrap tests for $c_{2,0} = c_2$, using wavelet coefficients and Leaders, for fBm ($c_{2,0} = c_2 = 0$). Nominal significances and results closest to nominal values are marked in bold.

9.2.3 Performance under H_A: Power

A second set of experiments is run to evaluate the powers of the procedures. For that, we test the hypothesis $H_0 : c_p = c_{p,0}$ when an alternative $H_A : c_p = c_{p,A} \neq c_{p,0}$ is true. The actual powers of the tests on $H_0 : c_p = c_{p,0}$ against a certain alternative $c_{p,A} \neq c_{p,0}$ are estimated as:

$$\hat{\beta}_{MC}(c_{p,A}, \alpha) = \hat{\mathbb{E}}_{MC}\{\hat{d}_\alpha | c_{p,0}, c_p = c_{p,A}\}, \tag{9.21}$$

and the larger, the better. With $c_p = c_{p,0} = 0$ and $p = 2$, this essentially evaluates the ability of the tests to reject self-similarity in favor of multifractal, and with $p = 3$, to reject a simple (log Normal) multifractal model against a more complicated one.

9.3 Performance of Bootstrap Tests

9.3.1 Testing monofractal versus multifractal model

The goal of this section is to validate tests $H_0 : c_2 = 0$ aiming at rejecting the monofractal model. To this end, we make use of fractional Brownian motion (fBm, cf. Section 2.2.3-a)) and multifractal random walk (mrw, cf. Section 2.7.2-b)), chosen because they provide us with simple yet representative examples of monofractal (Gaussian H-sssi) processes (for fBm, $c_2 \equiv 0$ and $H_0 : c_2 = c_{2,0} = 0$ is true), and non Gaussian multifractal processes (for mrw, $c_2 \neq 0$ and $H_0 : c_2 = c_{2,0} = 0$ is not true), respectively. This work has been reported in [174].

Experimental set-up. Results are obtained for sample sizes $N \in \{2^{12}, 2^{15}\}$, using Daubechies wavelets with $N_\psi = 3$ vanishing moments. Estimation is performed

c_2	fBm			
	d_X		L_X	
	2^{12}	2^{15}	2^{12}	2^{15}
nor	0.43	0.48	0.42	0.36
bas	0.44	0.48	0.39	0.32
per	0.42	0.47	0.41	0.41
stu	0.44	**0.49**	0.42	0.34
$adjb$	**0.46**	0.46	0.43	0.31
$adjp$	0.44	0.46	**0.47**	**0.42**

Table 9.3: **Mean p-value of bootstrap tests on c_2 for fBm.** Mean p-value $\hat{p}_{MC}$ of bootstrap tests for $c_{2,0} = c_2$, using wavelet coefficients and Leaders for fBm ($c_{2,0} = c_2 = 0$). Results closest to the theoretical value $\mathbb{E}\, p = 1/2$ are marked in bold.

with weighted (w^1) linear regressions over scales $(j_1, j_2) = (3, \log_2(N) - 4)$. Bootstrap estimations are based on T-B bootstrap (cf. Section 7.2.1) with $R = 599$, $S = 50$ for $N = 2^{12}$, and $R = 399$, $S = 25$ for $N = 2^{15}$. Nominal significances are set to $\alpha \in \{0.05, 0.1, 0.15, 0.2, 0.25\}$. The process parameter for fBm is set to $H = c_1 = 0.8$. For the simulations of the powers against multiple alternatives (cf. Fig. 9.2 and Tab. 9.4), we make use of mrw for the range of parameter settings $\beta^2 = -c_2 \in \{0.01, 0.02, \cdots, 0.08\}$ ($H = 0.72$).

Significance and p-value. Tab. 9.2 summarizes actual significances for tests on c_2 for fBm ($c_{2,0} = c_2 = 0$). The numerical results indicate that overall, $\hat{\alpha}_{MC}$ is satisfactorily close to nominal α for the proposed methods. The tests employing wavelet coefficients reproduce the nominal α slightly better than those using wavelet Leaders, in particular for large sample size. Also, while for tests based on wavelet coefficients, no clear preference can be given to any of the particular acceptance regions, the per and the $adjp$ method perform slightly better than the others for tests based on wavelet Leaders.

The average actual p-values $\hat{p}_{MC}$ of the tests are summarized in Tab. 9.3 for c_2. They suggest that the expected uniform (mean $1/2$) distributions are satisfactorily reproduced for all acceptance regions. Therefore, the error rate interpretation of the estimate $\hat{p}$ of p is approximately valid for the proposed procedures. Also, while the adjusted methods generally improve results – as expected – for small sample size $N = 2^{12}$, they appear to be much less decisive for larger sample size. This may be due to the smaller number of double bootstrap resamples ($S = 25$) used in the latter case.

These results (actual significances close to nominal, p-values approximately uniformly distributed) indicate that the nonparametric bootstrap null distribution estimations Eq. (9.9) are valid approximations to the real null distributions of the test statistics.

Power. Tab. 9.4 summarizes the powers $\hat{\beta}_{MC}(c_{2,A}, \alpha)$ of the bootstrap tests $H_0 : c_2 = 0$ for the particular alternative $H_A : c_2 = c_{2,A} = -0.08$. It shows that wavelet Leader based tests on c_2 are much more powerful than wavelet coefficient based ones: Whereas the tests employing wavelet Leaders achieve powers up to 0.7 for $N = 2^{12}$ and small α, and approximately 1 for $N = 2^{15}$, wavelet coefficient based procedures have only powers of 0.2 and 0.55, respectively. As expected, test power increases with increasing sample size N. Tab. 9.4 also shows that the percentile and adjusted percentile methods have consistently the largest powers.

$\hat{\beta}_{MC}(c_{2,A},\alpha)$	d_X - $N=2^{12}$					d_X - $N=2^{15}$				
α nominal	0.05	0.10	0.15	0.20	0.25	0.05	0.10	0.15	0.20	0.25
nor	0.16	0.25	0.31	0.38	0.42	0.45	0.55	0.62	0.67	0.71
bas	0.13	0.21	0.28	0.35	0.40	0.40	0.50	0.57	0.62	0.66
per	**0.22**	**0.30**	**0.36**	**0.41**	**0.47**	0.48	0.59	**0.66**	**0.71**	**0.75**
stu	0.13	0.18	0.25	0.32	0.37	0.29	0.41	0.50	0.57	0.61
$adjb$	0.12	0.18	0.24	0.30	0.35	0.46	0.53	0.58	0.63	0.66
$adjp$	0.19	0.27	0.33	0.38	0.43	**0.55**	**0.60**	**0.66**	**0.71**	0.74
$\hat{\beta}_{MC}(c_{2,A},\alpha)$	L_X - $N=2^{12}$					L_X - $N=2^{15}$				
α nominal	0.05	0.10	0.15	0.20	0.25	0.05	0.10	0.15	0.20	0.25
nor	0.61	0.71	0.78	0.82	0.85	0.97	**0.99**	**0.99**	0.99	0.99
bas	0.53	0.63	0.69	0.74	0.78	0.96	0.97	0.98	0.99	0.99
per	**0.70**	**0.80**	**0.84**	**0.87**	**0.89**	0.98	**0.99**	**0.99**	**1.00**	**1.00**
stu	0.45	0.56	0.62	0.67	0.71	0.90	0.94	0.96	0.97	0.98
$adjb$	0.49	0.56	0.61	0.67	0.71	0.97	0.97	0.98	0.98	0.99
$adjp$	0.66	0.71	0.76	0.81	0.84	**0.99**	**0.99**	**0.99**	0.99	0.99

Table 9.4: **Power of bootstrap tests of** $H_0 : c_2 \equiv 0$**.** Actual power $\hat{\beta}_{MC}(c_{2,A},\alpha)$ vs. significance of bootstrap tests of $H_0 : c_2 \equiv 0$ for mrw ($c_{2,A}$ = c_2 =-0.08), using wavelet coefficients (top) and Leaders (bottom). Best results are marked in bold.

The superiority of the wavelet Leader based procedures in terms of test power is also clearly demonstrated in Fig. 9.2, summarizing the powers of the bootstrap tests bas (top, for $N = 2^{12}$ and $N = 2^{15}$, $R = 199$), per and stu (bottom, $N = 2^{15}$), $R = 199, S = 50$) of $H_0 : c_2 = 0$ for a set of alternatives $c_2 = c_{2,A} \in \{-0.08, -0.07, \cdots, -0.01\}$. For $N = 2^{12}$, the wavelet Leader based bas test achieves practically useful power as soon as $c_{2,A} \leq -0.05$, whereas the powers of wavelet coefficient based procedure do not differ by much from the significance α for the whole range of alternatives. Similarly, for $N = 2^{15}$, the powers of the wavelet Leader based tests approach 1 relatively quickly as c_2 departs from zero and remain large over a wide range of alternatives, whereas those of wavelet coefficient based procedures remain low – of the order of the power of wavelet Leader based tests for the (one order of magnitude) smaller sample size $N = 2^{12}$. Also, and most importantly for applications, Fig. 9.2 indicates that the powers of the wavelet Leader based tests remain significantly above those of the wavelet coefficient based tests for alternatives close to the null value $c_{2,0} = 0$. Finally, Tab. 9.4 suggests that the per tests are consistently more powerful than the bas and the double bootstrap stu tests.

Conclusions. We conclude that both wavelet Leader and coefficient based tests of mono- vs. multi-fractal satisfactorily reproduce target significance. Yet, the wavelet Leader based procedures significantly outperform wavelet coefficient based ones in terms of test power. Therefore, tests of mono- vs. multi-fractal models should be based on wavelet Leaders. Also, the simple per test, having overall best performance, is in practice to be preferred over computationally more demanding double bootstrap tests. For this reason, we will mainly concentrate on the per test in the rest of this chapter.

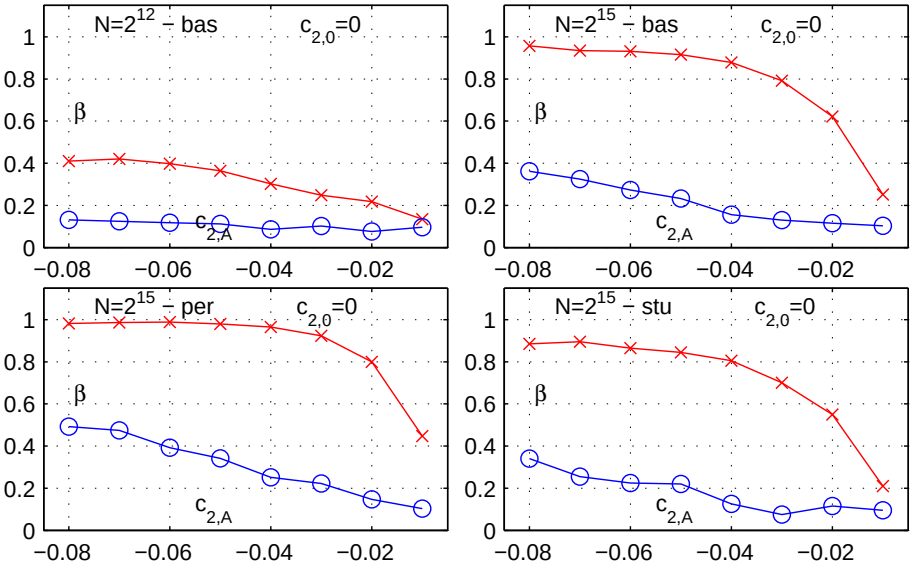


Figure 9.2: **Power of bootstrap tests of $H_0 : c_2 \equiv 0$.** Actual power $\hat{\beta}_{MC}(c_{2,A}, \alpha)$ of bootstrap test of $H_0 : c_2 \equiv 0$ for mrw against various alternatives $c_{2,A} = c_2$, obtained for nominal significances $\alpha = 0.05$: basic test (top) for sample size $N = 2^{12}$ (top left) and $N = 2^{15}$ (top right), percentile test (bottom left) and studentized test (bottom right) for sample size $N = 2^{15}$. Blue solid lines and '$\circ$' stand for wavelet coefficients, red solid lines and '$\times$' for wavelet Leaders based tests.

9.3.2 Testing c_2 for a multifractal process

The aim of this section is to investigate whether the proposed bootstrap tests on c_2 reproduce nominal performance for multifractal processes, that is, when the null hypothesis $c_{2,0} = c_2 \neq 0$ is true. To this end, we make use of multifractal random walk (mrw, cf. Section 2.7.2-b)), with process parameters set to $(H, \beta) = (0.72, \sqrt{0.08})$, i.e. $(c_1, c_2) = (0.8, -0.08)$. We use the experimental set-up as described in Section 9.3.1.

Results for test performance under $H_0 : c_2 = -0.08$ are given in Tab. 9.5 (actual significances), Tab. 9.6 (average actual p-values), and Fig. 9.3 (empirical distributions of p-values). Overall, they lead to the same conclusions as those obtained in the previous section for test performance under H_0 for fBm: Wavelet coefficients and Leaders based tests reproduce target significance equivalently well, both having actual significances satisfactorily close to nominal α. No clear preference can be given to any of the particular acceptance regions for wavelet coefficients, yet per and the stu perform slightly better for wavelet Leaders based tests. Also, the average actual p-values $\hat{p}_{MC}$ of the tests reproduce the expected uniform (mean $1/2$) distributions satisfactorily, hence approximate validity of the error rate interpretation of the estimate $\hat{p}$ of p, and practically satisfactory nonparametric bootstrap null distribution estimations Eq. (9.9).

Conclusions. We conclude that both wavelet Leader and coefficient based tests on c_2 satisfactorily reproduce nominal significance under H_0 for multifractal mrw. This further emphasizes the relevance of the test for mono- versus multi-fractal proposed in the previous section. Additional results for significance and power of tests on c_2, obtained for a different multifractal process (CPM-MF-fBm), are reported in Tab. 9.7 and lead to

c_2 mrw	d_X - $N = 2^{12}$					d_X - $N = 2^{15}$				
α nominal	**0.05**	**0.10**	**0.15**	**0.20**	**0.25**	**0.05**	**0.10**	**0.15**	**0.20**	**0.25**
nor	**0.06**	0.12	0.17	0.22	0.28	**0.05**	**0.10**	**0.15**	**0.20**	**0.25**
bas	0.07	0.12	0.18	0.24	0.29	**0.05**	0.12	0.18	0.23	0.29
per	0.07	0.13	0.19	0.24	0.30	**0.05**	**0.10**	0.14	0.19	0.24
stu	0.07	0.12	0.17	0.23	0.27	0.04	**0.10**	0.16	0.22	0.28
adjb	**0.06**	**0.09**	0.14	0.19	0.24	0.08	0.13	0.18	0.24	0.28
adjp	0.07	**0.11**	**0.15**	**0.20**	**0.25**	0.08	**0.10**	0.14	**0.20**	0.24
c_2 mrw	L_X - $N = 2^{12}$					L_X - $N = 2^{15}$				
α nominal	**0.05**	**0.10**	**0.15**	**0.20**	**0.25**	**0.05**	**0.10**	**0.15**	**0.20**	**0.25**
nor	0.06	0.12	0.17	0.23	0.27	0.07	0.14	0.19	0.24	0.30
bas	0.10	0.16	0.21	0.26	0.31	0.10	0.18	0.24	0.30	0.36
per	**0.05**	**0.10**	**0.14**	**0.20**	**0.25**	**0.05**	**0.10**	**0.16**	**0.21**	**0.26**
stu	0.06	0.11	**0.16**	0.21	**0.25**	0.06	0.12	0.18	0.23	0.28
adjb	0.09	0.12	**0.16**	0.19	0.23	0.15	0.16	0.23	0.26	0.32
adjp	0.04	0.06	0.09	0.12	0.16	0.08	0.08	**0.14**	0.16	0.22

Table 9.5: **Actual significances of bootstrap tests on c_2: mrw.** Actual significances $\hat{\alpha}_{MC}$ of bootstrap tests for $c_{2,0} = c_2$, using wavelet coefficients and Leaders, for mrw ($c_{2,0} = c_2 = -0.08$). Nominal significances and results closest to nominal values are marked in bold.

similar conclusions.

9.3.3 Testing the multifractal log-Normal model

Applying the bootstrap test procedures with the choice of null hypothesis $H_0 : c_3 = c_{3,0} \equiv 0$ provides indications for deciding whether a log-Normal multifractal process or a more sophisticated multifractal model is to be used to describe the data under study. Indeed, for a log-Normal process, such as mrw, $\forall p \geq 3, c_p = 0$. Although in itself, $c_3 = 0$ does not prove that data follow a log-Normal process, this is a very valuable and practically useful information (cf. [179] and Chapter 12). Therefore, the goal of this section is to study the

c_2	mrw			
	d_X		L_X	
	2^{12}	2^{15}	2^{12}	2^{15}
nor	0.47	0.49	0.49	0.47
bas	0.47	0.48	0.46	0.43
per	0.47	**0.50**	**0.50**	**0.50**
stu	0.48	0.49	**0.50**	0.47
adjb	0.49	0.46	**0.50**	0.44
adjp	**0.50**	0.49	0.56	0.52

Table 9.6: **Mean p-value of bootstrap tests on c_2.** Mean p-value $\hat{p}_{MC}$ of bootstrap tests for $c_{2,0} = c_2$, using wavelet coefficients and Leaders for mrw ($c_{2,0} = c_2 = -0.08$). Results closest to the theoretical value $\mathbb{E}\, p = 1/2$ are marked in bold.

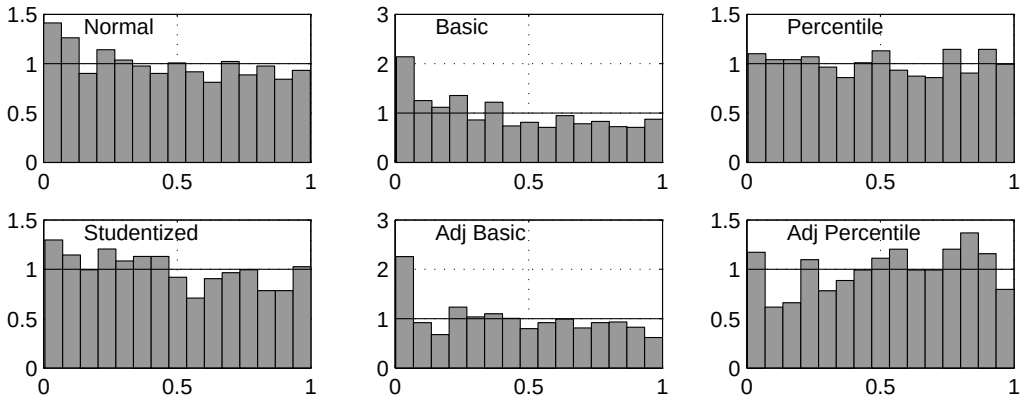


Figure 9.3: **Empirical distributions of p-values of bootstrap tests for mrw.** Empirical distributions of the p-values $\hat{p}$ of the bootstrap tests for $c_{2,0} = c_2$, obtained for mrw ($c_{2,0} = c_2 = -0.08$, $N = 2^{15}$) using Leaders.

relevance and statistical performance of such tests and to investigate whether they can detect departure from zero values for c_3. To this end, we use the multifractal process CPM-MF-fBm (see Section 2.7.2-a)), whose $c_3 \neq 0$ can be set a priori. This work has been reported in [179].

Experimental set-up. Results are obtained for sample size $N = 2^{15}$, using Daubechies wavelets with $N_\psi = 3$ vanishing moments and weighted (w^1) linear regressions over the range of scales $(j_1, j_2) = (3, \log_2(N) - 4)$. Bootstrap estimations use T-B bootstrap (cf. Section 7.2.1) with $R = 399$, $S = 50$, and nominal significance is set to $\alpha = 0.05$. The parameters of CPM-MF-fBm are fixed such that $(c_1, c_2, c_3) = (0.8, -0.08, 0.0311)$.

Results. Results for actual significances $\hat{\alpha}_{MC}$ (by setting $c_{3,0} = c_3 = 0.0311$) and powers $\hat{\beta}_{MC}$ ($c_{3,0} = 0$) are summarized in Tab. 9.7 (right). They indicate that for any of the tests on c_3, the nominal significance of $\alpha = 0.05$ is satisfactorily reproduced. The wavelet Leader based tests have actual significances slightly closer to nominal α than their wavelet coefficient based counterparts, and are hence more reliable for testing c_3. The difference between per and stu test actual significances is not decisive.
Second, the results demonstrate that wavelet Leader based tests of $H_0 : c_3 = 0$ against $c_{3,A} \neq 0$ are much more powerful than wavelet coefficient based ones: Whereas wavelet coefficient based test powers practically equal the actual significance, wavelet Leader based test powers can in practice allow to detect true departure from zero values for c_3. This is in particular the case for the per test, which is 30% more powerful than stu.
Finally, the results demonstrate that detection of true departure from zero is much more difficult for c_3 than for c_2: While the wavelet Leader based per test for $H_0 : c_2 = 0$ has power larger than 0.9 for the alternative $c_{2,A} = -0.03$ (cf. Fig. 9.2, bottom left), the corresponding test for $H_0 : c_2 = 0$ has power smaller than 0.4 for the alternative $c_{3,A} = 0.0311$.

9.3.4 Non Gaussian finite variance H-sssi processes and bootstrap tests

The results and discussions of the previous two sections demonstrate that the nonparametric bootstrap tests defined in Section 9.1.3 are efficient for discriminating between

CPM-MF-fBm		$c_2 = -0.08$				$c_3 = 0.0311$			
$N = 2^{15}$		d_X		L_X		d_X		L_X	
H_0		per	stu	per	stu	per	stu	per	stu
$c_{p,0} = c_p$	$\hat{\alpha}_{MC}$	0.07	**0.05**	**0.05**	**0.05**	**0.10**	**0.10**	0.07	**0.06**
$c_{p,0} = 0$	$\hat{\beta}_{MC}$	**0.57**	0.40	**1.00**	0.96	**0.13**	0.11	**0.39**	0.28

Table 9.7: **Actual significance and power of bootstrap tests for CPM-MF-fBm.** Actual significances $\hat{\alpha}_{MC}$ (top) and powers $\hat{\beta}_{MC}(c_{p,A}, \alpha)$ (bottom) for CPM-MF-fBm and $p = 2$ (left) and $p = 3$ (right). The nominal significance is $\alpha = 0.05$. Best results are marked in bold.

Gaussian finite variance H-sssi and non Gaussian MF processes (tests on c_2), as well as between simple and more elaborate multifractal models (tests on c_3). In the present section, we further investigate the performance of the test procedures when applied to *non Gaussian* finite variance H-sssi processes, and compare them to those for Gaussian fBm. The Rosenblatt process (ROS, see Section 2.2.3-b)) will be used as a very example for such processes. The goal of this study is to validate whether the proposed test procedures remain reliable for non Gaussian H-sssi processes, and whether they can discriminate non Gaussian H-sssi from non Gaussian multifractal processes. For sake of completeness, we also report results for tests on c_1. This work has been reported in [175].

Experimental set-up. The bootstrap test procedures are applied to $N_{MC} = 500$ realizations of length $N = 2^{15}$ of fBm and ROS with $H = 0.7$. Results are obtained using Daubechies wavelets with $N_\psi = 3$ and weighted (w^1) regressions over the range of scales $(j_1, j_2) = (3, 8)$. Bootstrap estimation is performed using both T-B and TS-B bootstrap (Sections 7.2.1 and 7.2.2, respectively) and $R = 199$ bootstrap samples per realization. The nominal significance is set to $\alpha = 0.05$.

Significance. For H-sssi processes, the bias can be dominating over variability in estimation performance for Leaders (cf. Chapter 4, e.g. Tab. 4.5 or 4.6), hence potentially altering actual significances of the tests. Therefore, we prefer in this section, in analogy to bias-corrected empirical coverage for confidence intervals (cf. Eq. (8.19)), the alternative actual significance definition $\hat{\alpha}_{MC} = \hat{\mathbb{E}}_{MC}\{\hat{d}_\alpha | c_{p,0} = \hat{\mathbb{E}}_{MC}\hat{c}_p\}$.
Results are reported in Tab. 9.8 and indicate that although actual significances are overall satisfactory for both processes, it is closer to nominal value α for fBm than for ROS. This is in particular true for c_1, for which actual significances are relatively far from the nominal value for ROS. For c_2 and c_3, the key attributes in discriminating mono- vs. multi-fractal models, the difference in actual significance between fBm and ROS is smaller. Second, wavelet coefficient based tests reproduce nominal significance closer than wavelet Leader based procedures for both H-sssi processes and in particular so for ROS. This is consistent with findings on confidence interval coverage for fBm (cf. Section 8.3.4). Finally, Tab. 9.8 suggests that TS-B bootstrap results in more accurate tests (actual significances closer to nominal) than T-B for wavelet Leaders, whereas there is such no difference for wavelet coefficients.

Power. Fig. 9.4 shows rejection rates (powers) $\hat{\beta}_{MC}$ of tests $H_0 : c_p = c_{p,0}$ for fBm (top row) and ROS (bottom row) for both wavelet Leaders (solid red lines) and wavelet coeffi-

cients (dashed blue lines), based on TS-B bootstrap resampling. The powers $\hat{\beta}_{MC}$ of the tests equal the probabilities of rejecting $c_{p,0}$ (given by the abscissa) when the alternatives c_p (i.e. $(c_1, c_2, c_3) = (0.7, 0, 0)$) are true. Thus, for $p = 2, 3$, $\hat{\beta}_{MC}$ quantifies the capability of the test to reject a multifractal hypothesis when the data are self-similar fBm or ROS. Fig. 9.4 demonstrates that wavelet Leader based tests display larger powers than wavelet coefficient based ones, and significantly so for the parameters c_2 and c_3 which discriminate mono- from multi-fractal, and simple from more elaborate multifractal models, respectively. However, switching from Gaussian fBm to non Gaussian ROS results in a non negligible decrease in test powers. Therefore, it is significantly more difficult to reject a multifractal hypothesis when the data are non Gaussian H-sssi processes than when they are Gaussian H-sssi.

Conclusions. From actual significance results, we conclude that although test performance remains satisfactory, bootstrap estimation is more difficult for the non-Gaussian Rosenblatt process than for fBm, resulting in less accurate nonparametric bootstrap tests. Alternative and additional results on bootstrap estimation and confidence interval estimation performance for fBm and ROS have been obtained by us and reported in [175] and lead to similar conclusions: Bootstrap estimation is more difficult and less accurate for non Gaussian than for Gaussian H-sssi processes. These results motivate studying the dependence structure of wavelet coefficients and Leaders for non Gaussian H-sssi processes. This is discussed in Chapter 11.
Irrespectively of the efficiency of bootstrap procedures, the non Gaussian nature of ROS causes a significant decrease in test powers (equivalently, increase in confidence interval sizes), resulting in increased difficulties in discriminating between multifractal and non Gaussian self-similar models. This is consistent with findings in Chapter 4, demonstrating substantially better estimation performance for fBm than for ROS (see e.g. Tab. 4.1).
Finally, wavelet Leader based tests do have significantly larger powers than wavelet coefficient based ones also for non Gaussian H-sssi processes.

fBm	d_X		L_X	
bias corrected	T-B	TS-B	T-B	TS-B
c_1	0.06	0.06	0.12	0.08
c_2	0.06	0.08	0.07	0.04
c_3	0.07	0.09	0.05	0.06
ROS	d_X		L_X	
bias corrected	T-B	TS-B	T-B	TS-B
c_1	0.12	0.11	0.24	0.14
c_2	0.06	0.07	0.15	0.11
c_3	0.08	0.09	0.15	0.12

Table 9.8: **Bootstrap test significances for fBm and ROS.** Actual significances $\hat{\alpha}_{MC}$ of T-B and TS-B based bootstrap bootstrap tests for wavelet coefficient (left) and Leader (right) based estimation, for fBm (top) and ROS (bottom). The nominal significance is $\alpha = 0.05$.

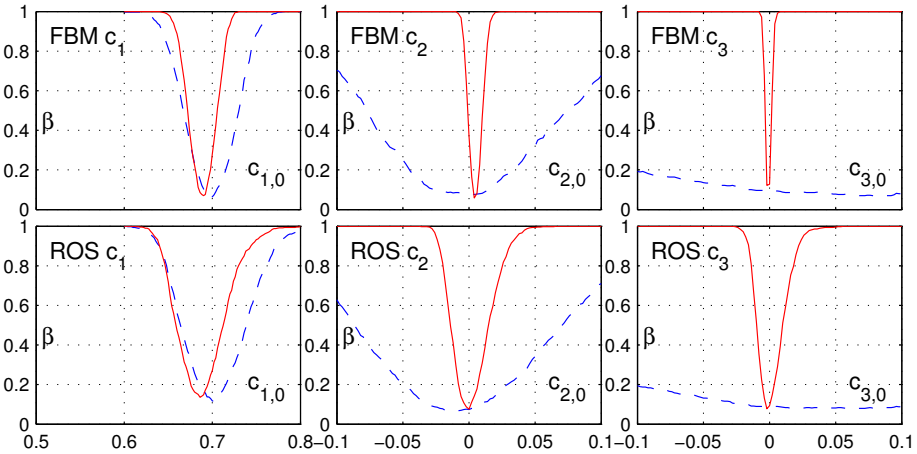


Figure 9.4: **Bootstrap test rejection rates for fBm and ROS.** Rejection rate $\hat{\beta}_{MC}$ of bootstrap tests for fBm (top) and ROS (bottom) for wavelet coefficients (blue dashed) and Leaders (red solid). The true values c_p are $(c_1, c_2, c_3) = (0.7, 0, 0)$, the null values $c_{p,0}$ are given by the abscissa. The target significance of the tests is set to $\alpha = 5\%$. It is significantly more difficult to reject a multifractal hypothesis when data are non Gaussian ROS than when they are Gaussian fBm.

9.4 Conclusions and Perspectives

We have constructed a practical procedure that enables to test a given a priori chosen multifractal property: $H_0 : c_p = c_{p,0}$. Obviously, the choice $H_0 : c_2 = c_{2,0} \equiv 0$ can be seen as a test of mono- versus multi-fractal (indeed, it is conjectured, that $c_2 = 0 \Rightarrow \forall p \geq 3, c_p = 0$). Also, the choice $H_0 : c_3 = c_{3,0} \equiv 0$ can be seen as a practical test of a simple (log-Normal) multifractal model against a more elaborate one. We showed from numerical simulations on synthetic multifractal processes that such tests possess satisfactory statistical performance.

The results discussed above show that the proposed nonparametric bootstrap procedures for testing $H_0 : c_p = c_{p,0}$ present satisfactory performance in reproducing the targeted significances and p-values, equivalently for wavelet coefficients and wavelet Leaders.

Wavelet Leader based tests have consistently by far larger powers than the wavelet coefficient based ones and are thus clearly preferable. A wavelet coefficient based analysis would have poor performance in detecting that the analyzed data do depart from a (monofractal or simple multifractal) model. Conversely, a wavelet Leader based test rejects H_0 with high probability, even in situations where the alternative is close to the null value.

The choice of the type of acceptance region has little impact on the actual significances $\hat{\alpha}_{MC}$ and on the empirical distributions of the p-values, with a slight preference however for the percentile and adjusted percentile methods. These methods obtain as well the largest powers, hence, overall, they are to be preferred. Furthermore, the adjusted method requires the calculation of double bootstrap resamples, increasing the computational cost for the bootstrap by a factor S, without bringing significant improvements. For instance, on a standard PC, the bootstrap estimation and test procedures for c_1, c_2 for

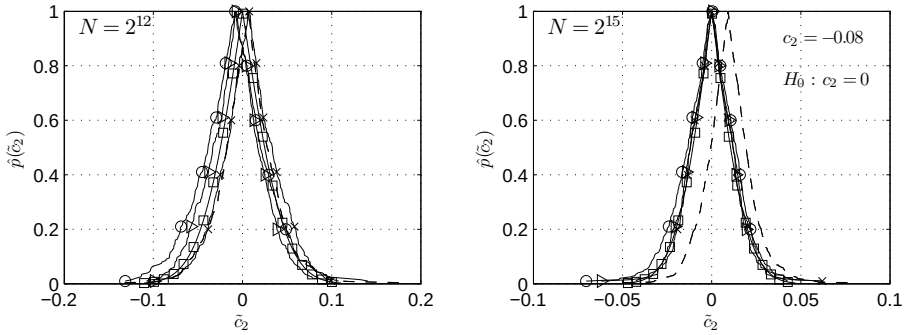


Figure 9.5: **Bootstrap p-value function estimation.** P-value function estimate from Monte Carlo simulation (dashed) and bootstrap estimates from single realization (solid) vs. a potentially observable value $\tilde{c}_2$, obtained for mrw ($c_2 = -0.08$, $c_{2,0} = 0$) using wavelet Leaders. The symbols (□ , ▷ , × , ○) stand for (nor, bas, per, stu) acceptance regions, respectively.

a single observation of sample size $N = 2^{12}\,(2^{15})$ requires approximately $1.6\,(2.7)$ seconds for simple bootstrap methods and $21.3\,(48.4)$ seconds for double bootstrap methods ($R = 399, S = 25$). It is possible that the results obtained with double bootstrap methods could be slightly improved by using a larger number S of double bootstrap resamples, however at the cost of further increasing computational load considerably.
The use of TS-B bootstrap brings (slightly) better actual significances, mainly for wavelet Leader based tests, at the price of increased complexity with respect to T-B bootstrap.
We conclude that tests for $H_0 : c_p = c_{p,0}$ should be based on wavelet Leaders and percentile acceptance regions and TS-B bootstrap.

Practical test procedure. A MATLAB© procedure, designed by the author, implements the proposed multifractality test procedures. To the best of our knowledge, this is the first and only practical multifractal model test that can actually be applied to a single observation of data with finite length. We see this result as an important contribution to empirical multifractal analysis.

Further developments. In addition to obtaining $\hat{d}_\alpha$ and $\hat{p}$, our practical test procedure also outputs, from a single realization, an approximate p-value as a function of a potentially observable value $\tilde{c}_p$. This is done by (numerically) inverting the estimated null distribution:

$$\hat{p}(\tilde{c}_p) = 2 \cdot \min\left(\hat{P}_0^T(\gamma); 1 - \hat{P}_0^T(\gamma)\right) \tag{9.22}$$

with $\gamma = \tilde{c}_p - c_{p,0}$ for test statistic $\hat{t}_B$, and $\gamma = \frac{\tilde{c}_p - c_{p,0}}{\hat{\sigma}^*}$ for $\hat{t}_S$. Examples of such p-value functions are depicted in Fig. 9.5, together with an estimate from numerical simulation. For all acceptance regions, we observe that the functions $\hat{p}(\tilde{c}_p)$ match satisfactorily well the one obtained from Monte Carlo simulations. Such p-value functions can be seen as a useful help for the practitioner. Indeed, the narrower the functions, the more powerful the tests.

Along another line, alternative test statistics such as $T = |c_p - c_{p,0}|$ and one-sided tests have also been studied. Results yield similar conclusions as for the test statistics considered above and are therefore not reported here.

Perspectives and open issues. When analyzing real data, power functions such as those proposed in Fig. 9.2 can usefully complete the test procedure. They can be estimated by numerical simulations on synthetic multifractal processes whose parameters and size fit those of the data under analysis. Methods for estimating the power of a test against specific alternatives from a single realization can further improve the practical test procedures [34].

Finally, the results on bootstrap test performance for non Gaussian H-sssi processes demand for closer investigation of the statistical properties of wavelet coefficients and Leaders for such processes. Results on their dependence structure (cf. Chapter 11) represent a first step in this direction.

Chapter 10

Bootstrap Test for the Time Constancy of Multifractal Attributes

Contents

In the practical multifractal analysis of empirical data, there has been, and there still is, an important controversy: Do scaling actually exist in data, or are they rather the consequence of non stationarities that conspire to mimic scaling behavior? To contribute to answering this question, let us first clarify the issue. There exist two major classes of stochastic processes used to model scale invariance: Self-similar and multifractal multiplicative cascade processes. Both classes consist of non stationary processes, and there is hence no contradiction between scale invariance and non stationarity in that respect. The controversy between scale invariance and non stationarity can in fact be cast in the following three categories: First, scaling actually exist but a smooth trend (in the mean or variance, for example), hence a non stationarity, is superimposed to the data and is likely to impair the analysis; Second, scaling exist in data but their parameters exhibit some form of variability with respect to time, for instance due to a change in experimental conditions; Third, scaling are not present in data but a strong non stationary variability is confused with a scaling property.
The first category has been addressed in a number of research papers (cf. [169] and the references therein) and will not be further considered here. The second and third categories are much more involved as *a non stationary variability* can correspond to many different realities. Nevertheless, their detection is of crucial practical importance, since the blind analysis of such time series is likely to produce misleading interpretations of scaling.
The discrimination of true scaling against various forms of non stationary variability can be addressed with the following heuristic: When data possess true scaling properties, scaling exponents estimated over the entire time series or over non overlapping adjacent windowed time series are statistically consistent. Conversely, when scaling exponents obtained over non overlapping adjacent subsets of the data are not statistically consistent, this can only be the signature of some form of non stationarity, whatever its precise and a priori unknown nature. This heuristic is illustrated in Fig. 10.1. Therefore, the issue of testing scale invariance against non stationarity can be meaningfully recast into a test of time constancy of scaling exponents estimated over adjacent non overlapping subsets of the analyzed time series. This is precisely the intuition developed in [170], where a time constancy test is proposed for the (wavelet coefficient based) second order estimation of the Hurst parameter of Gaussian H-sssi processes. The test is briefly described in Section 10.1.1.
Yet, the use of this test formulation is exclusively limited to the singular case of Gaussian H-sssi processes, and no equivalent procedure is available for validating true scaling for the more valuable class of multifractal processes. Therefore, the aim of this chapter is to extend such an approach to testing for the time constancy of attributes of *multifractal* process. This implies the following major changes in test goals and methodology: First, the second order ($q = 2$) scaling exponent – sufficient for the Gaussian self-similar case considered in [170] – has to be replaced with a whole collection of multifractal attributes, related to both positive and negative statistical orders q. Second, estimations are no longer based on wavelet coefficients but on wavelet Leaders. Third, multifractal processes make up for a large class of stochastic processes with a plurality of poorly studied statistical properties. Consequently, the design of statistical tests is significantly more complicated for multifractal processes. Moreover, the statistics underlying the test can – as opposed to the Gaussian H-sssi case – no longer be obtained analytically.
To cope with these issues, we propose the use of an original nonparametric block bootstrap test design, based on the bootstrap procedures defined in Section 7.2. The test pro-

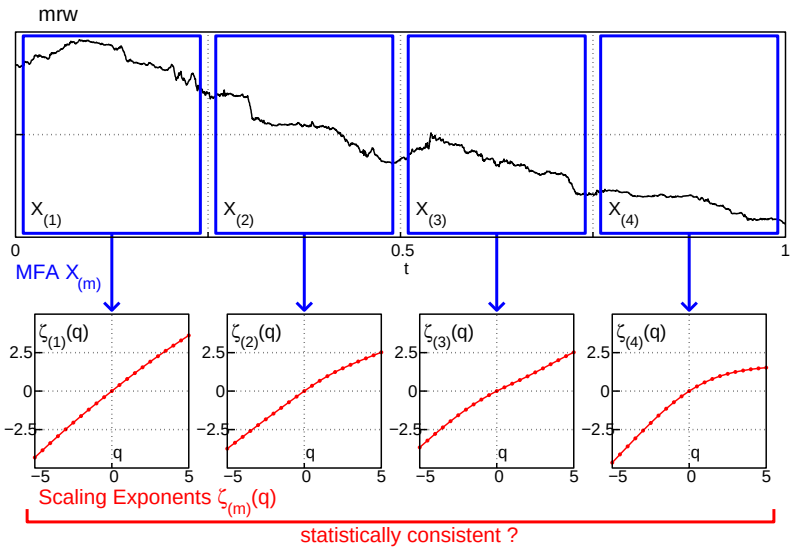


Figure 10.1: **Illustration of heuristic for time constancy of multifractal attributes.** When data possess true scaling properties, multifractal attributes estimated over non overlapping adjacent windowed time series are statistically consistent (shown are here scaling exponents $\zeta_{(m)}(q)$, obtained over $M = 4$ windows.). Conversely, if they are not, this is the signature of some form of non stationarity. The illustration is based on one single realization of mrw ($N = 2^{12}$, $(c_1, c_2) = (0.7, -0.08)$).

cedure employs an original "bootstrap on subsets" and "subsets on bootstrap" approach, which enables the effective estimation of a relevant null distribution for test formulation. We end up with a practical and operational nonparametric test procedure, that exhibits satisfactory statistical performance and that can be applied to a single observation of empirical data to assess the true existence of scaling.
The work presented in this chapter has been reported in [176].

10.1 Bootstrap Time Constancy Test

10.1.1 A test for time constancy of the Hurst parameter

In [170], a uniformly most powerful invariant test for the time constancy of the Hurst parameter H of Gaussian H-sssi processes is devised and analyzed. The test is constructed from wavelet coefficient based estimates $\hat{H}_{(m)}$, obtained from adjacent non overlapping subsets $X_{(m)}$ of X, and relies on Gaussianity, independence and (analytically) known statistics of the estimates. Notably, the variance of the estimates is known a priori and is independent of the true H. The test statistic reads:

$$T_H = \sum_{m=1}^{M} \frac{1}{\sigma^2_{(m)}} \left(\hat{H}_{(m)} - \frac{\sum_{n=1}^{M} \frac{\hat{H}_{(n)}}{\sigma^2_{(n)}}}{\sum_{n=1}^{M} \frac{1}{\sigma^2_{(n)}}} \right)^2 . \tag{10.1}$$

Under the null hypothesis (H constant), its distribution is known exactly, which enables the formulation of the test.

10.1.2 Extension to multifractal processes and bootstrap test principle

To adapt the test to multifractal processes, we have to extend it to any multifractal attributes $\theta \in \{\zeta(q), c_p\}$, whose estimations are based on wavelet Leaders. Let θ denote the multifractal attribute under test. From the time series X to be analyzed, M wavelet Leader based subset estimates $\hat{\theta}_{(m)}$ of θ are obtained from adjacent non overlapping subsets $X_{(m)}$. Assessing the time constancy of θ then amounts to testing the hypothesis that the random variables $\{\hat{\theta}_{(m)}, m = 1, \ldots, M\}$ have identical mean:

$$H_0 : \quad \theta_{(1)} = \theta_{(2)} = \cdots = \theta_{(M)}. \tag{10.2}$$

Yet, this generalization of the test induces two severe difficulties:

1. Variances $\sigma^2_{(m)}$ for the $\hat{\theta}_{(m)}$ are no longer known a priori and are likely to depend on the parameter values (see Section 4.1.7).

2. The null distribution of the test statistics T_θ is no longer known a priori.

To overcome these severe difficulties in the test formulation, we propose to base the test on resamples of wavelet Leaders $\{\mathcal{L}^*_X\}$, obtained from the sample of Leaders $\{\mathcal{L}_X\}$ by a specific resampling approach which is detailed in the following Sections 10.1.3 and 10.1.4.

10.1.3 Bootstrap test statistic

For convenience, let us denote resamples and samples of wavelet Leaders by $\{L_X(j,k)^*\}$ and $\{L_X(j,k)\}$, respectively. The test statistic consists of a modified version of Eq. (10.1). It is based on bootstrap variance estimates for the unknown variances, and on the Graybill Deal estimator instead of the maximum likelihood estimator of the consensus mean:

$$T_\theta = \sum_{m=1}^{M} \frac{1}{\hat{\sigma}^{2*}_{(m)}} \left(\hat{\theta}_{(m)} - \frac{\sum_{n=1}^{M} \frac{\hat{\theta}_{(n)}}{\hat{\sigma}^{2*}_{(n)}}}{\sum_{n=1}^{M} \frac{1}{\hat{\sigma}^{2*}_{(n)}}} \right)^2 . \tag{10.3}$$

The test statistic estimation procedure is sketched in Tab. 10.1 (left):
First, the set of Leaders $\{L_X(j,k)\}$ is cut into M subsets $\{L_{X_{(m)}}(j,k)\}$, corresponding to the subsets $X_{(m)}$. The subset estimates $\hat{\theta}_{(m)}$ are computed by applying Eqs. (2.65) and (2.71) to the $\{L_{X_{(m)}}(j,k)\}$.
Second, the variance estimates $\hat{\sigma}^{2*}_{(m)}$ for each $\hat{\theta}_{(m)}$ are obtained by T-B and TS-B bootstrap resampling from each subset $\{L_{X_{(m)}}(j,k)\}$, yielding R bootstrap samples:

$$\{(L_{X_{(m)}}(j,k))^{*(r)}\},\ r = 1, \cdots, R$$

per subset. Then the estimation procedures defined in Section 2.6 are used on each of these resamples to obtain the bootstrap subset estimates $\hat{\theta}_{(m)}{}^{*(b)}$. Finally, the variance estimates $\hat{\sigma}^{2*}_{(m)}$ for $\hat{\theta}_{(m)}$ are given by the bootstrap sample variances: $\hat{\sigma}^{2*}_{(m)} = \widehat{\mathrm{Var}}^{*}\, \hat{\theta}_{(m)}{}^{*(\cdot)}$.

10.1.4 Bootstrap null distribution estimation

A bootstrap estimate of the distribution of the test statistic T_θ under H_0 is obtained from the empirical distribution of the bootstrap counterpart of Eq. (10.3):

$$T_\theta^* = \sum_{m=1}^{M} \frac{1}{\hat{\sigma}^{2**}_{(m)}} \left(\hat{\theta}^*_{(m)} - \frac{\sum_{n=1}^{M} \frac{\hat{\theta}^*_{(n)}}{\hat{\sigma}^{2**}_{(n)}}}{\sum_{n=1}^{M} \frac{1}{\hat{\sigma}^{2**}_{(n)}}} \right)^2 . \tag{10.4}$$

The null distribution estimation procedure is summarized in Tab. 10.1 (right):
The subset estimates $\hat{\theta}^{*(b)}_{(m)}$ are obtained by, first, T-B or TS-B bootstrap sampling from the *complete* set $\{L_X(j,k)\}$ of Leaders, yielding the R resamples:

$$\{L_X^{*(r)}(j,k)\}, r = 1, \cdots, R,$$

and then, cutting each of these resamples into M subsets:

$$\{(L_X^{*(r)}(j,k))_{(m)}\},$$

from which $\hat{\theta}^{*(r)}_{(m)}$ are obtained by the estimation procedures defined in Section 2.6. Let us emphasize that resampling from the complete set of Leaders, rather than from subsets, is a crucial issue, as it ensures that the $\hat{\theta}^*_{(m)}$ all have the same conditional distributions and thus that T^*_θ reproduces the statistics of T_θ under H_0, shall X satisfy H_0 or H_1 (this is illustrated in Fig. 10.3 and discussed in Section 10.2.4-c) below).
The variance estimates $\hat{\sigma}^{2**(r)}_{(m)}$ of $\hat{\theta}^{*(r)}_{(m)}$ are obtained by first applying T-B or TS-B bootstrap sampling to each $\{L_X^{*(r)}(j,k)\}$, giving the S double bootstrap resamples $\{L_X^{**(r,s)}(j,k)\}$, $s = 1, \cdots, S$, per bootstrap sample $\{L_X^{*(r)}(j,k)\}$. Each of these double bootstrap samples is in turn cut into M subsets $\{(L_X^{**(r,s)}(j,k))_{(m)}\}$, enabling the computation of the double bootstrap subset estimates $\hat{\theta}^{**(r,s)}_{(m)}$. Finally, the double bootstrap sample variance estimates are computed: $\hat{\sigma}^{2**(r)}_{(m)} = \widehat{\mathrm{Var}}^{**} \hat{\theta}^{**(r,\cdot)}_{(m)}$.

10.1.5 Bootstrap test

The test is now readily formulated as:

$$d_\theta = \begin{cases} 1 & \text{if } T_\theta > T^*_{\theta,C} \\ 0 & \text{otherwise,} \end{cases} \tag{10.5}$$

where the test critical value $T^*_{\theta,C}$ is the $(1-\alpha)$ quantile of the empirical distribution of T^*_θ, for a certain preset significance level α. The critical value of α for which the observed test statistic T_θ would be regarded as just decisive against H_0 is called the p-value p_θ of T_θ.

10.2 Performance Assessment and Results

10.2.1 Numerical simulation study

To evaluate the performance of the proposed test procedures, we apply them to a large number N_{MC} of realizations of length N of multifractal random walk (mrw) (see Section

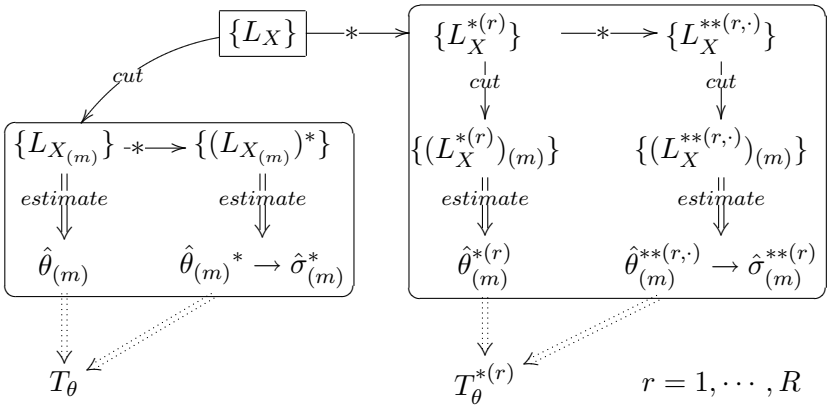


Table 10.1: **Bootstrap resampling scheme for time constancy test.** Procedure for obtaining T_θ (left) and T_θ^* (right) from the wavelet Leaders $\{L_X(j,k)\}$ of X. "cut", "$estimate$" and "$*$" stand for cutting a set into M subsets, computing estimates $\hat{\theta}$ from Eqs. (2.65) and (2.71), and T-B or TS-B bootstrap resampling as in Section 7.2, respectively.

2.7.2-b)), chosen here because it is easy to simulate, and because its multifractal attributes can be conveniently controlled. The simulation parameters are set to $N_{MC} = 1000$, $N = 2^{15}$, $R = S = 99$ and $\alpha = 0.1$. For multifractal attribute estimation, we use Daubechies wavelets with $N_\psi = 3$. Weighted (w^1) regressions are performed over the range of scales (j_1, j_2): $j_1 = 3$ and $j_2(M) = \log_2 N - \log_2 M - (2N_\psi - 1)$ (cf. [179]). The multifractal parameters (specified below) are chosen to correspond to realistic situations observed in actual data (for instance, $c_2 \approx -0.025$ is a commonly accepted value in turbulence, cf. [44, 179]).

10.2.2 Performance assessment

The performance of the test procedures, given that a certain hypothesis $H_{(\cdot)}$ is true, are assessed by their mean rejection rates and p-values,

$$\bar{d}_\theta^{H_{(\cdot)}} = \widehat{\mathbb{E}}_{MC}\{d_\theta | H_{(\cdot)}\} \tag{10.6}$$

$$\bar{p}_\theta^{H_{(\cdot)}} = \widehat{\mathbb{E}}_{MC}\{p_\theta | H_{(\cdot)}\}, \tag{10.7}$$

where $\widehat{\mathbb{E}}_{MC}$ stands for the mean over Monte Carlo simulations. We choose to consider here only results obtained for $\theta \in \{c_1, c_2\}$ and $M \in \{2, 4\}$ with equal subset lengths. Similar results can be obtained for c_p, $p \geq 3$ and $\zeta(q)$, other choices of M, and splitting into subsets of non equal length.

10.2.3 Performance under H_0

The performance under H_0 are studied on processes with constant multifractal attributes (c_1, c_2). Tab. 10.2 summarizes results for three different sets of parameters: $(c_1, c_2) = (0.75, -0.01)$ (left), $(c_1, c_2) = (0.8, -0.02)$ (center) and $(c_1, c_2) = (0.75, -0.08)$ (right) for T-B bootstrap (top) and TS-B bootstrap (bottom). They indicate that the mean rejection rates $\bar{d}_\theta^{H_0}$ are overall close to the preset significance level $\alpha = 0.1$ for both c_1 and c_2, and for all three parameter settings. Furthermore, the mean p-values $\bar{p}_\theta^{H_0}$ are close to 0.5, indicating a satisfactory null distribution estimation. Indeed, under H_0, the p-value

H_0	T-B						
	$(c_1,\, c_2)$	$(0.75, -0.01)$		$(0.8, -0.02)$		$(0.75, -0.08)$	
	θ	c_1	c_2	c_1	c_2	c_1	c_2
$M = 2$	$\bar{d}_\theta^{H_0}$	0.113	0.143	0.075	0.139	0.022	0.052
	$\bar{p}_\theta^{H_0}$	0.478	0.469	0.530	0.485	0.618	0.551
$M = 4$	$\bar{d}_\theta^{H_0}$	0.177	0.224	0.103	0.173	0.014	0.056
	$\bar{p}_\theta^{H_0}$	0.432	0.379	0.512	0.432	0.687	0.560
H_0	TS-B						
	$(c_1,\, c_2)$	$(0.75, -0.01)$		$(0.8, -0.02)$		$(0.75, -0.08)$	
	θ	c_1	c_2	c_1	c_2	c_1	c_2
$M = 2$	$\bar{d}_\theta^{H_0}$	0.125	0.137	0.122	0.143	0.128	0.110
	$\bar{p}_\theta^{H_0}$	0.470	0.469	0.468	0.462	0.463	0.501
$M = 4$	$\bar{d}_\theta^{H_0}$	0.214	0.216	0.193	0.227	0.138	0.116
	$\bar{p}_\theta^{H_0}$	0.426	0.406	0.423	0.400	0.463	0.477

Table 10.2: **Mean rejection rates under null hypothesis.** Mean rejection rates $\bar{d}_\theta^{H_0}$ and p-values $\bar{p}_\theta^{H_0}$ of the tests under H_0 for $(c_1,\, c_2) = (0.75,\, -0.01)$ (left), $(c_1,\, c_2) = (0.8,\, -0.02)$ (center) and $(c_1,\, c_2) = (0.75, -0.08)$ (right) for T-B bootstrap (top) and TS-B bootstrap (bottom). Under H_0, it is expected that $\bar{d}_\theta^{H_0} \approx \alpha = 0.1$, and $\bar{p}_\theta^{H_0} \approx 0.5$.

would be uniformly distributed between 0 and 1 if the test was based on the exact null distribution of the test statistic.
Yet, there are minor discrepancies in the observed test sizes that slightly depend on process parameters: For T-B, actual significances consistently become smaller (and p-values larger) as c_2 becomes more negative (and hence, the process is further from monofractal). For TS-B, observed test sizes are slightly above nominal (and p-values below 0.5), independently of process parameters. Results not reported here show that these slight discrepancies are mainly due to small differences between the variance $\hat{\sigma}^2_{(m)}$ of $\hat{\theta}_{(m)}$ (as measured from numerical simulations) and its bootstrap estimate $\hat{\sigma}^{*2}_{(m)}$: Whereas $\hat{\sigma}^{**2}_{(m)} \approx \hat{\sigma}^{*2}_{(m)}$ for both T-B and TS-B, $\hat{\sigma}^{*2}_{(m)} \geq \hat{\sigma}^2_{(m)}$ for T-B, and $\hat{\sigma}^{*2}_{(m)} \leq \hat{\sigma}^2_{(m)}$ for TS-B. The origins of these differences can be interpreted in the light of result obtained in Chapter 11.
The results lead us to the conclusion that the empirical distribution of T^*_θ is a practically satisfactory approximation of the null distribution of T_θ under H_0, and that it is robust with respect to the precise values of the multifractal parameters.

10.2.4 Performance under H_1

To study the powers of the proposed tests, we need to define an alternative hypothesis. One could imagine many forms of non stationary processes, a number of them being likely to mimic scaling behaviors when analyzed blindly over the entire time series. Here, we study one of the simplest such alternatives: processes possess piece-wise constant multifractal attributes. H_1 is thus analyzed with an alternative consisting of the concatenation of two truly multifractal processes of equal length with different multifractal attributes $(c_1^{(i)}, c_2^{(i)})_{i=1,2}$. Two cases are investigated.

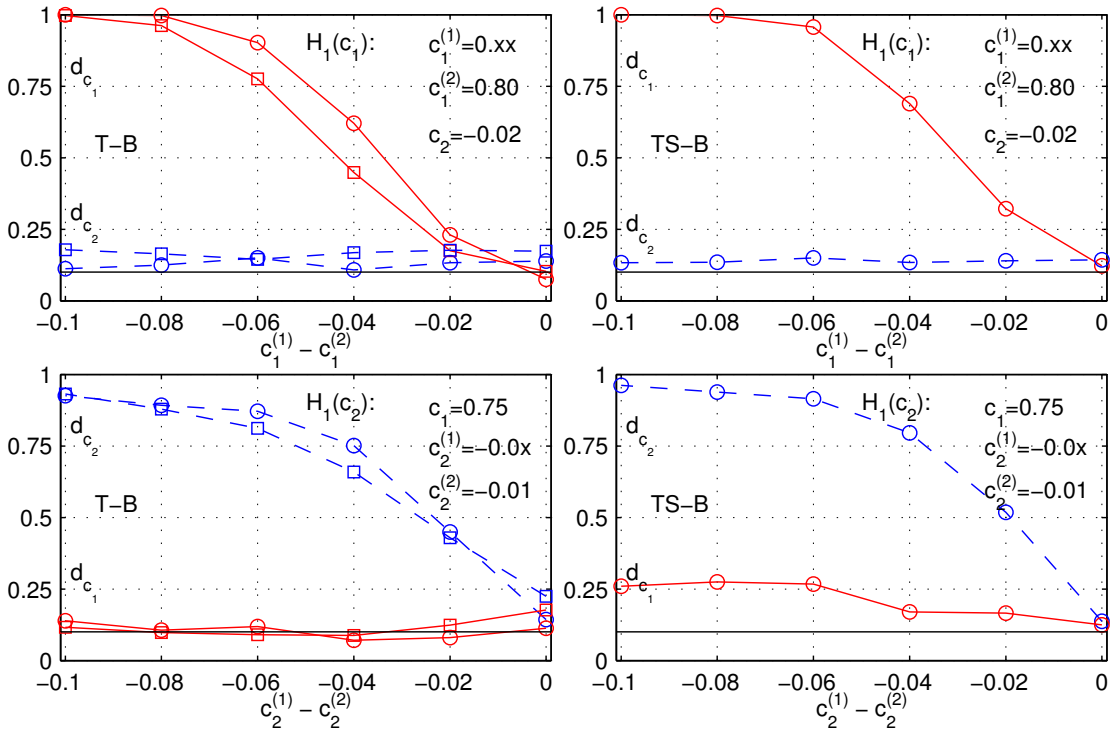


Figure 10.2: **Mean rejection rates under alternative hypothesis: Test power.** Test decisions $\bar{d}_{c_1}$ (solid red lines) and $\bar{d}_{c_2}$ (dashed blue lines) for T-B (left) and TS-B (right) under $H_1(c_1)$ (top) and $H_1(c_2)$ (bottom) for $M = 2$ (circles) and $M = 4$ (squares, T-B only) as a function of $c_p^{(1)} - c_p^{(2)}$. The horizontal solid black line indicates the preset significance level α.

10.2.4-a) Non constant c_1, constant c_2

In the first case, which we denote $H_1(c_1)$, we set $c_1^{(1)} \neq c_1^{(2)}$ and $c_2^{(1)} \equiv c_2^{(2)} = c_2$, i.e., c_2 is constant, while c_1 is not. Thus, $\bar{d}_{c_1}$ assesses the power of the test for time constancy of c_1, and $\bar{d}_{c_2}$ should ideally reproduce the preset significance α. The parameters are set to $c_2 = -0.02$ and $c_1^{(1)} = \{0.70, 0.72, \cdots, 0.80\}$, $c_1^{(2)} = 0.8$. Fig. 10.2 (top) shows test decisions $\bar{d}_{c_1}$ (solid red line) and $\bar{d}_{c_2}$ (dashed blue line) as a function of the step size $c_1^{(1)} - c_1^{(2)}$. The rightmost points $c_1^{(1)} - c_1^{(2)} = 0$ correspond to the mean rejection rates under H_0 of Tab. 10.2 (right). Fig. 10.2 (top) demonstrates that $\bar{d}_{c_1}$ increases fast with $|c_1^{(1)} - c_1^{(2)}|$ and thus that the test is powerful: When $c_1^{(1)} - c_1^{(2)} = -0.04$ ($c_1^{(1)} = 0.76$, $c_1^{(2)} = 0.8$), corresponding to values that are in practice considered as being very close, the test rejects the time constancy hypothesis for c_1 with a probability above 0.6 ($M = 2$) and close to 0.5 ($M = 4$). Conversely, the mean test decisions $\bar{d}_{c_2}$ reproduce closely the preset significance level α and remain constant when $c_1^{(1)} - c_1^{(2)}$ varies, indicating that the time constancy test for c_2 is not subject to cross-influence from changes in c_1. We conclude, first, that the test for time constancy of c_1 is powerful, and second, that the test for constancy of c_2 closely reproduces the level α, independently of $c_1^{(1)} - c_1^{(2)}$.

10.2.4-b) Constant c_1, non constant c_2

In the second case, which we denote $H_1(c_2)$, we set $c_2^{(1)} \neq c_2^{(2)}$ and $c_1^{(1)} \equiv c_1^{(2)} = c_1$, i.e., c_1 is constant, while c_2 is not. Therefore, $\bar{d}_{c_2}$ assesses the power of the test for time constancy of c_2. The parameters are set to $c_1 = 0.75$ and $c_2^{(1)} = \{-0.11, -0.09, \cdots, -0.01\}$, $c_2^{(2)} = -0.01$. Fig. 10.2 (bottom) shows test decisions $\bar{d}_{c_1}$ (solid red line) and $\bar{d}_{c_2}$ (dashed blue line) as a function of the step size $c_2^{(1)} - c_2^{(2)}$. Exchanging the roles of $\bar{d}_{c_1}$ and $\bar{d}_{c_2}$, conclusions are similar to those obtained under $H_1(c_1)$: Satisfactory power of the test for time constancy of c_2, and insensitivity of the test on c_1 with respect to level change $c_2^{(1)} - c_2^{(2)}$ for T-B bootstrap. For TS-B bootstrap, $\bar{d}_{c_1}$ slightly increases with increasing $|c_2|$.

10.2.4-c) Null distribution estimation under H_1

Fig. 10.3 shows bootstrap test critical values (as defined in Eq. (10.5)) $T^*_{c_1,C}$ under $H_1(c_1)$ (left) and $T^*_{c_2,C}$ under $H_1(c_2)$ (right). The circles and the bars correspond, respectively, to $\widehat{\mathbb{E}}_{MC} T^*_{\theta,C}$ and to $1.64 \cdot \widehat{\mathrm{Std}}_{MC} T^*_{\theta,C}$. We observe that the $T^*_{c_p,C}$ do not depend on the step size $c_p^{(1)} - c_p^{(2)}$ and thus on the precise hypothesis H_1. Moreover, the $T^*_{c_p,C}$ equal the critical values under H_0 (given by the rightmost points). This demonstrates that the empirical distribution of T^*_θ (Eq. (10.3)) under H_1 provides us with a robust and accurate null distribution estimation, as test design demands. Also, this illustrates that the specific bootstrap "resample and cut" procedure is efficient for approximate null distribution estimation, and that discrepancies in test size under $H_0(c_1)$ and $H_0(c_2)$ are due to small differences between the variance of $\hat{\theta}_{(m)}$ and bootstrap variance estimations $\hat{\sigma}^{*2}_{(m)}$, as discussed above.

10.3 Conclusions and Perspectives

We have devised a practical procedure for discriminating the existence of true scaling properties against non stationarities. It consists of a bootstrap based test for the constancy along time of wavelet Leader based multifractal parameter estimates. We have shown, by means of numerical simulations, that this bootstrap based test procedure is reliable and powerful. Notably, the empirical distribution of T^*_θ under H_1 yields a robust estimation of the null distribution, a central feature for relevant test design. Our procedure successfully addresses this nontrivial issue by combining a "split then bootstrap" for T_θ and a "bootstrap then split" for T^*_θ. Notably, it possesses significant power in detecting the occurrence of a change in a given multifractal parameter. The procedure can be used to test for time constancy of any multifractal attribute and can be applied to a single observation of real data with unknown statistical characteristics. It has heavy computational cost (due to double bootstrap) but remains, to the best of our knowledge, the only procedure practically available. The impact of choosing M remains to be discussed in terms of trade-off for type-II errors: A test with too small M may miss non stationarities, choosing M too large results in a lack of power due to poor estimations, hence the existence of an optimal M for a given but unknown alternative hypothesis. The procedure can be further extended to testing the constancy along time of the whole structure functions ($S(2^j, q)$, $\hat{C}(2^j, p)$) or to testing jointly the constancy of a vector of multifractal attributes.

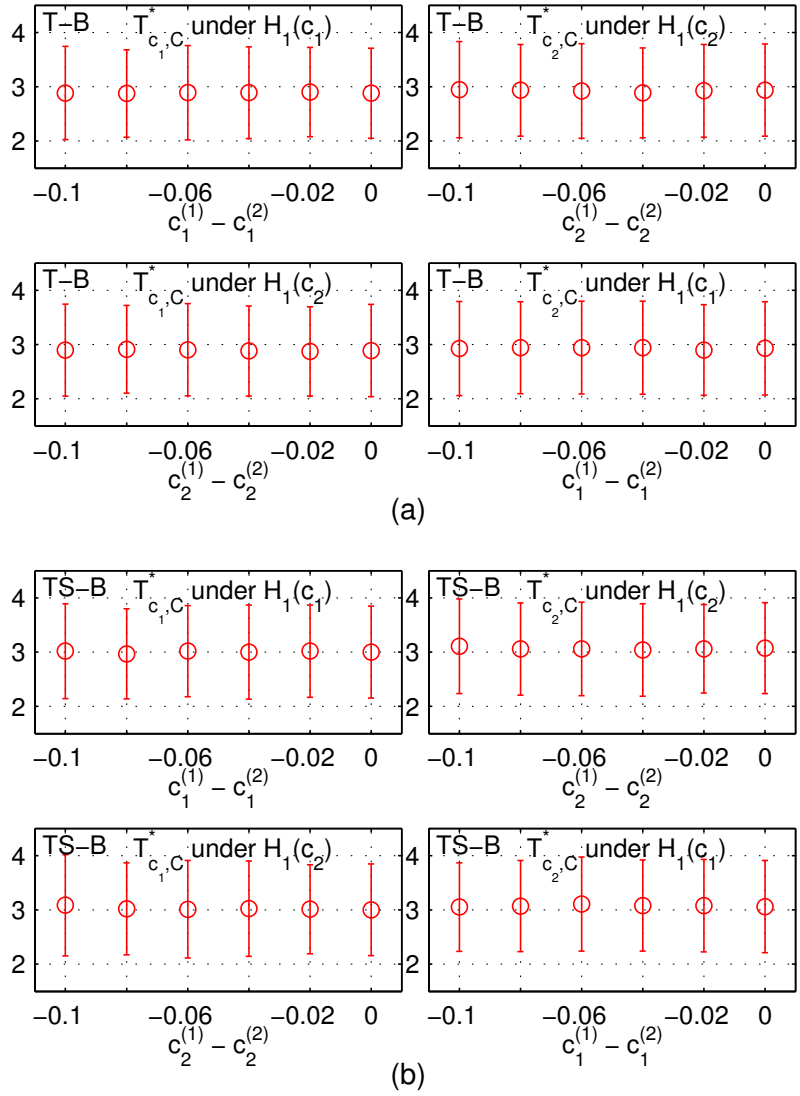


Figure 10.3: **Critical values of time constancy tests.** Bootstrap test critical values $T^*_{\theta,C}$ under H_1 (mean value '○', $1.64\sigma_{T^*_{\theta,C}}$ bars, obtained through Monte Carlo simulations): $T^*_{c_1,C}$ under $H_1(c_1)$ (top left) and $H_1(c_2)$ (bottom left), $T^*_{c_2,C}$ under $H_1(c_2)$ (top right) and $H_1(c_1)$ (bottom right) for T-B (a) and TS-B (b) ($M = 2$).

Chapter 11

Dependence Structure of Multiresolution Quantities

Contents

An important issue in practical applications concerns the dependence structure of multiresolution quantities, such as increments, wavelet coefficients or wavelet Leaders, for scale invariant (H-sssi and MMC) processes. Despite their being widely used in practical scaling and multifractal analysis, little is known on their dependence, and results mostly concentrate on the covariance or correlation of wavelet coefficients. In his seminal work, Flandrin [70] showed that wavelet coefficients of fBm are short range dependent (SRD) — i.e. fast decay of correlation, in contrast to long range dependent (LRD), cf. Def. 2.1 — on condition that the analyzing wavelet has enough vanishing moments N_ψ:

$$N_\psi > H + \frac{1}{2}, \tag{11.1}$$

be the process LRD or not. This is a direct consequence of the correlation structure of finite variance (fv) H-sssi processes (Eq. (2.16)) and hence also valid for non Gaussian fv H-sssi processes (cf. Section 2.3.4, **P3**). Also, increments are a special case of wavelets (referred to as "poor man wavelets", cf. Section 2.5.5), and their order P acts as their number of vanishing moments N_ψ. Therefore, higher order increments of order P fulfill Eq. (11.1) with $N_\psi = P$ [117]. For Gaussian fBm, Condition (11.1) also implies weak dependence, i.e. weak correlation only of higher orders of coefficients. Processes other than fBm received only marginal attention (see, a contrario, [77] for the covariance of wavelet coefficients for a specific multifractal processes class, and [15] for the correlation of the logarithm of wavelet coefficients for a specific multifractal cascade process), and the results in [70] have in practice commonly been heuristically associated with other processes than fBm. Also, to the knowledge of the author, no dependence results beyond (first order) correlation/covariance are available. Therefore, the following key practical questions remain unanswered:
Are results similar to those obtained in [70] for fBm valid for the correlation of multiresolution quantities of (finite variance) non Gaussian H-sssi processes? What if the process is a multifractal multiplicative cascade? Most importantly, do such results extend to the *dependence* structure of multiresolution quantities of such processes?
Answering such questions is interesting in itself and deserves per se being studied. Since estimators are based on sample moments such as Eqs. (2.7) and (2.58), this has major importance in applications because empirical moments have poor performance for LRD data ([33, 153], see Section 2.1.1-a)). Besides other practical implications, the answers may give new insights in the bootstrap performance reported in the previous three chapters. Also, they may provide a starting point for considering the theoretical validity and performance of the bootstrap procedures defined in Chapter 7.

Therefore, the first goal of the present chapter is to analyze the role N_ψ takes in controlling the *correlation* of wavelet coefficients of non Gaussian H-sssi processes and of multifractal multiplicative cascade processes, and to compare it to the key role it plays for Gaussian H-sssi processes by Eq. (11.1). The second goal is to investigate whether this result extends to the correlation of the absolute value of wavelet coefficients, $|d_X|$, and to correlation of higher orders q of wavelet coefficients, $|d_X|^q$, thus to their *dependence structure*. In the same spirit, we investigate the dependence, hence correlations and correlations of higher orders q, of wavelet Leaders L_X and L_X^q. These are the topics of Sections 11.1 to 11.3. We obtain both analytic and numerical results for the dependence of wavelet coefficients and wavelet Leaders, for both H-sssi and MMC processes. They indicate that, under conditions such as Eq. (11.1), wavelet coefficients are indeed only *weakly correlated*, but they are *in general strongly dependent* – and so are wavelet

Leaders – in the sense that their absolute values and q-th orders have LRD-like power law type correlations.

The second goal of this chapter is the study of the implications of such results on wavelet domain block bootstrap for multifractal analysis. This is discussed in Section 11.4.

The publication of the results presented in this chapter is in preparation [139, 168]. Some of the results in Sections 11.1 to 11.3 have been obtained in collaboration with Béatrice Vedel (LAMA, Université de Bretagne Sud, Vannes). This is stated explicitly in the text.

11.1 Finite Variance H-sssi Processes

The correlation function of wavelet coefficients of fv H-sssi processes is given by [70]:

$$\mathbb{E}d_X(j,k)d_X(j',k') \sim |2^j k - 2^{j'}k'|^{2H-2N_\psi}, \quad |2^j k - 2^{j'}k'| \to \infty, \tag{11.2}$$

and hence decays fast under the condition given by Eq. (11.1). Let us examine the generality of this property of the dyadic wavelet transform – dramatic decrease of correlation for fv H-sssi processes – for the absolute value of wavelet coefficients $|d_X|$, and for Leaders L_X.

11.1.1 Fractional Brownian motion. Gaussian finite variance H-sssi process

Let us first consider the Gaussian case and investigate correlations of the absolute value of wavelet coefficients $|d_X|$ and Leaders L_X of fractional Brownian motion.

11.1.1-a) Covariance of the absolute value of wavelet coefficients

For ease of notation, let us write $X = d_X(j,k)$, $Y = d_Y(j',k')$, $\sigma_x^2 = \mathrm{Var} d_X(j,k)$ and $\sigma_y^2 = \mathrm{Var} d_Y(j',k')$. Since the wavelet coefficients of fBm are Gaussian ([70], see Section 2.3.4, property **P4**), we can write their bivariate distribution:

$$f_{X,Y}(x,y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} \exp\left(-\frac{z}{2(1-\rho^2)}\right), \tag{11.3}$$

where $\rho = \dfrac{\mathsf{Cov}(X,Y)}{\sigma_x\sigma_y}$ and $z = \dfrac{x^2}{\sigma_x^2} - \dfrac{2\rho xy}{\sigma_x\sigma_y} + \dfrac{y^2}{\sigma_y^2}$. Then, the correlation of the absolute value of two wavelet coefficients is calculated as:

$$\begin{aligned}
\mathbb{E}|X||Y| &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |x||y| f_{X,Y}(x,y)dxdy = \\
&= \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |x||y| \exp\left(-\frac{\frac{x^2}{\sigma_x^2} - 2\rho\frac{xy}{\sigma_x\sigma_y} + \frac{y^2}{\sigma_y^2}}{2\sqrt{1-\rho^2}}\right) dxdy = \\
&= \frac{\sigma_x}{2\pi^{3/2}\sigma_y^2\sqrt{1-\rho^2}} \int_{-\infty}^{\infty} 2\sigma_y\sqrt{\pi}(1-\rho^2)|y| \exp\left(-\frac{y^2}{2\sigma_y^2(1-\rho^2)}\right) + \cdots \\
&\quad \cdots + \sqrt{2}\pi\rho\sqrt{1-\rho^2}y|y| \exp\left(-\frac{y^2}{2\sigma_y^2}\right) erf\left(\frac{\rho y}{\sqrt{2}\sigma_y\sqrt{1-\rho^2}}\right) dy = \\
&= \frac{\sigma_x\sqrt{1-\rho^2}}{\pi\sigma_y} \int_{-\infty}^{\infty} |y| \exp\left(-\frac{y^2}{2\sigma_y^2(1-\rho^2)}\right) dy + \cdots \\
&\quad \cdots + \frac{2\sigma_x\rho}{\sqrt{2\pi}\sigma_y^2} \int_0^{\infty} y^2 \exp\left(-\frac{y^2}{2\sigma_y^2}\right) erf\left(\frac{\rho y}{\sqrt{2}\sigma_y\sqrt{1-\rho^2}}\right) dy = \\
&= \frac{2\sigma_X\sigma_y(1-\rho^2)^{3/2}}{\pi} + \frac{\sigma_X\sigma_y\rho}{\pi\sqrt{1-\rho^2}} \left(2\sqrt{1-\rho^2}\arctan\left(\frac{\rho}{\sqrt{1-\rho^2}}\right) + 2\rho(1-\rho^2)\right) \\
&= \frac{2\sigma_X\sigma_y}{\pi}\left(\sqrt{1-\rho^2} + \rho\arctan\left(\frac{\rho}{\sqrt{1-\rho^2}}\right)\right).
\end{aligned}$$

The covariance of the absolute value of two wavelet coefficients is therefore:

$$\mathsf{Cov}(|X|,|Y|) = \mathbb{E}|X||Y| - \mathbb{E}|X|\mathbb{E}|Y| = \frac{2\sigma_X\sigma_y}{\pi}\left(\sqrt{1-\rho^2} - 1 + \rho\arctan\left(\frac{\rho}{\sqrt{1-\rho^2}}\right)\right), \tag{11.4}$$

and the correlation coefficient is given by:

$$\tilde{\rho} = \frac{\mathsf{Cov}(|X|,|Y|)}{\mathsf{Std}|X|\mathsf{Std}|Y|} = \frac{2}{\pi-2}\left(\sqrt{1-\rho^2} - 1 + \rho\arctan\left(\frac{\rho}{\sqrt{1-\rho^2}}\right)\right). \tag{11.5}$$

Therefore, as illustrated in Fig. 11.1, the absolute values of wavelet coefficients $|d_X|$ always have a correlation coefficient smaller than that of the wavelet coefficients d_X themselves:

$$\tilde{\rho} = \frac{\mathsf{Cov}(|X|,|Y|)}{\mathsf{Std}|X|\mathsf{Std}|Y|} \leq \rho = \frac{\mathsf{Cov}(X,Y)}{\sigma_X\sigma_Y}.$$

Let us finally study the behavior of the correlation for large time lags at fixed scale j. From Eq. (11.2), it is clear that (with $\tau = |2^j k - 2^{j'}k'|$):

$$\rho \sim \tau^{-\alpha}, \quad \tau \to \infty, \quad \alpha > 0$$

with $\alpha = 2N_\psi - 2H$, and that the d_X are hence LRD if $0 < \alpha < 1$. The development of Eq. (11.5) for $\tau \to \infty$ ($\rho \to 0$) is:

$$\begin{aligned}
\tilde{\rho} &= \frac{1}{\pi-2}\left(\rho^2 + \frac{1}{12}\rho^4\right) + \mathcal{O}(\rho^6), \quad \rho \to 0 &(11.6)\\
&\sim \tau^{-2\alpha} + \frac{1}{12}\tau^{-4\alpha} + \mathcal{O}(\tau^{-6\alpha}), \quad \tau \to \infty. &(11.7)
\end{aligned}$$

Therefore, we can distinguish three cases for the correlation of the absolute value of wavelet coefficients of fractional Brownian motion:

1. If the coefficients d_X are short range dependent ($1 < \alpha$ or $H < 0.5$), so are their absolute values $|d_X|$.
2. If the coefficients d_X are LRD with $0.5 < \alpha < 1$ (i.e. $0.5 < H < 0.75$), their absolute values $|d_X|$ are *short range dependent*: $\tilde{\rho} \sim \tau^{-\alpha'}$, $1 < \alpha' = 2\alpha$.
3. If the coefficients d_X are LRD with $0 < \alpha < 0.5$ (i.e. $0.75 < H < 1$), their absolute values $|d_X|$ are LRD with $0 < \alpha' < 1$.

We note that a similar calculation as described here for the covariance of $|d_X|$ for fBm has been obtained in [132].

11.1.1-b) Numerical study of large time lag correlation: Scaling analysis

Analytic calculations of the correlation of wavelet Leaders L_X seem to be difficult to obtain (see e.g. Section 3.2), even in the textbook case of Gaussian fBm. Also, analytic results for dependence of wavelet coefficients are difficult to obtain for the non Gaussian H-sssi Rosenblatt process and for the MMC processes considered in later sections of this chapter. Hence, for their exploration, we need to resort to numerical results.

Eq. (11.2) shows that for fv H-sssi processes, wavelet coefficients at a given scale are correlated with (fast) power law decay for large time lags, controlled by H and N_ψ. Also, numerical evidence in Fig. 3.5 suggests that (absolute values of) wavelet coefficients and wavelet Leaders at a given scale have power law like decay of correlation, hence indicating scale invariance (and potentially LRD-like correlations) for these time series. The state of the art tool for the estimation of the parameter of the power law like correlation of potentially LRD time series is discrete wavelet coefficient based estimation of $\zeta(2)$ as in Eq. (2.65) [4]. It relates to γ (Def. 2.1) and α (Eq. (2.3), and as above) as:

$$\alpha = 2 - \zeta(2) \quad \text{and} \quad \gamma = \zeta(2) - 1.$$

We apply this estimator, with $N'_\psi = 4$ vanishing moments, to the time series $\{d_X(J, \cdots)\}$, $\{|d_X(J, \cdots)|^q\}$ or $\{L_X(J, \cdots)^q\}$ of wavelet coefficients or Leaders at scale J, obtained with N_ψ vanishing moments from the process $X(t)$ under analysis. Rather than $\hat{\alpha}$ or $\hat{\gamma}$ – and with reminiscence to fv H-sssi processes (and in particular Section 2.2.2-b)) – we use the equivalent estimate[1]:

$$\hat{H} = \hat{\zeta}(2)/2. \tag{11.8}$$

The actual existence of scale invariance in these time series is validated for each case we consider by inspection of the corresponding structure functions. The final estimates are obtained as:

$$\tilde{H} = \hat{\mathbb{E}}_{N_{MC}} \hat{H}, \tag{11.9}$$

i.e., means over estimations for a large number N_{MC} of series of wavelet coefficients (Leaders) at scale J, calculated from N_{MC} different realizations of the process X. To distinguish $\tilde{H}$ for the different time series, we use the superscripts ^{d}, $^{|d|^q}$ and $^{L^q}$ for $\{d_X(J, \cdots)\}$, $\{|d_X(J, \cdots)|^q\}$ and $\{L_X(J, \cdots)^q\}$, respectively, and omit the double superscript $^{()^q}$ when $q = 1$. Similarly, we use a subscript $_{(\cdot)}$ for indicating the number of vanishing moments N_ψ of the wavelet used for analyzing the process $X(t)$.

[1] Hence, we speak of LRD in the time series of wavelet coefficients or Leaders if: $0.5 < \hat{H} < 1$, and say that they are weakly correlated or SRD otherwise.

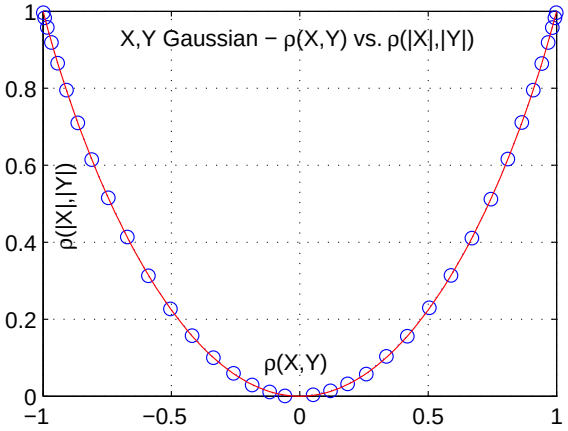


Figure 11.1: **Correlation coefficient of magnitude of Gaussian random variables.** Illustration of Eq. (11.5): Correlation coefficient $\rho(|X|,|Y|)$ of magnitude of correlated Gaussian r.v.s versus correlation coefficient $\rho(X,Y)$ of correlated Gaussian r.v.s (red solid line). Mean of estimates of $\rho(|X|,|Y|)$ for 500 realizations of Gaussian samples of size $N = 1000$ with correlation coefficient $\rho(X,Y)$ ('$\circ$').

11.1.1-c) Correlation of multiresolution quantities: Numerical results

Fig. 11.2 (top row) presents results obtained by this numerical simulation study for d_X ($\triangle$), $|d_X|$ ($\circ$) and L_X ($\times$) at scale $J = 3$ for fBm with $H_{fBm} = H = 0.8$. Equivalent results, not reported here, can be obtained at different scales J. It summarizes structure functions $S_J^d(j,2)$, $S_J^{|d|}(j,2)$ and $S_J^L(j,2)$ for $N_\psi = 1$ and $N_\psi = 2$ (left and center plots, respectively), and mean estimations of H Eq. (11.9) as a function of the number of vanishing moments $N_\psi \in \{1,2,3\}$ (plot on the right).

Fig. 11.2 (top left and center) indicates that the structure functions $S_J^d(j,2)$, $S_J^{|d|}(j,2)$ and $S_J^L(j,2)$ behave linearly in log-log coordinates for coarse scales, hence indicating scale invariance for the series of wavelet coefficients and Leaders and validating the estimation of $\tilde{H}$. First, we observe that for $N_\psi = 1$, the d_X are LRD with $\tilde{H}_1^d \approx H_{fBm} = 0.8$, as predicted by theory: Since a wavelet with $N_\psi = 1$ is identical to the first order increment, the wavelet coefficients d_X ressemble the increment process of fBm, hence fractional Gaussian noise (fGn) with corresponding $H = H_{fBm}$. The absolute values $|d_X|$ are less correlated with $\tilde{H}_1^{|d|} \approx 0.6 < \tilde{H}_1^d$, which is exactly conform to Eqs. (11.5) and (11.7) since $2\tilde{H}_1^d - 1 \approx 0.6$. Nonetheless, the $|d_X|$ are LRD for $N_\psi = 1$. Second, for $N_\psi \geq 2$, both the time series d_X and $|d_X|$ are only weakly correlated, with $\tilde{H}_{\geq 2}^d \approx \tilde{H}_{\geq 2}^{|d|} \approx 0.5$, in agreement with [70] and Eq. (11.5). Finally, the time series of wavelet Leaders L_X display similar correlation as $|d_X|$, $\tilde{H}_{(\cdot)}^L \approx \tilde{H}_{(\cdot)}^{|d|}$: Hence, wavelet Leaders of fBm form LRD sequences for $N_\psi = 1$, and are only weakly correlated (and not LRD) when $N_\psi \geq 2$.

We conclude that for fBm, the results of [70] on the role of N_ψ for weak correlation of wavelet coefficients translates to their absolute value and to their wavelet Leaders.

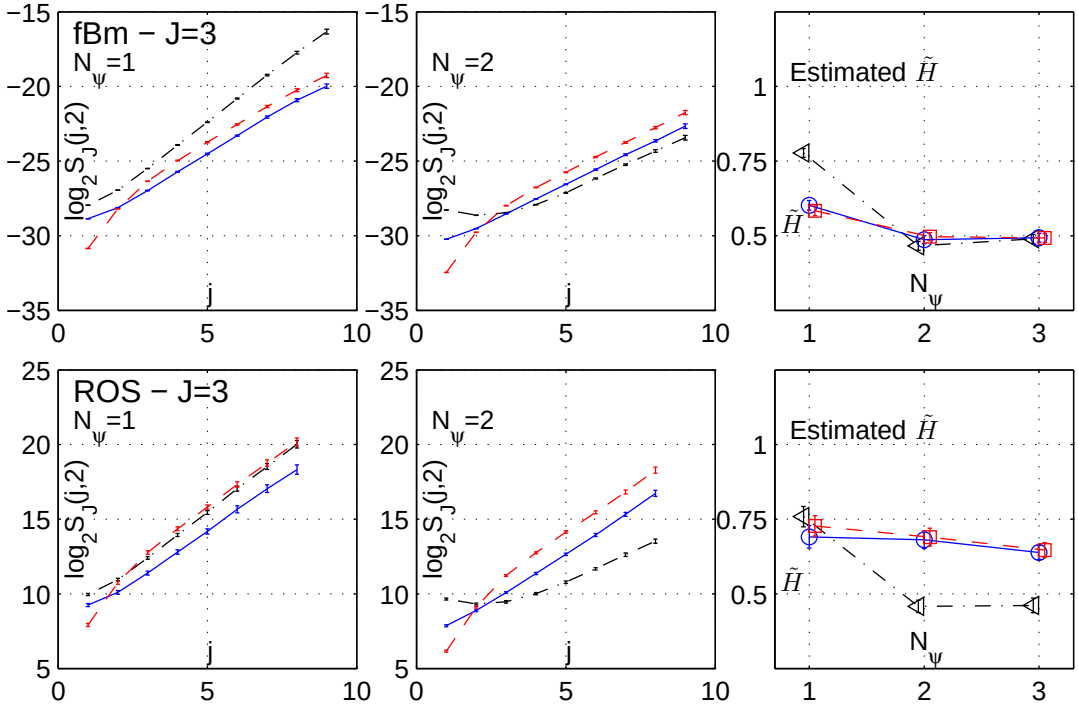


Figure 11.2: **Correlation of coefficients and Leaders of H-sssi processes.** Structure functions $\log_2 S(j, q = 2)$ for second order scaling analysis of wavelet coefficients $d_X(J = 3, \cdot)$ (black dashed-dotted lines and '◁'), $|d_X(J = 3, \cdot)|$ (blue solid lines and 'o'), and Leaders $L_X(J = 3, \cdot)$ (red dashed lines and '□') for fBm (top) and ROS (bottom) with $H = 0.8$ and vanishing moments $N_\psi = 1$ (left column) and $N_\psi = 2$ (center column) for obtaining d_X from the process $X(t)$. The column on the right summarize the estimates of the parameter $\tilde{H}$ (Eq. (11.9)) as a function of N_ψ. Results are obtained as means over $N_{MC} = 100$ realization of fBm and ROS ($H = 0.8$) of sample size $N = 2^{16}$, with regressions performed at coarse scales.

11.1.2 Rosenblatt process. Non Gaussian finite variance H-sssi processes

Fig. 11.2 (bottom row) presents results obtained by numerical simulations as conducted for fBm for the non Gaussian H-sssi Rosenblatt process with $H_{ROS} = H = 0.8$. Shown are structure functions $S_J^d(j, 2)$, $S_J^{|d|}(j, 2)$ and $S_J^L(j, 2)$ for $N_\psi = 1$ and $N_\psi = 2$ (first and second plot, respectively), and mean estimations of $\tilde{H}$ as a function of the number of vanishing moments $N_\psi \in \{1, 2, 3\}$ (plot on the right) for d_X (△), $|d_X|$ (o) and L_X (×) at scale $J = 3$.

The structure functions display linear behavior in log-log coordinates for coarse scales for d_X, $|d_X|$ and L_X and therefore indicate scale invariance for these time series, hence relevance of estimation of $\tilde{H}$ as proposed in Section 11.1.1-b).

For $N_\psi = 1$, the d_X are LRD with $\tilde{H}_1^d \approx H_{ROS} = 0.8$, interpretable in a similar way as for fBm: wavelet coefficients for $N_\psi = 1$ mimic the increment process of ROS, which is LRD with $H = H_{ROS}$. For $N_\psi \geq 2$, the time series d_X are – as for fBm – only weakly correlated, with $\tilde{H}_{\geq 2}^d \approx 0.5$. These findings confirm the extensions of the results in [70] to fv non Gaussian H-sssi processes – weak correlation of wavelet coefficients of fv H-sssi

processes under condition Eq. (11.1) – and are consistent with theory, since the covariance functions of any fv H-sssi process are identical. However, we observe that for ROS, the time series $|d_X|$ remain LRD when $N_\psi \geq 2$, with $\tilde{H}^{|d|}$ only slightly smaller than H_{ROS}. Also, $\tilde{H}^{|d|}$ does not significantly decrease with increasing N_ψ. Hence, the results for $|d_X|$ for fBm do not in general extend to non Gaussian H-sssi processes: While Eq. (11.1) ensures weak correlation only of wavelet coefficients for any H-sssi process, this comes together with weak dependence exclusively for the Gaussian fBm. Finally, the time series of wavelet Leaders L_X display similar correlation as those of $|d_X|$: $\tilde{H}^L \approx \tilde{H}^{|d|}$. Hence, wavelet Leaders of ROS form LRD sequences, regardless of how many vanishing moments N_ψ the wavelet used in the wavelet transform possesses.

Theoretical results. A formal proof of these numerical indications, based on increments, is under current investigation, in collaboration with B. Vedel. Its final version is not ready at the time this text is being written and is hence omitted. This is a joint work with B. Vedel who, after numerical simulations by the author, obtained preliminary analytical results for the covariance of increments of ROS. The results suggest that the covariance of the first and second order increments of ROS $X(t)$, $X_\tau^{(1)}(t) = X(t+\tau) - X(t)$ and $X_\tau^{(2)}(t) = X_\tau^{(1)}(t+\tau) - X_\tau^{(1)}(t)$, to the power $q = 2$ decays, in the limit of $\tau \to 0$ (or also, equivalently, $|t-s| \to \infty$), as:

$$\lim_{\tau \to 0} \frac{\mathbb{E}X_\tau^{(P)}(t)^2 X_\tau^{(P)}(s)^2 - \mathbb{E}X_\tau^{(P)}(t)^2 \mathbb{E}X_\tau^{(P)}(s)^2}{\tau^{4-\alpha_{ROS}}} = C|t-s|^{-\alpha_{ROS}}, \tag{11.10}$$

where $\alpha_{ROS} = 2 - 2H_{ROS}$. Hence, their covariance has algebraic LRD-like power law decay with $\tilde{H} = H_{ROS}$. The publication of these results is in preparation [139].

11.1.3 Conclusions

The analytic and numerical results presented in this section lead to the following conclusions. For Gaussian fBm, wavelet coefficients d_X are SRD if $N_\psi \geq H + \frac{1}{2}$. This result immediately translates to any finite variance H-sssi processes, since it is based only on the covariance structure Eq. (2.16) of such processes [70]. For Gaussian fBm, weak correlation also implies weak dependence. However, this is not necessarily similar for non Gaussian H-sssi processes, and numerical simulations indicate indeed that this is not the case: Higher orders of wavelet coefficients $|d_X|$ are in general long range dependent, regardless of the precise choice of N_ψ.
Finally, the numerical results suggest that the time series formed by the wavelet Leaders L_X have similar (large time lag) covariance behavior as $|d_X|$. Therefore, it is as well in general long range dependent, with the exception of fBm when $N_\psi \geq H + \frac{1}{2}$.
Hence, fBm represents a specific case because it is a Gaussian process, and weak correlation of wavelet coefficients for fv H-sssi processes should not in general be interpreted as weak dependence, since this is only valid for fBm.

11.2 Multifractal Multiplicative Cascade Processes

To our knowledge, the only results on dependence of multiresolution quantities for MMC processes (cf. Section 2.7) available in the literature are the work of Arneodo et al. [15] and of Gonçalves and Riedi [77]: In [15], the authors study the time correlation of the

logarithm of wavelet coefficients, $\log|d_X(j,\cdot)|$, for dyadic Random Wavelet Cascades (cf. Section 11.2.1), in a perspective of estimation of c_2 in cascade processes. They show that correlations between the logarithm of wavelet coefficients at distance τ decay as $\sim c_2 \log_2 \tau$. In [77], results for the large time lag correlation decay of wavelet coefficients $d_X(j,\cdot)$ are obtained for compound multifractal motion processes, such as CPM-MF-fBm. Note, however, that no such process was actually defined at the time of the publication of this paper. They obtain results on the influence of N_ψ on the correlation decay of wavelet coefficients d_X, similar to those obtained in [70] for fBm. For CPM-MF-fBm, their result suggests fast decay of correlation under condition that $N_\psi \geq \frac{\zeta(2)}{2}$, where $\zeta(q) = qH + \varphi(qH)$ are the scaling exponents of the CPM-MF-fBm (cf. Section 2.7.2-a)). Therefore, the goal of this section is to characterize the role of N_ψ and the correlation of (the q-th powers of) the absolute value of coefficients $|d_X|^q$ and Leaders L_X^q for multifractal multiplicative cascade processes. In Section 11.2.1, we calculate analytically the covariance of $|d_X(j,\cdot)|^q$ for (dyadic) Random Wavelet Cascades (RWC), in the same spirit as the results in [15] for $\log|d_X(j,\cdot)|$. This original result gives indications for the role of the function $\varphi(q)$ for the decay of (higher order) correlation for multiresolution quantities of MMC processes. In Section 11.2.2, we derive results for the decay of correlation of increments of first and second order, and of their absolute value taken to the q-th power, for compound Poisson motion (CPM). Together with our results for RWC in Section 11.2.1, this enables us to postulate a conjecture for the large time lag decay behavior of the correlation of (q-th powers of the absolute value of) wavelet coefficients and Leaders for MMC processes in general. Section 11.2.3 validates these analytic results by numerical results (cf. Section 11.1.1-b)) for d_X, $|d_X|$, $|d_X|^q$ and L_X, L_X^q of MMC processes.

11.2.1 Random wavelet cascades: Dependence structure of wavelet coefficients

11.2.1-a) Random wavelet cascades

The particular construction of RWC [15] (see Section 2.7.2-c)) implies that, on the one hand side, the marginal distributions of the wavelet coefficients can be written immediately (for laws stable w.r.t. multiplication, cf. Eq. (2.126)), making explicit analytic calculations feasible. On the other hand side, the time series of wavelet coefficients (or increments) at each scale j are non stationary, hence certain theoretical and practical difficulties for the investigation of dependence. Despite these particularities, analytic results obtained on RWC can give important indications for properties of other multifractal multiplicative cascade type processes. For instance, [15] calculates (some time average of) the correlation of $\log|d_X(j,k)|$ and $\log|d_X(j,k+2^j\tau)|$ and shows that it behaves as $\sim c_2 \log_2(\tau)$ for $2^{-j} \ll \tau < 1$, where c_2 is the second log-cumulant of the cascade and depends on the specific choice of the multipliers W. Exactly the same logarithmic decay behavior of correlations of the logarithm of wavelet coefficients has been observed on empirical time series such as instant velocity in hydrodynamic turbulence experiments.

11.2.1-b) Covariance for statistical orders q

In this subsection, we calculate analytically the covariance of two wavelet coefficients $|d_{j,k}|^q = |d_X(j,k)|^q$ and $|d_{j,k'}|^q = |d_X(j,k')|^q$ to the q-th power in RWC[2]. For convenience, we use the notations for geometry introduced in [15]:

$$\begin{aligned} \Delta x &= 2^{-j}\Delta k \\ \Delta k &= 2^p. \end{aligned}$$

Also, let us write $\gamma(q) = \mathbb{E}W^q$, hence: $\varphi(q) = -\log_2 \mathbb{E}W^q = -\log_2 \gamma(q)$ (cf. Eq. (2.127)). Let d denote the scale at which two coefficients $|d_{j,k}|^q$ and $|d_{j,k'}|^q$ have their first multiplier in common. Since the expectation of the product of independent r.v.s equals the product of their expectations, the correlation of $|d_{j,k}|^q$ and $|d_{j,k'}|^q$ is given by:

$$\mathbb{E}|d_{j,k}|^q|d_{j,k'}|^q = \prod_{\#\text{ common } W} \mathbb{E}W_{j'}^{2q} \prod_{\#\text{ not common } W} \mathbb{E}W_{j'}^{q} = \left(\mathbb{E}W^{2q}\right)^d \cdot \left(\mathbb{E}W^q\right)^{2(j-d)}, \quad (11.11)$$

and, since:

$$\mathbb{E}|d_{j,k}|^q\mathbb{E}|d_{j,k'}|^q = \prod_{j'=0}^{j-1} \mathbb{E}W_{j'}^{q} \prod_{j'=0}^{j-1} \mathbb{E}W_{j'}^{q} = \left(\mathbb{E}W^q\right)^{2j}, \quad (11.12)$$

their covariance is given by:

$$\text{Cov}\left(|d_{j,k}|^q, |d_{j,k'}|^q\right) = \left(\mathbb{E}W^q\right)^{2j}\left[\left(\frac{\mathbb{E}W^{2q}}{\left(\mathbb{E}W^q\right)^2}\right)^d - 1\right] = \gamma(q)^{2j}\left[\left(\frac{\gamma(2q)}{\gamma(q)}\right)^d\right], \quad (11.13)$$

and hence, the process $|d(j,\cdot)|$ is non stationary since d depends on k and k'. In [15], it is proposed to form time averages of the covariance of $\log|d_{j,k}|$, $\log|d_{j,k'}|$ over all positions k, k' with dyadic difference 2^p. The approach translates to $|d_{j,k}|^q$ as:

$$C_q(j,2^p) = 2^{-j}\sum_{k=1}^{2^j-2^p} \text{Cov}\left(|d_{j,k}|^q, |d_{j,k+2^p}|^q\right). \quad (11.14)$$

Denote by $N_{j,2^p}(d)$ the number of pairs of coefficients $d_{j,k}$ and $d_{j,k+2^p}$ that have a common parent at the same scale d and hence the same contribution to $C_q(j,2^p)$. Then, $C_q(j,2^p)$ simplifies to:

$$C_q(j,2^p) = 2^{-j}\sum_{d=0}^{j-1} N_{j,2^p}(d)\gamma(q)^{2j}\left[\left(\frac{\gamma(2q)}{\gamma(q)}\right)^d - 1\right]. \quad (11.15)$$

It can be shown that $N_{j,2^p}(d)$ has the following properties [15]:

$$\begin{aligned} N_{j,2^p}(d) &= 0 \text{ for } d \geq j-p & (11.16)\\ N_{j,2^p}(d) &= N_{j-p,1}(d) & (11.17)\\ N_{j,1}(d) &= 2^d, & (11.18) \end{aligned}$$

[2]As stated before in Section 2.8, we only consider covariance of coefficients within one integral scale. Beyond integral scale, correlations are trivially zero, since multiplicative cascades/trees are not connected.

and, by using these properties one after the other:

$$C_q(j,2^p) = 2^{-j}\sum_{d=0}^{j-p-1} N_{j,2^p}(d)\gamma(q)^{2j}\left[\left(\frac{\gamma(2q)}{\gamma(q)}\right)^d - 1\right] \quad (11.19)$$

$$= 2^{-j}\sum_{d=0}^{j-p-1} 2^p N_{j-p,1}(d)\gamma(q)^{2j}\left[\left(\frac{\gamma(2q)}{\gamma(q)}\right)^d - 1\right] \quad (11.20)$$

$$= 2^{p-j}\sum_{d=0}^{j-p-1} 2^d\gamma(q)^{2j}\left[\left(\frac{\gamma(2q)}{\gamma(q)}\right)^d - 1\right] \quad (11.21)$$

$$= 2^{p-j}\gamma(q)^{2j}\sum_{d=0}^{j-p-1} 2^d\left[\left(\frac{\gamma(2q)}{\gamma(q)}\right)^d - 1\right]. \quad (11.22)$$

Now, using $\sum_{n=0}^{N} a^n = \frac{a^{N+1}-1}{a-1}$ and grouping power law terms in $(\cdot)^p$, we have:

$$C_q(j,2^p) = 2^{p-j}\gamma(q)^{2j}\left[\frac{2^{j-p}\left(\frac{\gamma(2q)}{\gamma(q)^2}\right)^{j-p}-1}{2\frac{\gamma(2q)}{\gamma(q)^2}-1} - (2^{j-p}-1)\right] = \quad (11.23)$$

$$= \frac{\gamma(q)^2\gamma(q)^{2j}}{2\gamma(2q)-\gamma(q)^2}\left[\left(\frac{\gamma(2q)}{\gamma(q)^2}\right)^j\left(\frac{\gamma(q)^2}{\gamma(2q)}\right)^p + 2\cdot 2^{-j}\left(\frac{\gamma(2q)}{\gamma(q)^2}-\frac{1}{2}\right)2^p - 2\frac{\gamma(2q)}{\gamma(q)^2}+1\right] \quad (11.24)$$

$$= -\gamma(q)^{2j} + \frac{\gamma(q)^2\gamma(2q)^j}{2\gamma(2q)-\gamma(q)^2}\cdot\left(\frac{\gamma(q)^2}{\gamma(2q)}\right)^p + 2^{-j}\gamma(q)^{2j}\cdot 2^p. \quad (11.25)$$

Finally, substituting $\Delta x = 2^{-j}2^p$ and $\varphi(q) = -\log_2\gamma(q)$:

$$C_q(j,\Delta x) = -\gamma(q)^{2j} + 2^j\frac{\gamma(q)^2\gamma(2q)^j}{2\gamma(2q)-\gamma(q)^2}\Delta x^{2\log_2\gamma(q)-\log_2\gamma(2q)} + \gamma(q)^{2j}\Delta x \quad (11.26)$$

$$= A + B\cdot\Delta x^{\varphi(2q)-2\varphi(q)} + C\cdot\Delta x, \quad (11.27)$$

with constants:

$$A = -2^{-2j\varphi(q)} \quad (11.28)$$

$$B = 2^j\frac{2^{-2\varphi(q)}2^{-j\varphi(2q)}}{2\cdot 2^{-\varphi(2q)}-2^{-2\varphi(q)}} \quad (11.29)$$

$$C = 2^{-2j\varphi(q)}. \quad (11.30)$$

Since $\varphi(0) = \varphi(1) = 0$ and $\varphi(q)$ is a concave function, $\varphi(2q) - 2\varphi(q) < 0$ (at least) for $q \geq 1$. Hence, when Δx is small, $2^{-j} < \Delta x \ll 1$ and for $q \geq 1$:

$$C_q(j,\Delta x) \sim \Delta x^{\varphi(2q)-2\varphi(q)}. \quad (11.31)$$

If in addition $j \gg 1$, Eq. (11.31) is also valid for larger Δx, $2^{-j} < \Delta x < 1$, since:

$$\frac{C}{B} = \frac{2^{1-\varphi(2q)+2\varphi(q)}-1}{2^{j(1-\varphi(2q)+2\varphi(q))}} \approx 0, \quad (11.32)$$

and $1 - \varphi(2q) + 2\varphi(q) > 1$.
Therefore, the (time average Eq. (11.13) of the) covariance function of $|d_X(j,\cdot)|^q$ and $|d_X(j,k')|^q$ has *power law* like decay:

$$\mathrm{Cov}|d_{j,k}|^q|d_{j,k'}|^q \approx c|k-k'|^{\varphi(2q)-2\varphi(q)}. \quad (11.33)$$

Conclusions. The covariance of the absolute value of wavelet coefficients of RWC does not decay exponentially, but it has *algebraic* decay. Therefore, the wavelet coefficients $|d_X(j,\cdot)|^q$ have in general strongly *LRD-like power law type* covariance, since $\varphi(2q)-2\varphi(q) \in [-1,0]$ for some range of q (cf. Section 11.2.3, Eq. (11.60)). The precise decay rate is entirely controlled by the behavior of the process-dependent function $\varphi(q)$.

11.2.2 Compound Poisson motion: Dependence structure of increments

The specific construction of RWC makes explicit calculations of the (higher order) covariance of wavelet coefficients possible. For other MMC processes, however, such analytic calculations do not seem to be feasible for wavelet coefficients. Therefore, in this section, we derive analytical results for the correlation of the increments of compound Poisson motion $A(t) = \lim_{r\to 0}\int_0^t Q_r(u)du$ [22, 23] (Section 2.7.1-b), Eq. (2.106)). We suggest that results translate to wavelet coefficients through the role of the number of vanishing moments, since the P-th order increment has $N_\psi = P$ vanishing moments. This heuristic is validated numerically in Section 11.2.3.

11.2.2-a) Correlation of increments

Increments of order 1**.** Let us denote $\lambda = \varphi(2)$ (cf. Eq. (2.102)). For convenience, let us assume that $t > s+\tau$ and that $r < t-s-\tau$, such that we can drop the limit and the explicit reference to r in what follows. The first order increment of $A(t)$ is defined as:

$$A_\tau^{(1)}(t) = A(t+\tau) - A(t) = \int_t^{t+\tau} Q_r(u)du. \tag{11.34}$$

Since $\mathbb{E}Q_r(t)Q_r(s) = \exp\left(-\varphi(2)m(\mathcal{C}_r(t)\cap\mathcal{C}_r(s))\right) = |t-s|^{\varphi(2)}$ ([46, 47], cf. Eq. (2.105)), the correlation of $A_\tau^{(1)}(t)$ and $A_\tau^{(1)}(s)$ is:

$$\begin{aligned} C_{A_\tau^{(1)}}(t,s) &= \mathbb{E}A_\tau^{(1)}(t)A_\tau^{(1)}(s) = \mathbb{E}\int_t^{t+\tau}\int_s^{s+\tau} Q_r(u)Q_r(v)dudv = & (11.35)\\ &= \int_t^{t+\tau}\int_s^{s+\tau} \mathbb{E}Q_r(u)Q_r(v)dudv = \int_t^{t+\tau}\int_s^{s+\tau}(u-v)^\lambda dudv. & (11.36)\end{aligned}$$

Changing variables to $\alpha = u-v$ and $\beta = \frac{u+v}{2}$, the integral can be solved as:

$$\begin{aligned} C_{A_\tau^{(1)}}(t,s) &= \int_{t-s-\tau}^{t-s}\alpha^\lambda(-t+s+\tau+\alpha)d\alpha + \int_{t-s}^{t-s+\tau}\alpha^\lambda(t-s+\tau-\alpha)d\alpha = & (11.37)\\ &= \int_{t-s-\tau}^{t-s}\alpha^{\lambda+1} - (t-s-\tau)\alpha^\lambda d\alpha - \int_{t-s}^{t-s+\tau}\alpha^{\lambda+1} - (t-s+\tau)\alpha^\lambda d\alpha & (11.38)\\ &= \frac{1}{(\lambda+2)(\lambda+1)}\left[(t-s-\tau)^{\lambda+2} + (t-s+\tau)^{\lambda+2} - 2(t-s)^{\lambda+2}\right], & (11.39)\end{aligned}$$

which, for $t-s \gg \tau$ behaves as:

$$\begin{aligned} C_{A_\tau^{(1)}}(t-s) &\approx c(t-s)^\lambda\tau^2 + \mathcal{O}(\tau^4) & (11.40)\\ &\approx c(t-s)^{\varphi(2)} \quad (\tau \text{ fix}). & (11.41)\end{aligned}$$

Therefore, the correlation of increments of order 1 of CPM has algebraic (power law) decay and is – since mostly $\varphi(2) \in [-1, 0]$ – in general LRD like.

Increments of order 2. The second order increment of $A(t)$ is given by:

$$A_\tau^{(2)}(t) = A_\tau^{(1)}(t+\tau) - A_\tau^{(1)}(t) = \int_{t+\tau}^{t+2\tau} Q_r(u)du - \int_t^{t+\tau} Q_r(u)du. \tag{11.42}$$

In contrast to the first order increment $A_\tau^{(1)}(t)$, $A_\tau^{(2)}(t)$ can hence be negative. Let us suppose that $t - s > 2\tau$. The correlation of $A_\tau^{(2)}(t)$ and $A_\tau^{(2)}(s)$ is:

$$\begin{aligned} C_{A_\tau^{(2)}}(t,s) &= \mathbb{E}A_\tau^{(2)}(t)A_\tau^{(2)}(s) = & (11.43)\\ &= 2\mathbb{E}A_\tau^{(1)}(t)A_\tau^{(1)}(s) - \mathbb{E}A_\tau^{(1)}(t)A_\tau^{(1)}(s+\tau) - \mathbb{E}A_\tau^{(1)}(t+\tau)A_\tau^{(1)}(s) & (11.44)\end{aligned}$$

and can hence be expressed in terms of $C_{A_\tau^{(1)}}(\cdot,\cdot)$ in Eq. (11.39). Similarly, successive higher order increments can be calculated (see below). Substituting from Eq. (11.39) and grouping terms gives:

$$\begin{aligned} C_{A_\tau^{(2)}}(t-s) = \frac{1}{(\lambda+2)(\lambda+1)} \Big[&-(t-s-2\tau)^{\lambda+2} + 4(t-s-\tau)^{\lambda+2} - & (11.45)\\ &-6(t-s)^{\lambda+2} + 4(t-s+\tau)^{\lambda+2} - (t-s+2\tau)^{\lambda+2} \Big] & (11.46)\end{aligned}$$

which, in the limit $t - s \gg \tau$, behaves as:

$$\begin{aligned} C_{A_\tau^{(2)}}(t-s) &\approx c(t-s)^{\lambda-2}\lambda(\lambda-1)\tau^4 + \mathcal{O}(\tau^6) & (11.47)\\ &\approx c(t-s)^{\varphi(2)-2} \quad (\tau \text{ fix}). & (11.48)\end{aligned}$$

Therefore, since $\varphi(2) < 0$, the correlation of increments of order 2 of CPM has algebraic decay, but is not LRD like.

Increments of order P. The correlations $C_{A_\tau^{(P+1)}}(t-s)$ and $C_{A_\tau^{(P)}}(t-s)$ of the $(P+1)$st and Pth order increments, respectively, are related by the recursion:

$$C_{A_\tau^{(P+1)}}(t-s) = -C_{A_\tau^{(P)}}(t-s-\tau) + 2C_{A_\tau^{(P)}}(t-s) - C_{A_\tau^{(P)}}(t-s+\tau), \tag{11.49}$$

as can easily be seen from $A_\tau^{(P+1)}(t) = A_\tau^{(P)}(t+\tau) - A_\tau^{(P)}(t)$. It can hence be shown that:

$$C_{A_\tau^{(P)}}(t-s) = \frac{1}{(\lambda+2)(\lambda+1)} \sum_{n=0}^{2P} (-1)^n \binom{n}{P} [t-s+(n-P)\tau]^{\lambda+2}, \quad |t-s| > P\tau, \tag{11.50}$$

which, in the limit $t - s >> \tau$, behaves as:

$$C_{A_\tau^{(P)}}(t-s) \approx c(t-s)^{\lambda+2-2P}. \tag{11.51}$$

Therefore, the increment processes have algebraic power law type decay and are SRD on condition that:

$$N_\psi = P > \frac{\varphi(2)+2}{2} + 1/2 = \frac{\zeta(2)}{2} + 1/2. \tag{11.52}$$

Since $\varphi(2) < 0$, this practically means $N_\psi \geq 2$.

11.2.2-b) Correlation of wavelet coefficients

In the previous paragraph, we showed that it is possible to calculate exactly the correlation function for increments of compound Poisson motion. These can then be used for understanding the large time lag correlation behavior and deriving results such as Eq. (11.52). For other MMC processes or for wavelet coefficients, such an explicit calculation could not be obtained. However, it is possible to show that the covariance function of finite variance, stationary increment, and scale invariant processes (the MMC processes considered in this manuscript fall into this category, with the exception of CMC, for which increments are not stationary) necessarily is of the form:

$$f(t) + f(s) - f(t-s) \text{ with } f(t) = |t|^\gamma,\ 0 < \gamma < 2.$$

By a Fourier domain argument, similar to that used in [70], it can then be argued (cf. [77] for a particular example) that for such processes, wavelet coefficients d_X are SRD on condition that:

$$N_\psi > \frac{\zeta(2)}{2} + 1/2.$$

11.2.2-c) Higher order correlations of increments

Let us now investigate (higher order) correlations of the absolute value of the increments.

Increments of order 1. Since by construction $Q_r(t) > 0$, the first order increments are always positive and hence, the correlation of $|A_\tau^{(1)}(t)|$ and $|A_\tau^{(1)}(s)|$ equals the correlation of $A_\tau^{(1)}(t)$ and $A_\tau^{(1)}(s)$:

$$C_{|A_\tau^{(1)}|}(t-s) = C_{A_\tau^{(1)}}(t-s), \tag{11.53}$$

and hence, for $1 > t - s >> \tau$ (cf. Eqs. (11.40) and (11.41)):

$$C_{|A_\tau^{(1)}|}(t-s) \approx c(t-s)^{\varphi(2)}\tau^2 + \mathcal{O}(\tau^4). \tag{11.54}$$

Also, it is possible to characterize the correlation of $A_\tau^{(1)}(t)^q$ and $A_\tau^{(1)}(s)^q$:

Theorem 11.1 (Increments of order 1, $q \geq 1$ an integer) *Let $1 \leq q < q_c^+/2$ be an integer, $t > 0$, $s > 0$ and $\tau > 0$ be such that $s > t + \tau$ and $|t - s - \tau| < 1$. One has:*

$$|t-s+\tau|^{\varphi(2q)-2\varphi(q)} \leq \frac{\mathbb{E}A_\tau^{(1)}(t)^q A_\tau^{(1)}(s)^q}{C(q)\tau^{2(q+\varphi(q))}} \leq |t-s-\tau|^{\varphi(2q)-2\varphi(q)}.$$

In particular, there exists $C(q) > 0$ such that, for $0 < t - s < 1$, one has:

$$\lim_{\tau\to 0} \frac{\mathbb{E}A_\tau^{(1)}(t)^q A_\tau^{(1)}(s)^q}{|\tau|^{2(q+\varphi(q))}} = C(q)|t-s|^{\varphi(2q)-2\varphi(q)}. \tag{11.55}$$

Note that since $\varphi(1) = 0$, this corresponds exactly to what is obtained in Eq. (11.54) for $q = 1$. Moreover, the constant $C(q)$ can be calculated from $C(q)|\tau|^{2(q+\varphi(q))} = \left(\mathbb{E}A_\tau^{(1)}(1)^q\right)^2$.

For $q \notin \mathbb{N}$, an exact result for scaling has not been obtained. Yet, the the following inequalities can be obtained, showing that the exact power law behaviors for integer q extend to real q, at least in the limit $|t - s| \to 0$:

Theorem 11.2 (Increments of order 1, $q \geq 1$ real) *Let $1 \leq q < q_c^+/2$ be a real number. There exist $0 < C_1(q), C_2(q) < \infty$, depending only on q, such that, for $s > t > 0$ and $|t-s| < 1$, one has:*

$$C_1(q)|t-s|^{c(\varphi(2q)-2\varphi(q))} \leq \lim_{\tau \to 0} \frac{\mathbb{E} A_\tau^{(1)}(t)^q A_\tau^{(1)}(s)^q}{|\tau|^{2(q+\varphi(q))}} \leq C_2(q)|t-s|^{\varphi(2q)-2\varphi(q)}. \quad (11.56)$$

Eq. (11.56) can be rewritten as[3]:

$$\lim_{|t-s|\to 0} \lim_{\tau \to 0} \frac{\mathbb{E} A_\tau^{(1)}(t)^q A_\tau^{(1)}(s)^q}{|\tau|^{2(q+\varphi(q))}} \approx \mathcal{O}|t-s|^{\varphi(2q)-2\varphi(q)}. \quad (11.57)$$

It is conjectured that the double-inequality in Eq. (11.56) turns to an equality:

Conjecture 11.1 (Increments of order 1, $q \geq 1$ real) *Let $1 \leq q_c^+/2 \leq q$ be a real number. There exists $0 < C(q) < \infty$, depending only on q, such that, for $s > t > 0$ and $|t-s| < 1$, one has:*

$$\lim_{\tau \to 0} \frac{\mathbb{E} A_\tau^{(1)}(t)^q A_\tau^{(1)}(s)^q}{|\tau|^{2(q+\varphi(q))}} \leq C(q)|t-s|^{\varphi(2q)-2\varphi(q)}. \quad (11.58)$$

The detailed proof of Theorem 11.1 has been obtained by the author in collaboration with B. Vedel, that of Theorem 11.2 is due to B. Vedel. They are sketched in Appendices C.1 and C.2, respectively.

Increments of order 2. The same kind of geometric arguments used in the proof of Theorem 11.1 can be used for proofing the following proposition:

Proposition 11.1 (Increments of order 2, $q = 2$) *Let $q = 2$ and $|t-s+2\tau| > 0$. There exists a constant $C > 0$ such that:*

$$\lim_{\tau \to 0} \frac{\mathbb{E} A_\tau^{(2)}(t)^2 A_\tau^{(2)}(s)^2}{\mathbb{E} A_\tau^{(2)}(t)^2 \mathbb{E} A_\tau^{(2)}(s)^2} = \lim_{\tau \to 0} \frac{\mathbb{E} A_\tau^{(2)}(t)^2 A_\tau^{(2)}(s)^2}{C|\tau|^{2(2+\varphi(2))}} = |t-s|^{(\varphi(4)-2\varphi(2))}. \quad (11.59)$$

The detailed mathematical proof of Proposition 11.1 is due to B. Vedel. Its final version is not ready at this time and is hence not reported in this manuscript.

The publication of the results on the dependence structure of increments of CPM obtained in this section is in preparation [168].

11.2.3 Dependence structure for multifractal multiplicative cascade processes: Conjecture and numerical results

The results in Sections 11.2.1-b) and 11.2.2, derived for wavelet coefficients of RWC and for increments of CPM, suggest the following conjecture:

[3] The notation in Eq. (11.58) stands for: The exponent $\beta = (\varphi(2q) - 2\varphi(q))$ is the only β such that: $0 < \lim_{|t-s|\to 0} \lim_{\tau \to 0} \left(\frac{\mathbb{E} A_\tau^{(1)}(t)^q A_\tau^{(1)}(s)^q}{|\tau|^{2(q+\varphi(q))}} \right) / \left(|t-s|^\beta \right) < \infty$

Conjecture 11.2 *Let* $1 \leq q < q_c^+/2$. *There exists* $C(q) > 0$ *such that, for* $0 < |t-s| < 1$, *one has:*

$$\lim_{a \to 0} \frac{\mathbb{E}|T_X(a,k)|^q |T_X(a,k')|^q}{|a|^{2(q+\varphi(q))}} = C(q)|t-s|^{\varphi(2q)-2\varphi(q)}.$$

Hence, increasing the number of vanishing moments N_ψ does not impact the decrease of correlation.
The exponent $\varphi(2q) - 2\varphi(q)$ can be written as:

$$\varphi(2q) - 2\varphi(q) = \sum_{p \geq 2} c_p (2^p - 2) \frac{q^p}{p!}, \tag{11.60}$$

where the c_p are the log-cumulants. Therefore, the first log-cumulant c_1 (having the role that H has for H-sssi processes) does not intervene at all, and the decay of correlation is entirely controlled by the multifractal properties of the process, captured by $c_p, p \geq 2$. Also, for a log-normal MMC process (for which $\forall p \geq 3 : \ c_p \equiv 0$), the conjecture reduces to:

$$\varphi(2q) - 2\varphi(q) = c_2 q^2. \tag{11.61}$$

Therefore, taking the notation $\tilde{H} = \frac{\varphi(2q)-2\varphi(q)}{2} + 1$ of Section 11.1.1-b) (Eq. (11.9)): $\tilde{H} \to 1$ for $q \to 0$, $\tilde{H} \to 0.5$ for $q \to q_*^+/\sqrt{2} = \sqrt{-\frac{1}{c_2}}$ — hence, $0.5 < \tilde{H} < 1$ for $0 < q < q_*^+/\sqrt{2}$.

Since analytical results are limited to the ones derived in the previous subsections, we resort to empirical studies for validating the conjecture for other MMC processes, and for wavelet coefficients and Leaders. Results are obtained using numerical simulations, as described in Section 11.1.1-b) (wavelet based estimation on the time series $\{d_X(J, \cdots)\}$, $\{|d_X(J, \cdots)|^q\}$ or $\{L_X(J, \cdots)^q\}$ of wavelet coefficients or Leaders of the process $X(t)$ at scale J), using the following MMC processes and process parameter settings: CPM and CPM-MF-fBm ($\mu = -0.3$, $\sigma^2 = 0.0504$, $H = 0.7077$, $c = 1$; hence $(c_1, c_2, c_3) = (1.06, -0.14, 0.072)$ and $(0.75, -0.070, 0.036)$, respectively); mrw with $(c_1, c_2) = (0.8, -0.05)$ of sample size $N = 2^{16}$.
Results are presented in Fig. 11.3 and 11.4 (the results for CPM and mrw in Fig. 11.4 are obtained for larger sample size $N = 2^{21}$). Fig. 11.3 summarizes structure functions $S_J^d(j,2)$, $S_J^{|d|}(j,2)$ and $S_J^L(j,2)$ for $N_\psi = 1$ and $N_\psi = 2$ (first and second column, respectively), and mean estimations of H (Eq. (11.9)) as a function of vanishing moments $N_\psi \in \{1,2,3\}$ for d_X ($\triangle$, 11.3 only), $|d_X|$ ($\circ$) and L_X ($\times$) at scale $J = 3$ (third column) for CPM (top row), CPM-MF-fBm (center row) and mrw (bottom row). Fig. 11.4 presents mean estimations of $\tilde{H}$ (Eq. 11.9) for $|d_X|^q$ (blue solid lines) and L_X^q (red dashed lines) at scale $J = 3$ for CPM (top row), CPM-MF-fBm (center row) and mrw (bottom row), together with predictions from the conjecture Eq. (11.60) (fat black solid lines with dots), as a function of N_ψ for $q = 2$ (left column), and as a function of q for $N_\psi = 2$ (right column). Equivalent results, not reported here, can be obtained at different scales J.
Fig. 11.3 (left columns) shows that the structure functions $S_J^d(j,2)$, $S_J^{|d|}(j,2)$ and $S_J^L(j,2)$ do behave linearly in log-log coordinates for coarse scales, hence indicating scale invariance for the series of wavelet coefficients and Leaders and validating the estimation of the parameter $\tilde{H}$. Fig. 11.3 (right column) indicates that for $N_\psi = 1$, the time series of coefficients d_X possess LRD-like power law type correlations with $\tilde{H}_1^d \approx 0.75$ for CPM-MF-fBm and mrw, and $\tilde{H}_1^d \approx 0.95$ for CPM. For $N_\psi \geq 2$, the time series of coefficients display weak correlation only for all MMC processes considered here, with $\tilde{H}_{\geq 2}^d \approx 0.5$ (cf.

Fig. 11.3, right column). This is in agreement with the analytic result, obtained for increments and the process CPM, in Eq. (11.52), and with the arguments in Section 11.2.2-b). Fig. 11.3 (right column) and Fig. 11.4 (left column) suggest that the time series $|d_X|$ and $|d_X|^2$ do, in contrast, display (strong) LRD-like power law correlations, regardless of the precise choise of N_ψ: The estimated values for $\tilde{H}$ are between 0.8 and 0.9 and do not significantly decrease with increasing N_ψ. Also, Fig. 11.4 (left) shows that the estimated values for $\tilde{H}$ for different values of N_ψ are largely in agreement with the predictions from Conjecture 11.2.
Fig. 11.4 (right column) indicates that the time series of coefficients at orders q, $|d_X|^q$, are not SRD: They possess LRD-like power law type correlations, whose estimated $\tilde{H}$ can be large and close to 1. This confirms predictions from Conjecture 11.2, which are, for all orders q, in close agreement with estimations for CPM and mrw, and in satisfactory agreement for CPM-MF-fBm. Discrepancies between estimations and the conjecture can be explained by the fact that Conjecture Eq. 11.2 and the Eqs. (11.55) and (11.59) only contain the leading term $\varphi(2q) - 2\varphi(q)$ that dominates in the limit $\tau \to 0$ and $|t-s|$ large (respectively, $a \to 0$ and $|k-k'|$ large). In practice, the higher order terms play a role, since the limit can not be assessed numerically. For CPM-MF-fBm, these discrepancies are stronger because sample size is relatively small, due to limitations that the practical simulation of this process imposes – hence, we are practically further away from the limit.
Finally, Fig. 11.4 (right) suggests that the time series of wavelet Leaders L_X^q display correlations very similar to those of $|d_X|^q$, $\tilde{H}^{L^q} \approx \tilde{H}^{|d|^q}$. Hence, the same observations as for $|d_X|^q$ are obtained for L_X^q: Wavelet Leaders of MMC processes are in general not SRD but display strong LRD-like power law (higher order) correlations, regardless of the number of vanishing moments N_ψ. Also, estimations of $\tilde{H}$ for wavelet Leaders are in close agreement with predictions from the conjecture.

11.3 Discussion and Conclusion

The analytical and empirical results stated in the previous sections lead us to the following conclusions.
The key result for wavelet coefficients of *fractional Brownian motion* [70] is a *singular situation* and unfortunately *not general*: For fBm, wavelet coefficients are only weakly correlated if:

$$N_\psi > H + 1/2,$$

and Gaussianity implies short range dependence *at any order*.
For finite variance non Gaussian H-sssi processes and MMC processes with stationary increments, wavelet coefficients are also only *weakly correlated* if:

$$N_\psi > \frac{\zeta(2)}{2} + 1/2.$$

This is a consequence of the particular form of the covariance of such processes. However, for these processes, weak correlation *does not imply weak correlation at all orders*: The q-th order time series of coefficients $|d_X|^q$ display in general strong *LRD-like power law type correlation*. In particular, this is also the case for the time series d_X^2 involved in the estimation of the self-similarity parameter of H-sssi processes.
Moreover, LRD-like power law type correlations are *unaffected* by the precise *number of vanishing moments* N_ψ chosen for the analyzing wavelet, which proves to have no impact

on higher order correlation decay of wavelet coefficients.
These results have, to the best of our knowledge, never been clearly reported and investigated elsewhere and have significant implications in practical estimation procedures. Their consequences with respect to bootstrap estimation will be investigated in the next section.
Finally, the empirical results suggest that the dependence structure of time series of wavelet Leaders is similar to that of wavelet coefficients, since correlations of L_X^q are close to those of $|d_X|^q$. In particular, when LRD like power law type correlations are present for wavelet coefficients, their wavelet Leaders do not alter their decay.

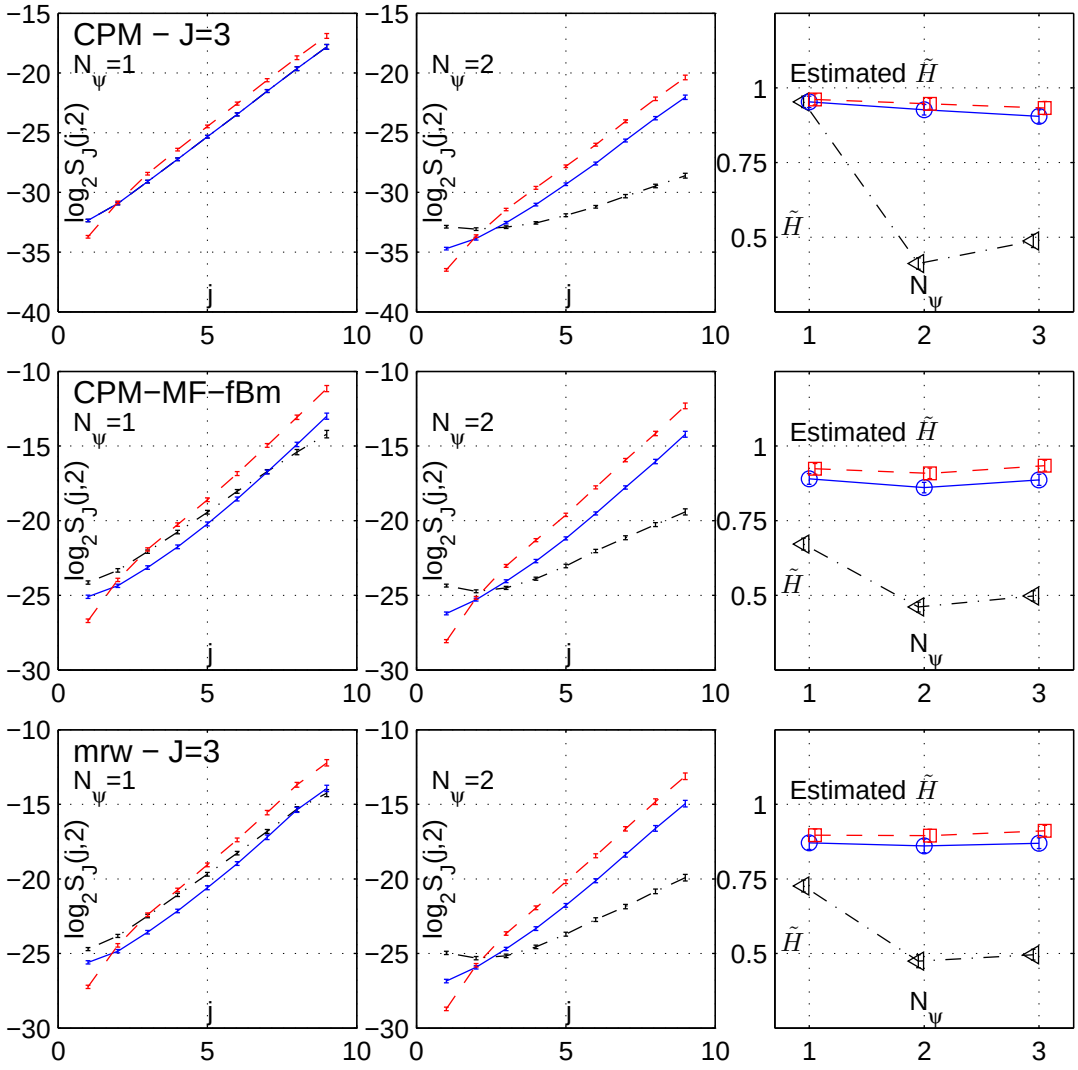


Figure 11.3: **Correlation of coefficients and Leaders of multifractal processes.** Structure functions $\log_2 S(j, q = 2)$ for wavelet coefficients of wavelet coefficients $d_X(J = 3, \cdot)$ (black dashed-dotted lines and '◁'), $|d_X(J = 3, \cdot)|$ (blue solid lines and 'O'), and of wavelet coefficients of wavelet Leaders $L_X(J = 3, \cdot)$ (red dashed lines and '□') for CPM (top row), CPM-MF-fBm (center row) and mrw (bottom row) with vanishing moments $N_\psi = 1$ (left columns) and $N_\psi = 2$ (center columns) of the wavelet analyzing the processes. The column on the right summarize the parameter $\tilde{H}$ (Eq. (11.9)) estimated on this structure functions, as a function of N_ψ of the wavelet analyzing the processes. Results are obtained as means over 100 realization of fBm and ROS of sample size $N = 2^{16}$, and with $N'_\psi = 4$ for analyzing $d_X(J = 3, \cdot)$, $|d_X(J = 3, \cdot)|$ and $L_X(J = 3, \cdot)$. Process parameters are set to $(\mu, \sigma^2) = (-0.3, 0.05)$ and $H = 0.708$ for CPM(-MF-fBm) and $(H, \beta^2) = (0.8, 0.05)$ for mrw.

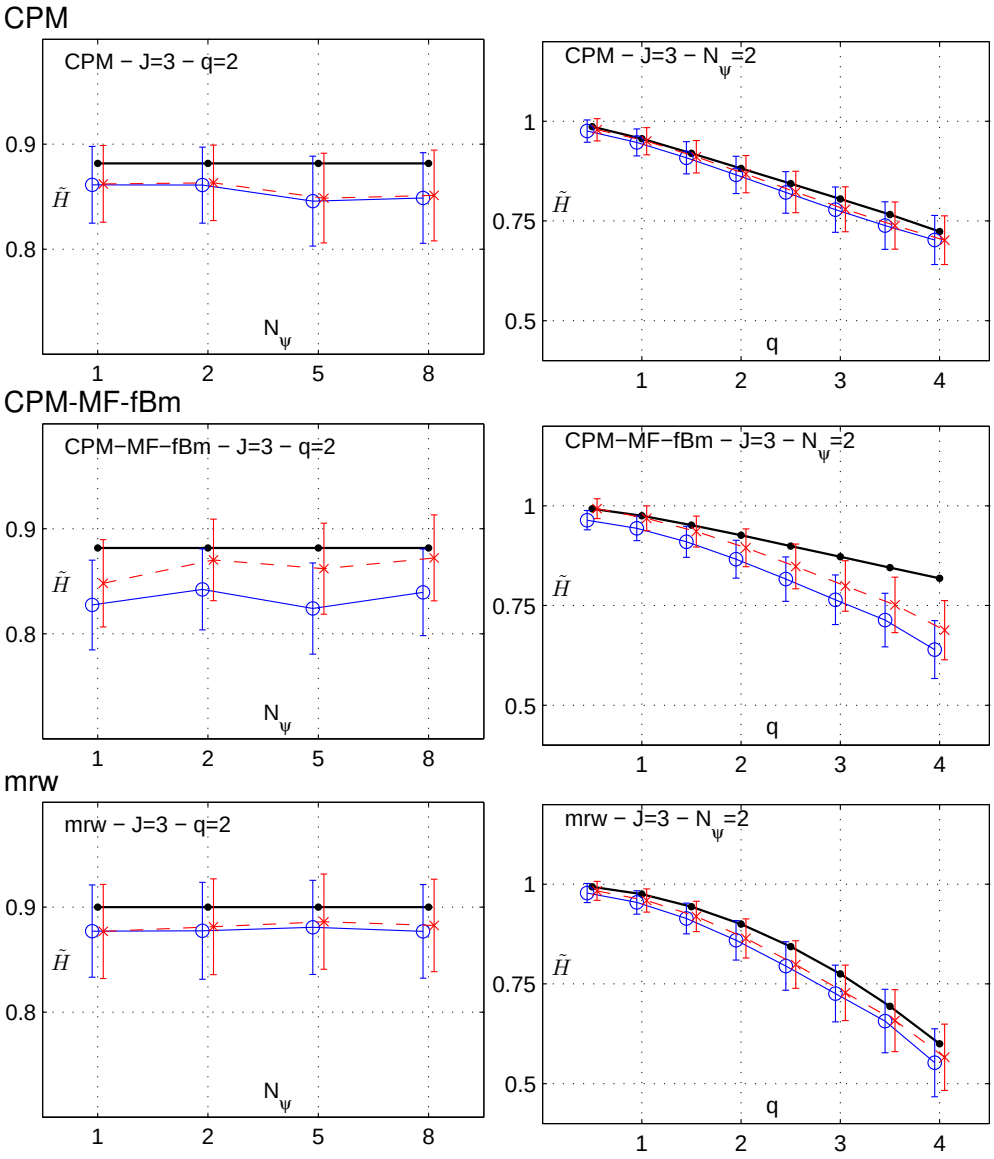


Figure 11.4: **Correlation of q-th order of coefficients and Leaders of multifractal processes.** Estimates of parameter $\tilde{H}$ (Eq. (11.9)) of wavelet coefficients $|d_X(J = 3, \cdot)|^q$ (blue solid lines and '$\circ$'), and of wavelet coefficients of Leaders $L_X(J = 3, \cdot)^q$ (red dashed lines and '$\times$') as a function of vanishing moments N_ψ of the wavelet analyzing the processes (left column, $q = 2$) and as a function of q (right column, $N_\psi = 2$) for CPM (top row), CPM-MF-fBm (center row) and mrw (bottom row). The black solid line with dots '$\bullet$' corresponds to the Conjecture 11.2. Shown are means and 1.96 standard deviation error bars obtained for 100 realizations of sample sizes $N = 2^{21}$ (CPM and mrw) and $N = 2^{18}$ (CPM-MF-fBm) of wavelet based ($N'_\psi = 4$) estimations of $\tilde{H}$. Process parameters are set to $(\mu, \sigma^2) = (-0.3, 0.05)$ and $H = 0.708$ for CPM(-MF-fBm) and $(H, \beta^2) = (0.7, 0.05)$ for mrw.

11.4 Implications for Wavelet-Domain Block Bootstrap

In the previous sections, we presented analytic and empirical results that show that wavelet coefficients and Leaders do in general – and with the singular exception of fractional Brownian motion – display complicated and severe dependence structures, mostly LRD-like power law type higher order correlations. These results potentially have important consequences for the wavelet domain block bootstrap estimation procedures we proposed in Chapter 7, and may shed new light on bootstrap estimation performance (cf. Chapters 8 to 10). The aim of this section is to investigate and clarify these issues.

11.4.1 Block bootstrap and long range dependence

Theory. Lahiri [103, 104] describes the behavior of block bootstrap methods when the time series violate the mixing conditions Eq. (6.2). More specifically, he considers the case of a zero mean real-valued stationary random process $\{X_n\}_{n\in\mathbb{Z}}$ that has autocovariance function $r(m) = \mathbb{E}X_iX_{i+m} \sim m^{-\alpha}$, $m \to \infty$ with $0 < \alpha < 1$ and that is therefore long range dependent (cf. Section 2.1.1-a), Def. 2.1 and Eq. (2.3)). Then, the variance of the sample mean $\bar{X}_N = \frac{1}{N}\sum_{n=1}^{N} X_n$ decays at rate slower than $O(N^{-1})$, and the scaling factor N/d_N for the centered sample mean:

$$T_N = N(\bar{X} - \mu)/d_N,$$

has to be of order smaller than $N^{1/2}$ in order to converge to a non-degenerate distribution. More precisely, it has to be defined as $d_N \sim \sqrt{N^{2-\gamma\alpha}}$, i.e. $N/d_N \sim N^{\gamma\alpha/2}$, where γ is the Hermite rank of the process $\{X\}$, which is defined as follows[4]. Suppose f is the instantaneous transform for which $\{X\}$ is obtained from a Gaussian process $\{Y\}$, $X_i = f(Y_i)$. Then, the Hermite rank γ of $\{X\}$ is the smallest $p \in \mathbb{N}$ for which the inner product of the Hermite polynomial[5] H_p of order p and the function $f' = f - \mathbb{E}X$ is not zero [62, 104, 164, 165]: $\gamma = \inf_p : \langle f', H_p\rangle \neq 0$.
Lahiri shows that the moving block bootstrap (MBB) version of the normalized sample mean:

$$T_N^* = N(\bar{X}^* - \bar{X})/d_N,$$

does not converge to the same limit distribution as normalized sample mean T_N. Whereas T_N has a non-degenerate limit distribution[6] for any $\gamma \geq 1$, T_N^* has a *degenerate* limit. Heuristically, this is because the bootstrap sample is composed of independent blocks, therefore forcing the variance of the bootstrap sample mean to go to zero too fast.
For the specific case when $\gamma = 1$ and hence when T_N has a Gaussian limit distribution, the problem can be circumvented by appropriate scaling of T_N^*, and the MBB version with modified scaling factor:

$$\tilde{T}_N^* = T_N^* \cdot \frac{d_N}{\sqrt{N/l \cdot d_l}}, \frac{d_N}{\sqrt{n/l \cdot d_l}} = F(\alpha, \gamma, N, l)$$

has the same limit distribution as T_N [103]. Note that the definition of such scaling factors requires that the LRD parameter α is known.

[4] Note that $0 < \alpha < 1/\gamma$ for a process of Hermite rank γ, cf. [62, 164, 165].

[5] The Hermite polynomial of order $p \in \mathbb{Z}^+$ is defined as $H_p(x) = (-1)^k \exp(x^2/2)\frac{d^k}{dx^k}\exp(-x^2/2)$, $x \in \mathbb{R}$.

[6] Under some additional conditions, cf. [62, 164, 165].

In the case $\gamma > 1$, no appropriate modification $\tilde{T}_N^*$ of the bootstrap version of the sample mean exists [103]. Therefore, any MBB version of the normalized sample mean has a limit distribution that is different from the limit distribution of T_N, and the MBB fails rather drastically.

Multifractal analysis. The non Gaussian finite variance H-sssi processes, such as ROS, are exactly in this framework and have Hermite rank $\gamma > 1$. Also, MMC processes do have LRD-like power law covariance (within the integral scale) and are hence in this framework. We conjecture that for most fv MMC processes, $\gamma = 1$ since they are close to log Normal processes whose Hermite rank is 1. We showed in the previous sections that wavelet coefficients $|d_X|^q$ and Leaders L_X^q of fv non Gaussian H-sssi and MMC processes potentially do have LRD-like power law dependence structure. Therefore, they are in the above framework. Although it is not clear how γ of X translates to the Hermite rank $\gamma^{|d_X|^q}$ and γ^{L^q} of $|d_X|^q$ and L_X^q, the results in [103] indicate that the wavelet domain block bootstrap procedures, which are defined *without* modified normalization constants d_N and d_l, are supposed to fail to reproduce the distribution of the structure functions:

$$\begin{aligned} S^d(j,q) &= \frac{1}{n_j}\sum_{k=1}^{n_j} |d_X(j,k)|^q \\ S^L(j,q) &= \frac{1}{n_j}\sum_{k=1}^{n_j} L_X(j,k)^q, \end{aligned}$$

regardless of the precise value of $\gamma^{|d_X|^q}$ and γ^{L^q}.

In contrast, the results in Chapters 8 to 10 show that bootstrap estimation performance is very satisfactory for scaling exponents (and log-cumulants):

$$\hat{\zeta}(q) = \sum_j w_j \log_2 S(j,q),$$

for a wide range of process parameters, sample sizes and statistical orders q. Therefore, it remains to be understood why empirical bootstrap estimation performance is satisfactory for $\hat{\zeta}(q)$, when it is supposed to theoretically fail to reproduce the limiting distribution of $S(j,q)$. This is the subject of the remainder of this section. In particular, we will investigate the following issues: What is the role of the $\log$, and does bootstrap estimation work for $Y(j,q) = \log_2 S(j,q)$? Which roles do the linear fits $\sum_j w_j Y(j,q)$ play?

11.4.2 Numerical simulation study

To answer such questions, we apply the estimation procedures for $S(j,q)$ and $\zeta(q)$ (Eqs. (2.57) and (2.65)) and the corresponding T-B and TS-B bootstrap estimation procedures (cf. Section 7.2) to a large number $N_{MC} = 1000$ of realizations of sample size N of (Gaussian H-sssi) fBm and (non Gaussian multifractal) mrw. We calculate (means $\widehat{\mathbb{E}}_{MC}$ over N_{MC} realizations of) bootstrap standard deviations:

$$\hat{\sigma}_Y^*(j,q) = \widehat{\mathbb{E}}_{MC}\, \widehat{\mathsf{Std}}^*\, \log_2 S^*(j,q), \tag{11.62}$$

$$\hat{\sigma}_\zeta^*(q) = \widehat{\mathbb{E}}_{MC}\, \widehat{\mathsf{Std}}^*\, \hat{\zeta}(q)^*, \tag{11.63}$$

and compare them to Monte Carlo standard deviations:

$$\hat{\sigma}_Y(j,q) = \widehat{\mathsf{Std}}_{MC} \log_2 S(j,q), \tag{11.64}$$

$$\hat{\sigma}_\zeta(q) = \widehat{\mathsf{Std}}_{MC} \hat{\zeta}(q), \tag{11.65}$$

obtained over Monte Carlo realizations. Process parameters are set to $H = 0.8$ for fBm, and $(c_1, c_2) = (0.8, -0.01)$ (weakly multifractal) and $(c_1, c_2) = (0.8, -0.1)$ (strongly multifractal) for mrw. The sample size is set to $N \in \{2^{10}, 2^{16}\}$, hence one small and one large sample size, the analyzing Daubechies wavelet has $N_\psi = 3$ vanishing moments, and weighted (w^1) linear regressions are performed over scales $(j_1, j_2) = (3, 7)$. Bootstrap estimations are obtained for $R = 99$ bootstrap samples per realization, and the block size l is varied over a large range: For T-B bootstrap estimation, the block sizes are $l_T \in \{1, 3, 6, 12, 24, 48, 96, 192, (384, 768, 1536, 3072)\}$ (the block sizes in brackets are only used for large sample size $N = 2^{16}$), and for TS-B, $l_{TS} \in \{16, 32, 64, 128, 256\}$ for $N = 2^{10}$, and $l_{TS} \in \{64, 128, 256, 512, 1024, 2048, 4096\}$ for $N = 2^{16}$. Results are reported in Fig. 11.5 to 11.8 and discussed below.

11.4.3 Bootstrap estimation and logarithm of structure functions

Fig. 11.5 summarizes Monte Carlo and bootstrap standard deviations $\hat{\sigma}_Y(j, q)$ and $\hat{\sigma}^*_Y(j, q)$ (Eqs. (11.64) and (11.62), respectively) of $\log_2 S^d(j, q = 2)$ (left column), $\log_2 S^L(j, q = -2)$ (center column) and $\log_2 S^L(j, q = 2)$ (right column). Bootstrap estimations are obtained from T-B (1st and 3rd line) and TS-B (2nd and 4th line) bootstrap samples. Results are shown as a function of scale j, and parametrized by block size l.

fBm. The results for fBm (Fig. 11.5, top) show that:

$$\hat{\sigma}^*_Y(j, q) \approx \hat{\sigma}_Y(j, q)$$

as soon as block sizes are large enough to capture residual dependence (typically, $l \geq 2N_\psi = 6$). This indicates that block bootstrap estimations for $\log_2 S(j, q)$ are efficient, and is in agreement with results in Section 11.1.1, showing that wavelet coefficients and Leaders of fBm are not LRD[7].

mrw. Results for mrw in Fig. 11.5 (bottom) demonstrate that:

$$\hat{\sigma}^*_Y(j, q) < 0.5\hat{\sigma}_Y(j, q),$$

and therefore, the T-B and TS-B bootstrap procedures fail to correctly estimate the standard deviations of $\log_2 S(j, q)$: Bootstrap standard deviation estimations are at best 50% of the actual (Monte Carlo) standard deviations. These best case results for $\log_2 S(j, q)$ are achieved with largest block size tested ($l = 1536$ for T-B and $l = 4096$ for TS-B) and are far from bootstrap estimation performance for multifractal attributes as reported in Chapters 8 to 10.

Conclusions. We conclude that bootstrap estimation for the logarithm of structure functions, $Y(j, q) = \log_2 S(j, q)$, works as expected for fBm, and fails rather drastically for mrw, for which the results obtained in Section 11.2 indicate that wavelet coefficients and Leaders at a given scale display LRD-like power law type dependence structure.

[7] We note that increased bootstrap estimation errors for large scales j are due to decreasing number of coefficients n_j with increasing scale, eventually forcing the adaptation of block size: As soon as the sample size is of the order of twice the block size, $n_j \approx 2l$, the block size is halved for this and all higher scales: $l_{new} = l_{old}/2$.

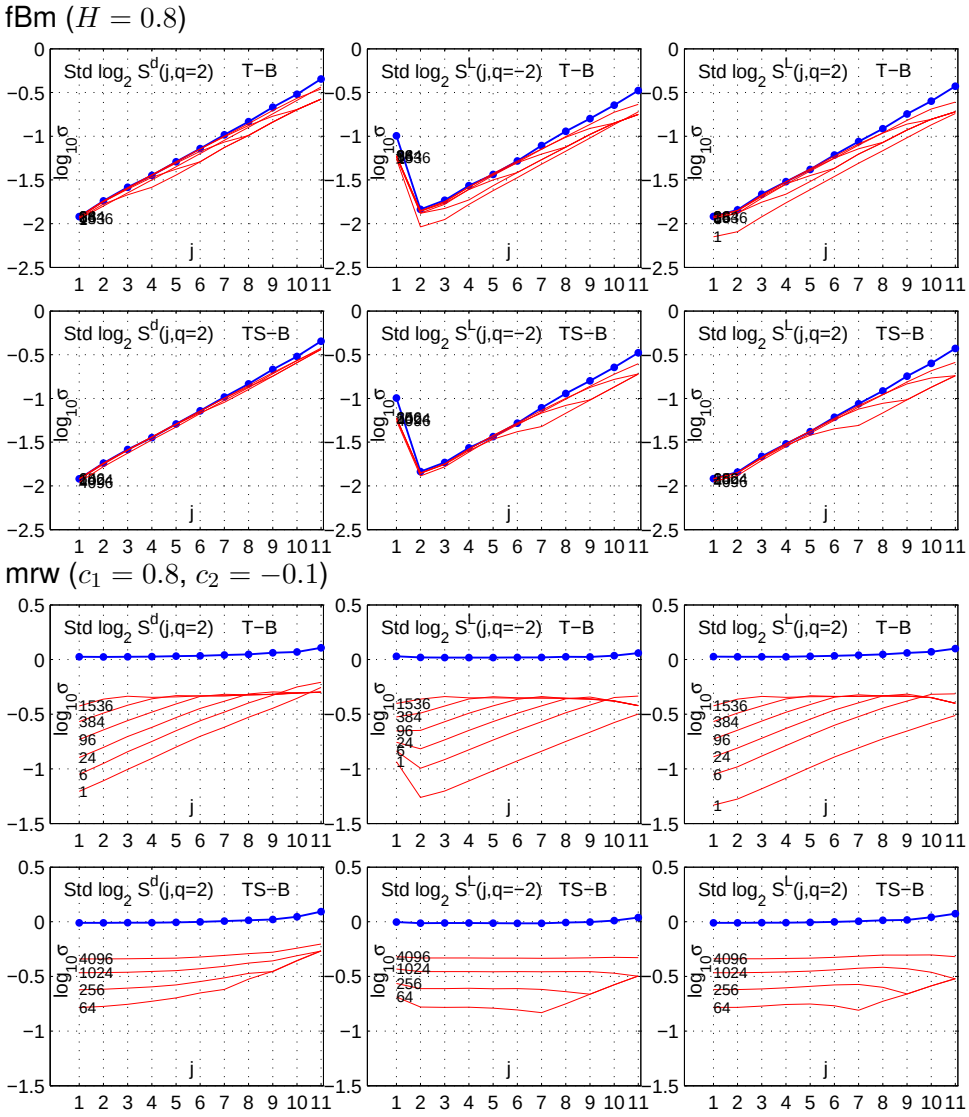


Figure 11.5: **Standard deviation and bootstrap estimation for $Y(j,q) = \log_2 S(j,q)$.** Monte Carlo standard deviation estimations of $Y(j,q) = \log_2 S(j,q)$, $\hat{\sigma}_Y(j,q)$ (bold blue solid lines, '•') and mean of bootstrap standard deviation estimations $\hat{\sigma}^*_Y(j,q)$ (red solid lines) for fBm (top, $H = 0.8$) and mrw (bottom, $(c_1, c_2) = (0.8, -0.1)$) and sample size $N = 2^{16}$. The first and third line correspond to T-B, the second and fourth to TS-B bootstrap estimation; the left, center and right column correspond to wavelet coefficient ($q = 2$) and Leader ($q = -2$ and $q = 2$) based estimation, respectively. Bootstrap estimation curves are indexed by the block size l employed.

11.4.4 Bootstrap estimation and linear fits: Scaling exponents

Fig. 11.6 summarizes Monte Carlo and bootstrap standard deviations $\hat{\sigma}_\zeta(q)$ and $\hat{\sigma}^*_\zeta(q)$ for $\zeta^d(q=2)$ (left column), $\zeta^L(q=-2)$ (center column) and $\zeta^L(q=2)$ (right column) for T-B bootstrap sampling as a function of block size l_T (graduated in $\log_2(l_T/N_\psi)$; the leftmost points correspond to $l=1$). The subplots correspond to fBm (top, $H=0.8$) and mrw ($(c_1,c_2)=(0.8,-0.01)$, center, and $(c_1,c_2)=(0.8,-0.1)$, bottom). Shown are results for different sample sizes $N=2^{10}$ and $N=2^{16}$ (first and second lines of the subplots, respectively). Fig. 11.7 shows the corresponding results for TS-B resampling (block size l_{TS} graduated in $\log_2(l_{TS})$).

fBm. Fig. 11.6 and 11.7 (top) illustrate that for fBm and for certain block sizes l_T and l_{TS} (discussed below):

$$\hat{\sigma}^*_\zeta(q) \approx \hat{\sigma}_\zeta(q),$$

for large sample size. For small sample size, bootstrap estimation performance remains satisfactory. This indicates that the T-B and TS-B block bootstrap estimation procedures are efficient for $\zeta(q)$ of fBm. These empirical findings are consistent with and confirm results of the previous Section 11.4.3, and Section 11.1.1.

mrw. Fig. 11.6 and 11.7 (center and bottom) show that for mrw:

$$\hat{\sigma}^*_\zeta(q) \approx \hat{\sigma}_\zeta(q),$$

for certain block sizes l_T and l_{TS} (cf. below). This is consistent with bootstrap estimation performance for multifractal attributes reported in Chapters 8 to 10. Together with the findings for $Y(j,q)$ in the previous subsection, these empirical results indicate that the linear fits in $\hat{\zeta}(q)=\sum_j w_j \log_2 S(j,q)$ must play an important role for bootstrap estimation for multifractal attributes of MMC processes.

Block size. Fig. 11.6 illustrates that the optimal block size l_T for T-B bootstrap estimation for $\zeta(q)$ slightly depends on the process type, parameters and sample size, but is of the order $\approx 2N_\psi$ (in agreement with the heuristic choice Eq. (7.6)). It is important to note that for mrw, the optimal l *decreases* with increasing sample size N or departure from monofractal (i.e., increasing $|c_2|$), hence bootstrap samples of coefficients capture *less* of their dependence: For large l_T, the T-B bootstrap *overestimates* the variance of $\zeta(q)$.
The best block size l_{TS} for TS-B bootstrap estimation for $\zeta(q)$ (cf. Fig. 11.6) corresponds exactly to the heuristic choice proposed in Eq. (7.7), regardless of sample size, process type and process parameters. Varying the block size changes performance only slightly, and there is no range of block sizes for which bootstrap variance estimations "overshoot": Block size variation consistently changes performance in the same way for both fBm and mrw, regardless of sample size and precise choice of process parameters.
Finally, for mrw and for both T-B and TS-B, optimal block sizes for estimation of $\hat{\sigma}_\zeta$ are *smaller* than for estimation of $\hat{\sigma}_Y$: Hence, best performance for $\hat{\sigma}_\zeta$ involve deliberate under-estimations for $\hat{\sigma}_Y$.

Conclusions. We conclude that bootstrap estimation for $\zeta(q)$ provides satisfactory results not only for fBm, but also for mrw. The TS-B procedure has an edge over T-B in terms of robustness with respect to changes in mono-/multifractal and sample size. Since for mrw, bootstrap estimation for $S(j,q)$ and $Y(j,q)$ are not consistent, the linear fits must play an important role for satisfactory estimation performance for $\zeta(q)$.

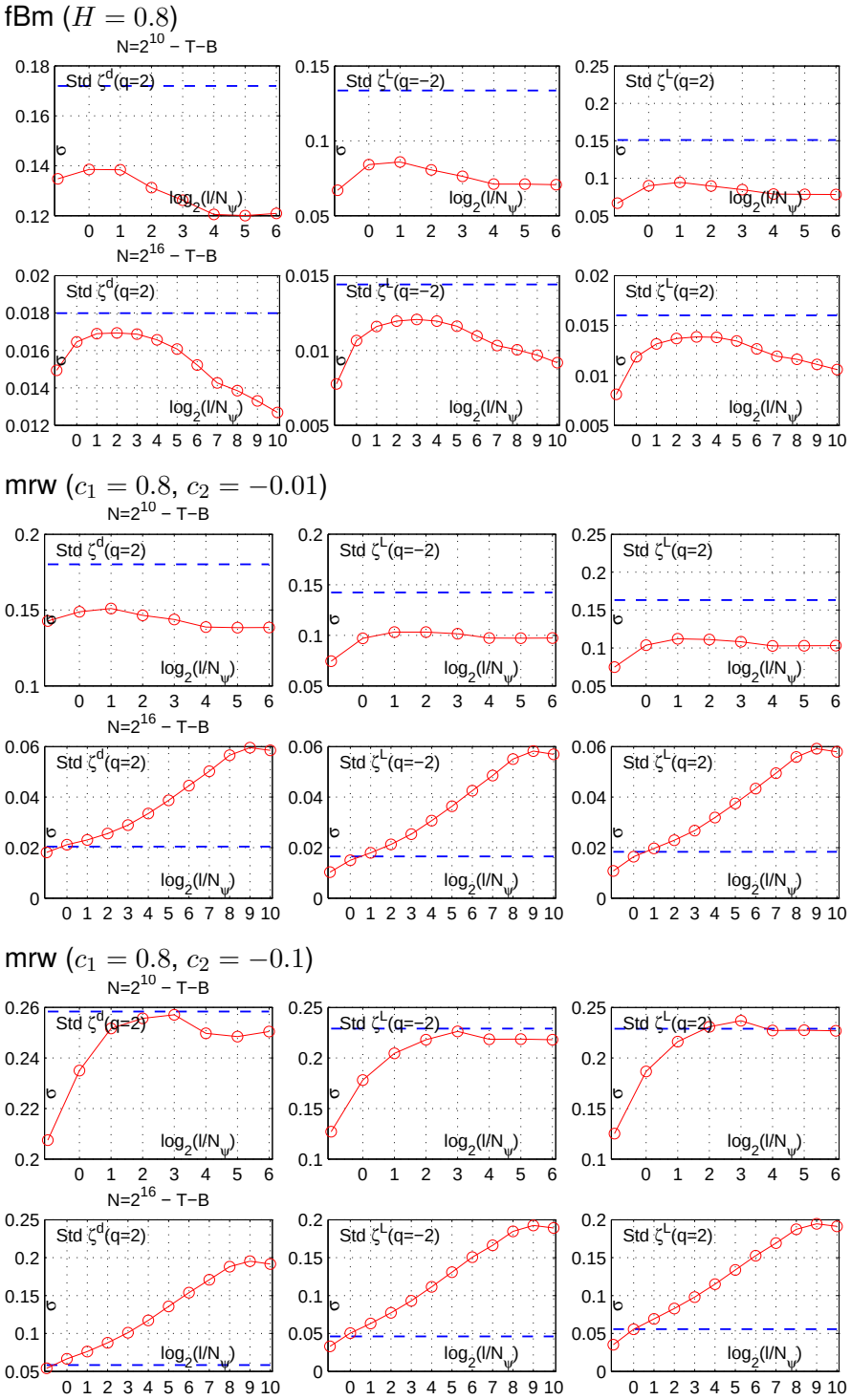


Figure 11.6: **Scaling exponent standard deviation and T-B bootstrap estimation.** Monte Carlo standard deviation estimations $\hat{\sigma}_\zeta(q)$ (blue dashed lines) and means of T-B bootstrap standard deviation estimations $\hat{\sigma}^*_\zeta(q)$ (red solid lines and '○') for $\zeta(q)$ as a function of block size l: fBm (top, $H = 0.8$), mrw (center, $(c_1, c_2) = (0.8, -0.01)$), mrw (bottom, $(c_1, c_2) = (0.8, -0.1)$). The first line corresponds to sample size $N = 2^{10}$, the second line to sample size $N = 2^{16}$. The left, center and right column correspond to wavelet coefficient ($q = 2$) and Leader ($q = -2$ and $q = 2$) based estimation, respectively.

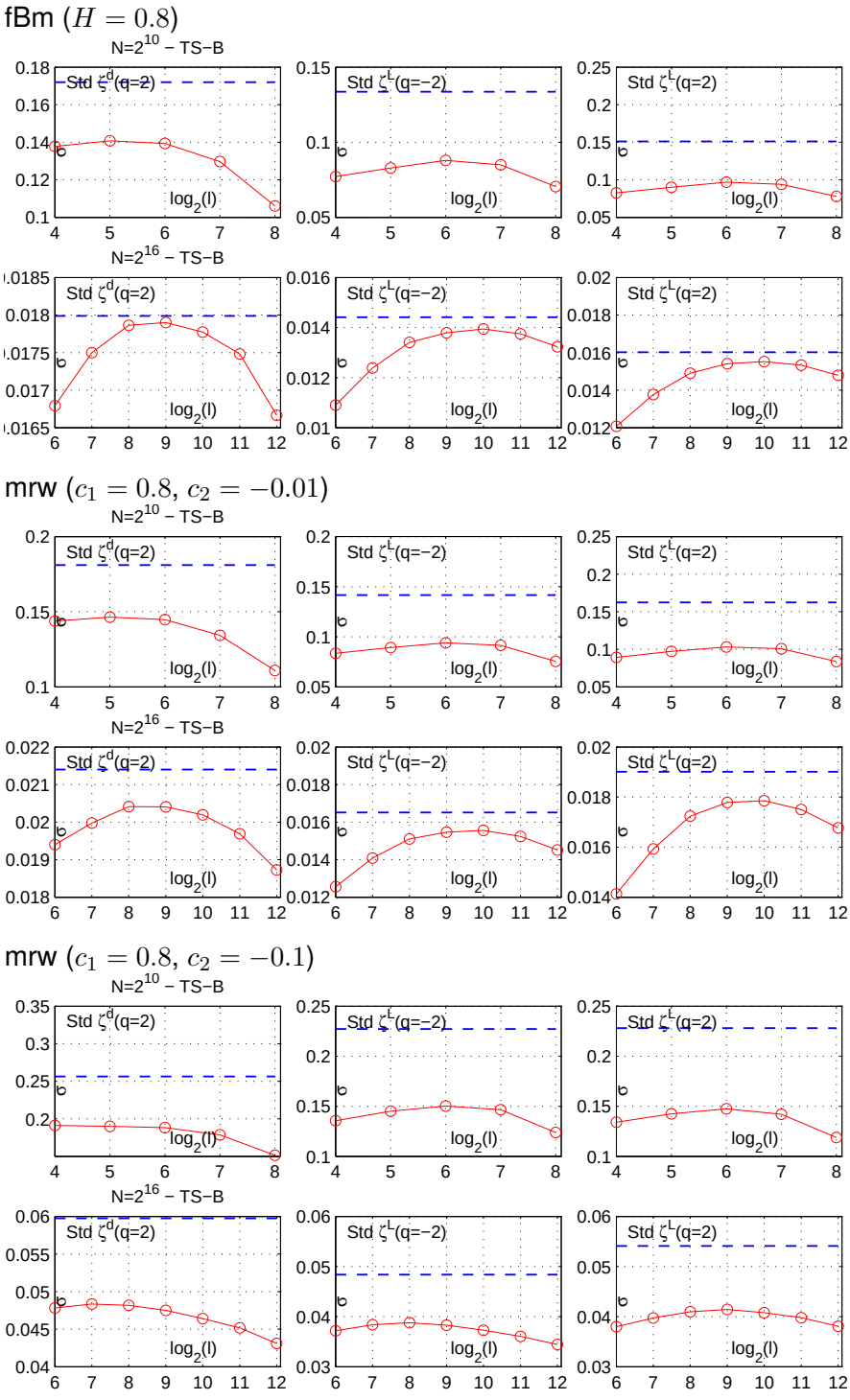


Figure 11.7: **Scaling exponent standard deviation and TS-B bootstrap estimation.** Monte Carlo standard deviation estimations $\hat{\sigma}_\zeta(q)$ (blue dashed lines) and means of TS-B bootstrap standard deviation estimations $\hat{\sigma}^*_\zeta(q)$ (red solid lines and '○') for $\zeta(q)$ as a function of block size l: fBm (top, $H = 0.8$), mrw (center, $(c_1, c_2) = (0.8, -0.01)$), mrw (bottom, $(c_1, c_2) = (0.8, -0.1)$). The first line corresponds to sample size $N = 2^{10}$, the second line to sample size $N = 2^{16}$. The left, center and right column correspond to wavelet coefficient ($q = 2$) and Leader ($q = -2$ and $q = 2$) based estimation, respectively.

11.4.5 Bootstrap estimation and inter-scale dependence

From the definition of $\hat{\zeta}(q)$ (cf. Eq. (2.65)):

$$\hat{\zeta}(q) = \sum_j w_j Y(j,q) \quad \left(= \sum_j w_j \log_2 \left(\frac{1}{n_j} \sum_{k=1}^{n_j} |T_X(j,k)|^q \right) \right),$$

it is clear that the variance $\hat{\sigma}_\zeta^2(j,q)$ of $\hat{\zeta}(q)$ and its bootstrap estimation $\hat{\sigma}_\zeta^{2*}(j,q)$ can be written as:

$$\hat{\sigma}_\zeta^2(q) = \sum_j w_j^2 \hat{\sigma}_Y^2(j,q) + \sum_j \sum_{j' \neq j} w_j w_{j'} \mathrm{Cov} Y(j,q) Y(j',q), \tag{11.66}$$

$$\hat{\sigma}_\zeta^{2*}(q) = \sum_j w_j^2 \hat{\sigma}_Y^{2*}(j,q) + \sum_j \sum_{j' \neq j} w_j w_{j'} \mathrm{Cov}^* Y^*(j,q) Y^*(j',q). \tag{11.67}$$

Thus, it is composed of two terms: The first term $\sum_j w_j^2 \hat{\sigma}_Y^{2(*)}(j,q)$ is entirely due to the variances $\hat{\sigma}_Y(j,q)$ of the logarithm of the structure functions (and their bootstrap counterparts, respectively). The second term is due to inter-scale dependence (ISD) of the multiresolution quantities and will be called the ISD term in what follows.
We can investigate the contribution of the ISD term indirectly by comparing the actual variances $\hat{\sigma}_\zeta^2(q)$ and $\hat{\sigma}_\zeta^{2*}(q)$ to the versions:

$$\tilde{\sigma}_\zeta^2(q) = \sum_j w_j^2 \hat{\sigma}_Y^2(j,q), \tag{11.68}$$

$$\tilde{\sigma}_\zeta^{2*}(q) = \sum_j w_j^2 \hat{\sigma}_Y^{2*}(j,q), \tag{11.69}$$

which assume independence between different scales, hence zero ISD.
Fig. 11.8 shows results for $\hat{\sigma}_\zeta(q)$ and $\tilde{\sigma}_\zeta(q)$ (blue dashed and black dashed-dotted horizontal lines, respectively), and for their TS-B bootstrap counterparts $\hat{\sigma}_\zeta^*(q)$ and $\tilde{\sigma}_\zeta^*(q)$ (red solid lines with circles and red dashed-dotted lines with crosses, respectively) as a function of block size l_{TS} (i.e. $\log_2(l_{TS})$). Results are obtained for fBm (top) and mrw (bottom, $(c_1, c_2) = (0.8, -0.1)$) and sample size $N = 2^{16}$. Note that results for T-B bootstrap estimation are not shown, since T-B bootstrap samples entirely blind out any ISD by construction and therefore: $\hat{\sigma}_\zeta^*(q) \equiv \tilde{\sigma}_\zeta^*(q)$.

fBm. We observe that for fBm, ISD slightly increases actual variances of $\zeta(q)$ since:

$$\tilde{\sigma}_\zeta(q) < \hat{\sigma}_\zeta(q).$$

The increase is of the order of 5% (12%) for wavelet coefficients (Leaders), thus slightly larger for wavelet Leaders.
For the block size choice $l_{TS} = 2^{10}$ (cf. above and Eq. (7.7)):

$$\hat{\sigma}_\zeta^*(q) - \tilde{\sigma}_\zeta^*(q) \approx \hat{\sigma}_\zeta(q) - \tilde{\sigma}_\zeta(q).$$

Therefore, the TS-B bootstrap samples completely capture the ISD for fBm.

mrw. For mrw, we observe that ISD has a very pronounced *negative* contribution to actual variances of $\zeta(q)$:

$$\tilde{\sigma}_\zeta(q) >> \hat{\sigma}_\zeta(q),$$

and the inter-scale covariance contribution strongly *decreases* variances of linear fits: $\hat{\sigma}_\zeta(q) \approx \frac{1}{8}\tilde{\sigma}_\zeta(q)$.
In contrast to what is observed for fBm, TS-B bootstrap samples capture the ISD for mrw *only partially*. The fraction that is captured in the bootstrap samples increases with increasing blocklength l_{TS} and is of the order $\hat{\sigma}^*_\zeta(q) \approx \frac{1}{3}\tilde{\sigma}^*_\zeta(q)$ ($l_{TS} = 2^{10}$).

Conclusions. We conclude that for fBm, TS-B bootstrap samples capture all the inter-scale dependence present in the original sample of coefficients or Leaders. For mrw, however, they capture ISD only partly. T-B bootstrap samples ignore all ISD by construction.

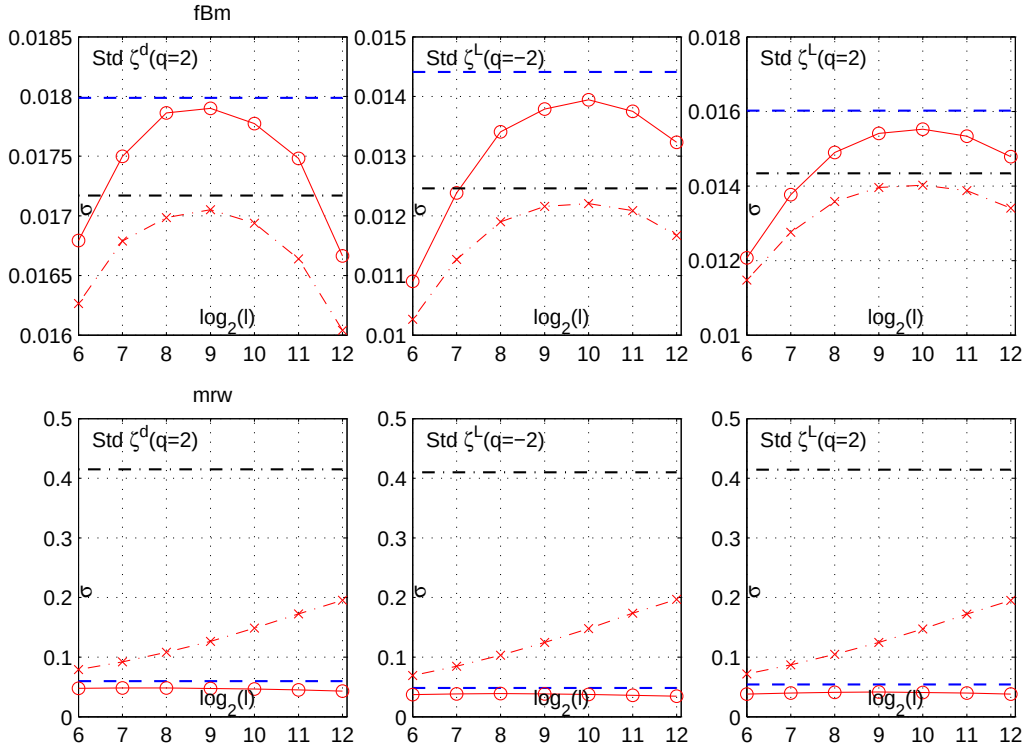


Figure 11.8: **Interscale dependence and bootstrap estimation: TS-B BS.** Standard devations $\hat{\sigma}_\zeta(q)$ and $\tilde{\sigma}_\zeta(q) = \sum_j w_j^2 \hat{\sigma}_Y(j,q)$ (blue dashed and black dashed-dotted horizontal lines, respectively), and TS-B bootstrap estimations $\hat{\sigma}^*_\zeta(q)$ and $\tilde{\sigma}^*_\zeta(q) = \sum_j w_j^2 \hat{\sigma}^*_Y(j,q)$ (red solid lines with '○', and red dashed-dotted lines with '×', respectively) as a function of block size l_{TS} (graduated as $\log_2 l_{TS}$). Results are obtained for fBm (top, $H = 0.8$) and mrw (bottom, $(c_1, c_2) = (0.8, -0.1)$) and sample size $N = 2^{16}$ for wavelet coefficients ($q = 2$, left column) and Leaders ($q = -2$ and $q = 2$, center and right column, respectively).

11.4.6 Discussion, conclusions and perspectives

The empirical findings presented above lead to the following conclusions. For fBm, dependence are (very) weak, and the proposed bootstrap procedures do not encounter any theoretical or practical difficulties: The variances of $\log_2 S(j,q)$ are correctly estimated since the range of time dependence of coefficients is short. The ISD term in linear fits, arising from inter-scale dependence of coefficients, is correctly reproduced by the TS-B resamples. Since it is very small, the error introduced by the T-B procedure can be ignored, and both block bootstrap procedures correctly estimate the variances of $\hat{\zeta}(q)$.
For MMC processes, the bootstrap procedures fail to estimate the variances of $\log_2 S(j,q)$, and bootstrap variance estimates $\hat{\sigma}^*_Y(j,q)$ are significantly smaller than actual (Monte Carlo) variances $\hat{\sigma}_Y(j,q)$. Nevertheless, bootstrap variance estimations for $\hat{\zeta}(q)$ are highly satisfactory. This can be interpreted as follows. For such processes, the ISD term is very significant and has a very strong *negative* contribution to variances of $\hat{\zeta}(q)$. Due to the LRD-like power law type dependence structure, the bootstrap samples can not reproduce the entire time dependence, and idem for the ISD, and the two under-estimation effects compensate each other for certain block sizes:

- For T-B bootstrap samples, who *entirely* ignore ISD, the block size has to be chosen very small, resulting in very strongly underestimated variances $\hat{\sigma}_Y(j,q)$ to compensate for completely blinding out the ISD term. It is shown above that this compensation effect, taking place for block size $l_T \approx 2N_\psi$, is rather stable with respect to sample size and process parameters.

- The TS-B bootstrap samples reproduce the ISD *partly*. This underestimation of ISD precisely counterbalances the effect of underestimation of $\hat{\sigma}_Y(j,q)$. This compensation is, in certain limits, robust with respect to block size l_{TS}. Most importantly, for the block size l_{TS} as defined in Eq. (7.7), this happens *independently* of sample size N and process parameters.

From these empirical findings, we can of course not conclude whether estimations are *asymptotically* consistent for $N \to \infty$, and theoretical results would be needed. Heuristically, for the T-B bootstrap, there is empirical evidence against this by noting that when passing from weakly (and small sample size) to strongly multifractal (and large sample size), the best performance block size reduces to $l_T \to 1$ to compensate for increased importance of ISD – this is the lower limit for block size and hence leaves no more margin for when departure from monofractal and sample size are further increasing. For the TS-B bootstrap, empirical evidence is in favor of asymptotic consistency: Results are robust with respect to changes in block size, and stable with respect to sample size increase.
To conclude, empirical results show that the block bootstrap estimations for multifractal attributes have satisfactory performance for the sample sizes and process parameter ranges typically encountered in practice and considered here for numerical study. Theory and asymptotic results as well as further numerical results are needed for a better understanding of these empirical findings for MMC processes, and for potential improvements of the block constructions, or for definitions of alternative inference procedures.

Part III

Bootstrap and Multifractal Analysis: Applications

In this work, we are also interested in a number of applications, involving real-world data of very different nature, some of which will be presented in more detail in this part of the manuscript. The aim is to demonstrate how the above developed methods can be fruitfully used for the analysis of real world data and in applications.

The tools we propose in this work have been used in an application aiming at detecting, from functional Magnetic Resonance Imaging (fMRI) data, which parts of the human brain are activated by a given stimulus type. This work was conducted in collaboration with the group of P. Ciuciu at NeuroSpin, CEA, and at Functional Neuroimaging Institute, Paris, France. They designed and performed the experiments and the fMRI data measurements, as well as the final detection step. Detection is based on multifractal attribute estimates, obtained from voxel-per-voxel fMRI time series with the tools proposed in this work. The tools enable, first, to evidence multifractal signatures in the fMRI signals, which is an original result in itself and has never been clearly stated before. Second, it provides relevant attribute estimates for the effective discrimination of ongoing ("background") brain activity against task-related brain activity. This application is reported in [50, 51].

Another application we considered involves the multifractal analysis of baro-reflex data, in collaboration with E. Pereira de Souza Neto (anesthesiologist at the Hospital of Lyon, France). The data consist of jointly recorded arterial pressure and cardiac rhythms. The application aims at differentiating between baro-reflex activity under different predefined medical conditions, and at practically characterizing these different situations. To account for the specific nature of the bivariate data in this application, we propose a wavelet inter-spectrum estimator, acting as a scale invariance - adapted alternative to a (Fourier domain based) inter-spectrum. This specific tool is not reported in this manuscript. Together with the multifractal analysis procedures proposed here, it enables a relevant characterization of baro-reflex data and can be fruitfully used in standard test procedures commonly accepted in the biomedical community. The publication of the method and the results for this application are in preparation.

Furthermore, the methods proposed in this work were involved in an application concerning the characterization of turbulence intermittency data, and for the classification of texture images. We choose to present these applications in detail below in Chapter 12 and Chapter 13, respectively.

Chapter 12

Application to 1d Signals: Turbulence

Contents

Most of the seminal ideas of the concept of multifractal were introduced in the study of hydrodynamic turbulence. One of the first models proposed for the description turbulence flows was based on monofractal fractional Brownian motion [100]. It is nowadays commonly accepted that velocity or dissipation turbulence fields are better modeled with multifractal multiplicative cascades. A number of multifractal models have been proposed, the most prominent being the log-Normal model [99], the log-Poisson (or She-Lévêque) model [159], and the log-stable cascade model [155]. However, it remains to validate which of these multifractal models better fits the data. A relevant answer to this contended question can contribute to a better theoretical understanding of the physical mechanisms responsible for occurrence of turbulence flows.
The goal of this chapter is to revisit this issue by applying wavelet Leader based multifractal analysis and bootstrap hypothesis tests to real-life data from two major hydrodynamic turbulence experiments. We choose to concentrate on the log-Normal and the log-Poisson multifractal model. Although discrimination of the log-stable cascade model would be of interest, it is not considered here since the log-cumulant expansion Eq. (2.70) is not relevant for this cascade (cf. Section 2.6.2).
The work presented in this chapter has been conducted in collaboration with Stéphane G. Roux (ENS Lyon) and reported in [179].

12.1 Multifractal Analysis in Hydrodynamic Turbulence

12.1.1 Turbulence and scaling: A short survey

Fluid motions are described by (partial differential) equations, such as Navier-Stokes, mass conservation, state and thermodynamic equations. However, for most natural flows, the fluctuations of the fluid parameters (velocity, density, pressure, temperature) appear highly erratic, unpredictable and random. After the early work of Richardson in the twenties [144], the heuristic understanding of hydrodynamic turbulence relates these erratic fluctuations to a transfer of energy from large flow scales (where it is injected by an external forcing) to small flow scales (where it is dissipated by viscous friction mechanisms). For thorough introductions to turbulence, the reader is referred to, e.g., [75, 128]. This energy cascade based heuristic analysis of turbulence flows is deeply associated with scale invariance: Between the coarse injection scale and the fine dissipation scale, no characteristic scale can be identified. It lead to the use of stochastic processes with built-in scaling properties for turbulence modeling. In 1941, Kolmogorov proposed one of the first stochastic descriptions of turbulence based on fractional Brownian motion, a Gaussian self-similar, hence strictly monofractal ($c_1 = H = \frac{1}{3}, c_p \equiv 0, p \geq 2$) process [100]. However, after the seminal work by Yaglom, the energy transfer from coarse to fine scales mechanism has often been modeled via split/multiply iterative random procedures, that match the physical intuitions beyond the vorticity stretching mechanisms at work in turbulence flows [75]. Mandelbrot in the seventies fruitfully gathered these models in the unified framework of multiplicative martingales and studied their properties [121] (see also [95]). Nowadays most practitioners agree on the existence of scale invariance in turbulence data and, following the analyses of Parisi and Frisch [76], on its multifractal nature. However, a major open issue consists in deciding which particular multifractal process better models turbulence flows. A large variety of cascades has been proposed over the last 30 years, each trying to better fit experimental data and/or to

Model	c_1	c_2	c_3
LN	0.3708	-0.0250	0.0000
LP	0.3814	-0.0365	0.0049

Table 12.1: **Multifractal models for turbulence.** Canonical values of c_p for the log-Normal Obukov-Kolmogorov 62 (LN) and She-Lévêque log-Poisson (LP) models.

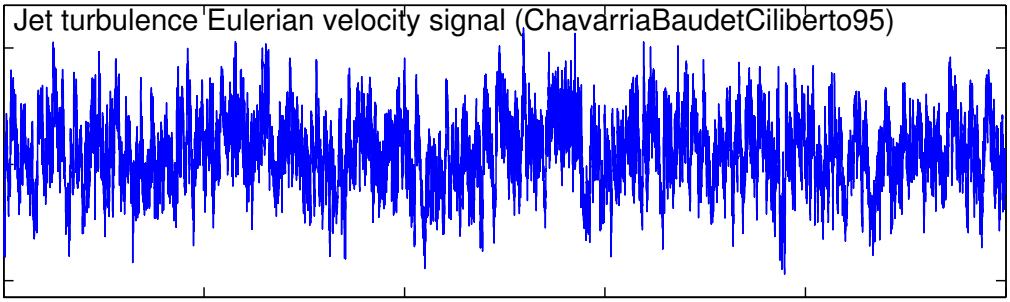


Figure 12.1: **Jet turbulence signal.** A part of the jet turbulence longitudinal Eulerian velocity signal in [150].

better account for a specific fluid flow property and yielding a different prediction for the multifractal spectrum. Let us concentrate on two amongst the most popular such models. In 1962, Obukhov and Kolmogorov [99, 133] proposed a model mostly based on a law of large numbers argument and referred to as the log-Normal multifractal model. It predicts that $\zeta(q) = c_1 q + c_2 q^2/2$ and hence that $c_p \equiv 0$ for $p \geq 3$. More recently, She and Lévêque [159] proposed an alternative construction based on the central assumption that energy dissipation gradients must remain finite within turbulence flows. It is referred to as the log-Poisson model and yields a multifractal process with all non zero c_ps. The canonical values of c_1, c_2, c_3 for the log-Normal Obukov-Kolmogorov 62 and She-Lévêque log-Poisson models are given in Tab. 12.1.
Discriminating between the log-Normal and log-Poisson models hence requires the use of tools providing us with an accurate estimate for the c_3 parameter and with a statistical test aiming at rejecting the null hypothesis $H_0 : c_3 \equiv 0$. These are, according to the numerical results reported in Chapter 4, provided by wavelet Leaders[1], and by the nonparametric bootstrap hypothesis tests in Chapter 9.

12.1.2 Turbulence data and estimation parameters

Data description. In the present contribution, we analyze large turbulence data sets from two different experiments. They consists of high quality, high sampling rate and long observation duration longitudinal Eulerian velocity signals, measured with hot-wire anemometry techniques. The first set is obtained from a jet turbulence experiment, with approximate (Taylor scale based) Reynolds number $R_\lambda \simeq 580$, [150]. A part of this signal is shown in Fig. 12.1. It has been made available to us by C. Baudet (LEGI, Université Joseph Fourier, INPG, CNRS, Grenoble, France). The second data set consists of wind-

[1]Wavelet Leaders have also been used for the measurement of multifractal spectra of major turbulence data sets in [111], in a similar perspective.

Data set	Jet	Wind-Tunnel
Taylor based Reynolds number	$\simeq 580$	$\simeq 2000$
Number of runs	$= 79$	$= 24$
Duration (in samples) per run	$= 2^{20}$	$= 2^{20}$
Duration (in integral scales)	$\simeq 2^7$	$\simeq 2^7$
Integral Scale (in samples)	$\simeq 2^{13}$	$\simeq 2^{13}$
Taylor Scale (in samples)	$\simeq 2^6$	$\simeq 2^4$

Table 12.2: **Turbulence data sets.** Description of the 1d Eulerian velocity data sets analyzed in this contribution.

tunnel turbulence (1995 campaign, cf. [94]), with $R_\lambda \simeq 2000$, made available to us by Y. Gagne (LEGI, Université Joseph Fourier, INPG/CNRS, Grenoble, France). The description of the data is summarized in Tab. 12.2 (it is worth mentioning that these data sets consist of 79 and 24 million samples, respectively!).

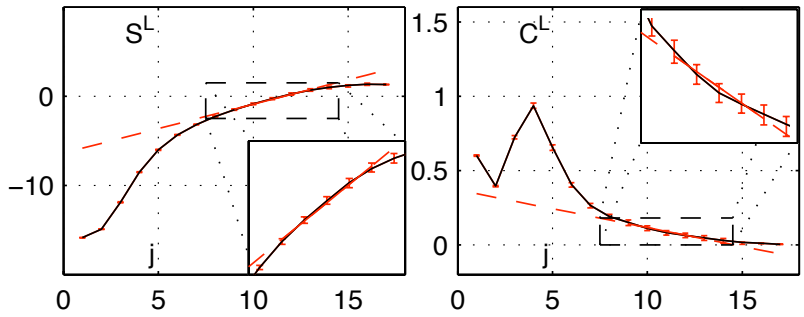


Figure 12.2: **Turbulence data structure functions.** Structure functions (on the left, $q = 2$) and Cumulant estimates (on the right, $p = 2$), based on wavelet Leaders, for one single run of jet turbulence data ($N = 2^{20}$). The bootstrap confidence limits (in solid red) are obtained with the percentile method and $\alpha = 0.05$. The confidence limits confirm the choice of regression range, $j \in [9, 13]$.

Scaling range selection. According to turbulence common understanding, the power law behavior associated with scale invariance takes place in the so called *inertial range* of scales. It spreads from above the Taylor scale to below the integral scale, which are estimated at (in sample numbers) 2^6 and 2^{13} for the jet data set and 2^4 and 2^{13} for the wind tunnel data set, respectively. The corresponding turbulence velocity structure functions, for one jet turbulence run with $N = 2^{20}$, are plotted in Fig. 12.2 for illustration purposes. However, choosing the range of scales where the linear regression needs to be performed remains an involved question. Two categories of arguments, leading to slightly different regression ranges, are opposed: Use of a goodness-of-fit tool, e.g. based on confidence intervals as shown in Fig. 12.2, to select a range of scales well inside the inertial range where the power law model holds; Selection of a range of scales such that $\zeta(3) = 1$ (as this is a theoretical requirement, the so-called Karman-Howarth results [75], in the ideal case of 3D data collected on homogeneous isotropic stationary turbulence). The latter choice amounts to performing the linear fit in a slightly lower range of scales than the former one. The use of wavelet Leaders to perform multifractal anal-

Jet turbulence data			
p	1	2	3
$\hat{c}_p$	0.3041	−0.0206	−0.0001
low$_{per}$	0.2866	−0.0255	−0.0020
high$_{per}$	0.3214	−0.0158	0.0018
low$_{stu}$	0.2835	−0.0259	−0.0023
high$_{stu}$	0.3257	−0.0140	0.0023
Wind tunnel data			
p	1	2	3
$\hat{c}_p$	0.3515	−0.0259	0.0006
low$_{per}$	0.3421	−0.0293	−0.0011
high$_{per}$	0.3609	−0.0223	0.0024
low$_{stu}$	0.3408	−0.0296	−0.0014
high$_{stu}$	0.3622	−0.0216	0.0028

Table 12.3: **Turbulence data multifractal attribute estimates.** Jet turbulence (top) and wind tunnel turbulence (bottom) wavelet Leader based estimates of log cumulants c_p, together with T-B bootstrap confidence limits. The results are averaged over runs.

ysis makes this question even more involved. In the present work, we chose the former solution and will not further discuss this issue which remains, however, strongly debated amongst practitioners. Careful analysis of structure functions and confidence intervals leads to the choices $(j_1, j_2) = (9, 13)$ for the jet turbulence data set and $(j_1, j_2) = (6, 10)$ for wind tunnel turbulence.

Estimation parameters. The remaining estimation parameters are set to: ordinary linear regressions (w^0), Daubechies wavelets with $N_\psi = 3$, and T-B bootstrap with $R = 399$, $S = 50$, $\alpha = 0.05$.

12.1.3 Results and conclusions

From the structure functions (cf. Fig. 12.2), estimates for the multifractal attributes $\zeta(q), D(h), c_p$ are computed for each run with $N = 2^{20}$. They are shown, for a single run, in Fig. 12.3, together with their 95% T-B bootstrap percentile confidence intervals. Estimates and confidence intervals confirm that the data are multifractal: $\zeta(q)$ is not a linear function of q, the multifractal spectrum $D(h)$ has support on an whole range of Hölder exponents h, and the confidence interval for c_2 excludes the zero value.

Wavelet Leader based estimates for c_1, c_2, c_3 averaged across the entire data sets for both jet and wind tunnel turbulence are reported in Tab. 12.3, together with their T-B bootstrap based confidence intervals. It shows that confidence limits based on percentile or studentized statistics are extremely close. Furthermore, it indicates that the estimated c_p are close but not equal for the two data sets: This can be either due to difficulties in the regression range selection or to the difference in Reynolds numbers (and can hence be related to the much debated issue in turbulence of universal values for multifractal

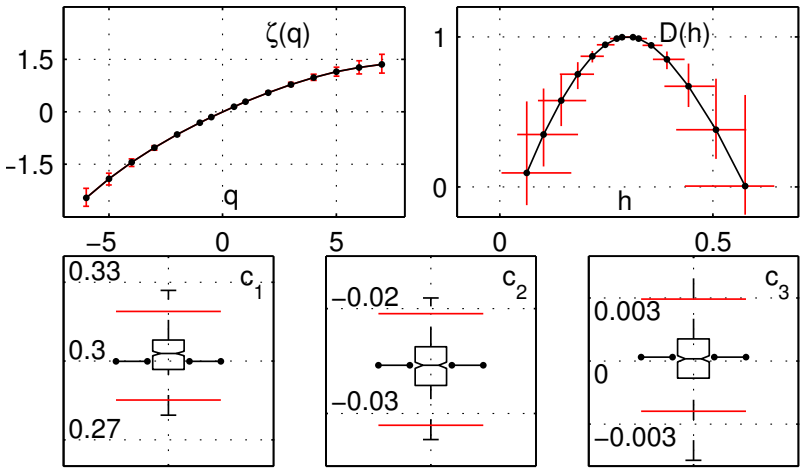


Figure 12.3: **Turbulence data multifractal attribute estimates.** Estimates (in solid black, '•') of scaling exponent $\zeta(q)$ (top left), multifractal spectrum $D(h)$ (top right) and log cumulants c_p (bottom row), based on wavelet Leaders, for one single run of jet turbulence data ($N = 2^{20}$). The bootstrap confidence limits (in solid red) are obtained with the percentile method and $\alpha = 0.05$. The boxplots (bottom row) are obtained on the bootstrap estimates $\hat{c}_p^*$ and show the lower and upper quartile, median, and support of their empirical distributions.

attributes at infinite Reynolds numbers, [75]).

Tab. 12.4 (left columns) shows the results of the Leader and T-B bootstrap based hypothesis tests on $c_2 = 0$. Both data sets unambiguously reject monofractality, with extremely low p-values. The percentile and the studentized tests are in close agreement. This is consistent with results in Tab. 12.3 where confidence intervals for c_2 clearly exclude the zero value. The relevance of these results is further strengthened by the statistical performance of the tests on synthetic processes, as reported and discussed before (cf. Tab 9.2): With only $N = 2^{15}$ samples, Leaders based tests possess sufficient power for rejecting $c_2 = 0$ when $c_{2,A}$ is in the order of magnitude of the $\hat{c}_2$ reported in Tab. 12.3. Hence, this confirms that turbulence data select multifractal multiplicative cascade models as opposed to monofractal ones.

Tab. 12.4 (right columns) shows the results of the Leader boostrap based hypothesis tests on $c_3 = 0$. For both data sets, only a low fraction of the runs rejects the hypothesis $c_3 = 0$. The corresponding p-values remain large, indicating a strong risk of incorrect rejection, if one decides to reject $c_3 = 0$. Percentile and studentized statistics based tests are in good agreement. This is consistent with results in Tab. 12.3 where confidence intervals for c_3 do include the zero value. Note moreover that the estimates for c_3 are in agreement with value zero up to the fourth digit for both data sets, as opposed to results for c_1 and c_2 that slightly differ from one data set to the other. Numerical simulations reported before clearly indicate that non-zero c_3 values can be estimated from data (e.g. Fig. 4.1). Moreover, they show that the tests possess satisfactory power even for small $c_{3,A}$ when only $N = 2^{15}$ samples are available, as opposed to the $N = 2^{20}$ samples of the turbulence data sets used here (cf. Tab. 9.7). Therefore, the results reported in the

Jet turbulence Data			
Model: $c_{2,0} = 0$		Model: $c_{3,0} = 0$	
Reject$_{per}$	98.8	Reject$_{per}$	18.8%
Reject$_{stu}$	98.8	Reject$_{stu}$	15.0%
p-value$_{per}$	0.0049	p-value$_{per}$	0.36
p-value$_{stu}$	0.0052	p-value$_{stu}$	0.37
Wind tunnel Data			
Model: $c_{2,0} = 0$		Model: $c_{3,0} = 0$	
Reject$_{per}$	100	Reject$_{per}$	20.8%
Reject$_{stu}$	100	Reject$_{stu}$	20.8%
p-value$_{per}$	0.005	p-value$_{per}$	0.33
p-value$_{stu}$	0.005	p-value$_{stu}$	0.36

Table 12.4: **Testing monofractality and simple multifractality.** Jet turbulence (top) and wind tunnel turbulence (bottom) wavelet Leader based T-B bootstrap hypothesis tests of monofractality $c_{2,0} = 0$ (left) and of simple multifractality $c_{3,0} = 0$ (right). The significance is $\alpha = 0.05$. The results are averaged over runs.

present contribution are strongly in favor of the conclusion that turbulence c_3 can be considered to be practically zero.

To finish with, hypotheses tests reported in Tab. 12.5 indicate that both data sets strongly reject the She-Lévêque log-Poisson model and that the log-Normal Obukhov-Kolmogorov 62 one is clearly preferred: The c_2 and c_3 values of the former are rejected for almost all runs, for both data sets and both by percentile and studentized bootstrap tests, with very small p-values, whereas the log-Normal model c_2 and c_3 values are rejected for only a small fraction of runs and have large p-values.

Our conclusions — $c_3 = 0$ — are in agreement with results reported in [18], confirming and strengthening them by the analysis of two different major turbulence data sets and by the use of a better mathematically grounded tool (wavelet Leader) and of a statistically more meaningful (bootstrap confidence intervals and hypothesis tests) framework. Similar results have been obtained in [111].

Jet turbulence Data		
Model: LN	$c_{2,0} = -0.0250$	$c_{3,0} = 0.0000$
Reject$_{per}$	42.5%	18.8%
Reject$_{stu}$	40.0%	15.0%
p-value$_{per}$	0.23	0.36
p-value$_{stu}$	0.22	0.37
Model: LP	$c_{2,0} = -0.0365$	$c_{3,0} = 0.0049$
Reject$_{per}$	98.8%	95.0%
Reject$_{stu}$	98.8%	87.5%
p-value$_{per}$	0.005	0.009
p-value$_{stu}$	0.005	0.026

Wind tunnel Data		
Model: LN	$c_{2,0} = -0.0250$	$c_{3,0} = 0.0000$
Reject$_{per}$	16.6%	20.8%
Reject$_{stu}$	8.3%	20.8%
p-value$_{per}$	0.43	0.33
p-value$_{stu}$	0.47	0.36
Model: LP	$c_{2,0} = -0.0365$	$c_{3,0} = 0.0049$
Reject$_{per}$	100%	95.8%
Reject$_{stu}$	100%	87.5%
p-value$_{per}$	0.005	0.034
p-value$_{stu}$	0.005	0.043

Table 12.5: **Testing the log-Normal and the log-Poisson Models.** Jet turbulence (top) and wind tunnel turbulence (bottom) wavelet Leader based T-B bootstrap hypothesis tests of c_2, c_3 of the log-Normal (LN) and the log-Poisson (LP) models. The significance is $\alpha = 0.05$. The results are averaged over runs.

Chapter 13

Application to 2d Signals: Texture Images

Contents

In this chapter, we are interested in *texture classification*. The characterization of textures is now often envisaged by measuring the fluctuations (with respect to space) of the regularity of the amplitude of the image. In an important number of research articles, it is argued that such regularity characterizations should rely on the mathematical framework of multifractal analysis [53, 96, 136, 158, 185]. For instance, Xu et al. [185] propose a texture descriptor, termed the multifractal spectrum vector (MFS) that aims at providing a viewpoint and illumination invariant description of the texture characteristics of images, based on histograms of pixel-per-pixel local estimates of exponents of power laws, for increasing pixel neighborhoods of some measurement functions (one example is the energy of the gradients).
The goal of this chapter is to propose a classification procedure that is based on the multifractal attributes obtained with the tools defined in Chapter 2, and to validate this approach. To this end, we make use of the large database of high resolution images in [185]. It is described in Section 13.1. We obtain results that show that the multifractal analysis tools provide us with attributes that enable the effective discrimination of images of different textures. The performance of the multifractal feature-based classification compares – though not using any specific feature selection or fine tuning – very favorably against those reported in [185].

13.1 High-resolution Texture Image Database

To evaluate the MFS approach, Xu et al. apply it to a large database of high resolution texture images that they set up themselves[1] [185]. It consists of 1000 digital 1280×960 pixel gray level images of 25 different non-traditional textures, such as fruits, plants, floor textures or fabric. Each of these 25 classes contains 40 un-calibrated images of the respective texture, taken from different viewpoints and distances, and for varying illumination conditions. The database has been made available to us by the authors, and will serve us in this chapter as a reference for real-world texture images. For convenience, we refer to the jth image in the ith class by the index (i, j), $i = 1, \cdots, 25$ and $j = 1, \cdots, 40$. Fig. 13.1 shows, for 8 out of the 25 classes of the database, one example image out of the 40 samples per class.

Multifractal attribute estimation: 2 examples. Fig. 13.2 shows wavelet Leader based structure functions and estimates of $\zeta(q)$ and $D(q), h(q)$ for two example images of the database. Estimation parameters are set to: pseudo-fractional integration of order $\alpha = 1$, Daubechies wavelets with $N_\psi = 2$, ordinary (w^0) regressions involving scales $j = [3, 6]$. Bootstrap estimates are obtained with TS-B bootstrap and $R = 99$.
The inspection of structure functions (Fig. 13.2, left) indicates that the images do indeed display scale invariance properties, since they exhibit linear behavior with scale j over a large range of scales. Also, the estimates of scaling exponents and the spectrum (Fig. 13.2, center and right, respectively) suggest that the images are characterized by multifractal properties: The functions $\zeta(q)$ are clearly non-linear in q, and the estimated spectrum has support on a large range of values h.
These results clearly indicate that a scale invariance or multifractal approach is pertinent for the characterization of texture images in this database. This motivates their use for the estimation of features for texture image classification.

[1] Their method is not further detailed here, and its description can be found in [185].

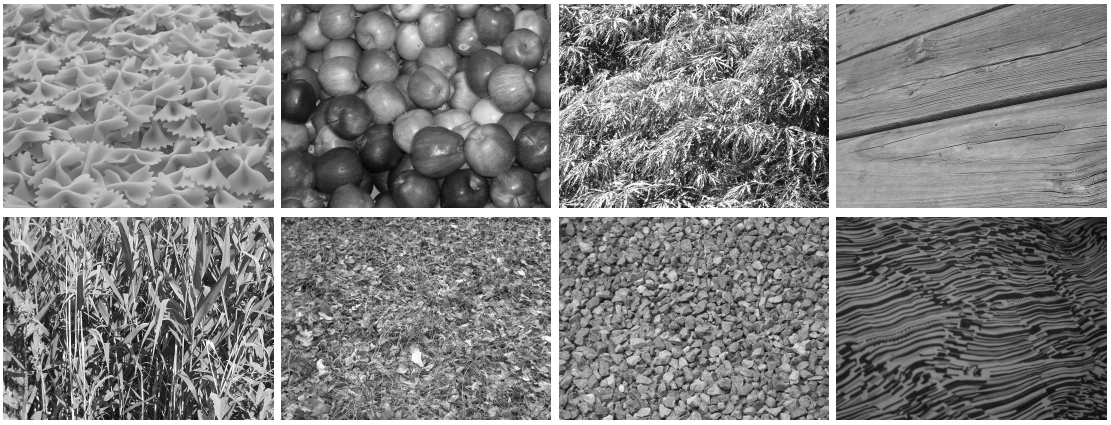


Figure 13.1: **High resolution texture image database.** One example image out of the 40 samples per class for 8 out of the 25 classes of the image database in [185]. From left to right, and from top to bottom: Farfalle, apples, shrubbery, wood deals, grass, fallen leaves, gravel, fabric.

13.2 Multifractal Analysis based Texture Image Classification

In this section, we propose to base texture image classification on feature vectors made up of the multifractal attributes of the images. The procedure is evaluated for the image database in [185] described in the previous section. To allow quantitative comparisons with the results obtained in [185], we use the same k-nearest neighbor (k-NN) classification as described therein.

As mentioned before (cf. Sections 2.5.4 and 5.1.3-b)), the estimates of scaling exponents, log-cumulants and the Legendre spectrum can be fruitfully used as relevant quantities for characterizing data, without explicit interpretation in terms of the Hölder singularities in the data. For automated classification tasks, as considered below, or for batch analyses on large databases, this is natural, since it is practically not feasible to inspect structure functions for each image in the database, and there is no guarantee that linear fits are justified for the a priori fixed range of scales[2]. Therefore, in this classification application, we use the multifractal attribute estimates without explicit reference to the validity of the multifractal formalism.

13.2.1 Feature vectors

The multifractal attribute feature vectors we propose consist of (wavelet Leader or wavelet coefficient based) estimates of $\zeta(q)/q$, $D(q)$, $h(q)$ ($q^L \in \{-4,-3,-2,-1,1,2,3,4\}$, $q^d \in \{1,2,3,4\}$) and c_1, c_2, c_3. Each feature vector $\mathbf{X}_{(i,j)}$ is composed of these multifractal attribute estimates for the image (i,j), obtained for Daubechies wavelets with $N_\psi = 2$ and $N_\psi = 3$ and for the ranges of scales in regressions: $(j_1, j_2) \in \{(1,3),(2,4),(3,5),(4,6)\}$, plus the intercepts of these regression lines at the origin[3].
The multifractal attribute estimates are obtained with pseudo-fractional integration of order $\alpha = 1$. This is justified by the uniform regularity characterization of the database,

[2]This is also due to the absence of methods for the automatic selection of scales over which scaling is observed, hence the necessity to fix the range of scales for linear regressions a priori.

[3]Hence, we consider at total 240 (432) attributes per wavelet coefficient (Leader) based feature vector.

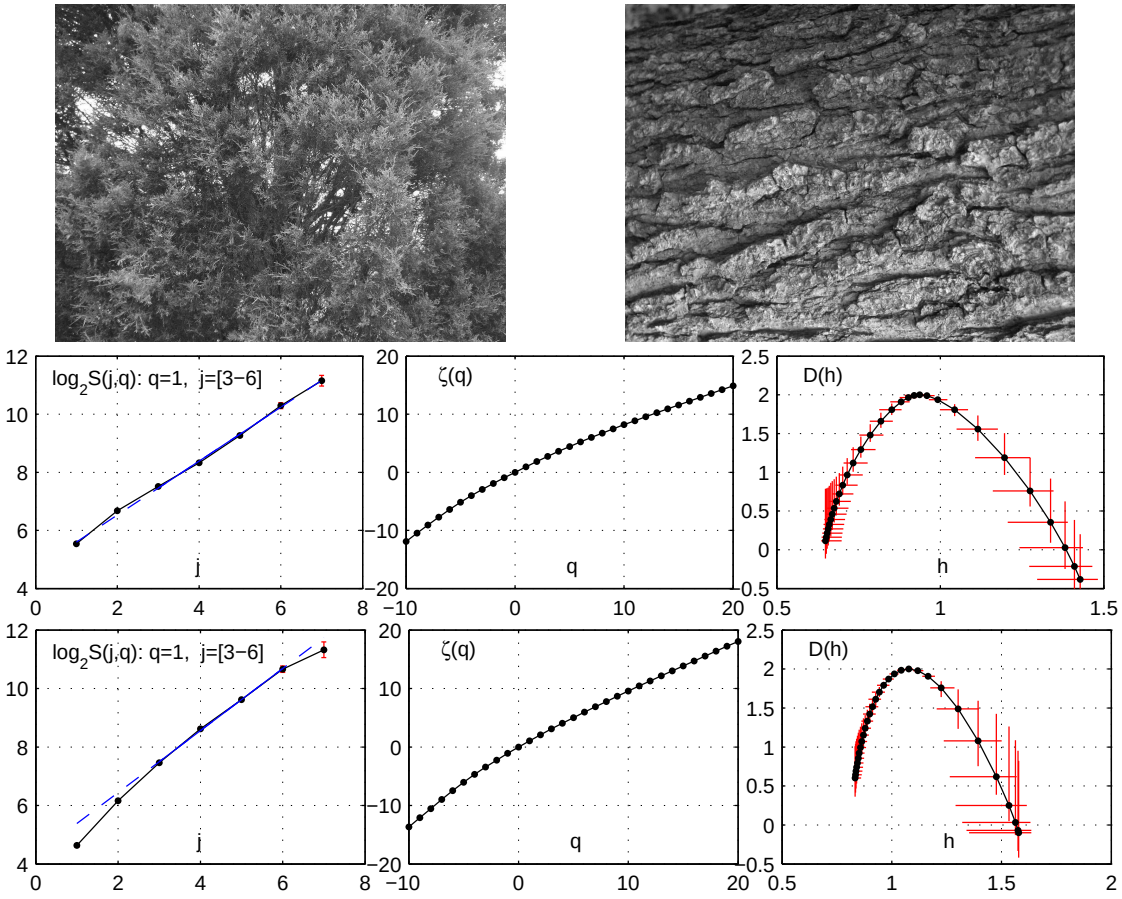


Figure 13.2: **Estimation for example texture images.** Wavelet Leader based estimates of structure functions, scaling exponents and spectra for two examples of images in the database in [185] (shown on top). The first line corresponds to the image shown on the left, the bottom line to the image on the right. Estimations are performed with pseudo-frational integration of order $\alpha = 1$, Daubechies wavelets with $N_\psi = 2$, ordinary (w^0) regressions involving scales $j \in [1, 3]$. Bootstrap estimates are obtained with TS-B bootstrap and $R = 99$, and a confidence level of 90%.

reported in Section 5.1 (Fig. 5.2, left), which shows that for most images in the database, the estimated $\hat{h}_{min} > -1$.

13.2.2 Nearest neighbor classification and performance evaluation

For convenience, we briefly recall the principle of k-NN classification, and state how classification performance is evaluated.
The images in each class i are divided into a set of T training images, which we choose to index by $(m, \tilde{j})$, where $m = i$ indicates the image class of the training images, and $\tilde{j} = 1, \cdots, T$. All the other images, which we call test images, remain to be classified. The k-NN classification of a test image (i, j) is obtained by the following procedure:

1. Calculate for the test image (i, j) the L^1 norm distances $d_{i,j}(m, \tilde{j})$ between its fea-

ture vector $\mathbf{X}_{(i,j)}$ and the feature vectors of all training images $\mathbf{X}_{(m,\tilde{j})}$:

$$d_{i,j}(m,\tilde{j}) = ||\mathbf{X}_{(i,j)} - \mathbf{X}_{(m,\tilde{j})}||_1.$$

2. Sort the distances $d_{i,j}(\cdot,\cdot)$ in increasing order:

$$d_{i,j}(m(1),\tilde{j}(1)) \leq d_{i,j}(m(2),\tilde{j}(2)) \leq d_{i,j}(m(3),\tilde{j}(3)) \leq \cdots,$$

and keep the k training images which are at smallest distance:

$$d_{i,j}(m(1),\tilde{j}(1)) \leq d_{i,j}(m(2),\tilde{j}(2)) \leq \cdots \leq d_{i,j}(m(k),\tilde{j}(k)).$$

3. Determine the class m' to which the majority of these k training images at smallest distance, $(m(n),\tilde{j}(n))$, $n = 1,\cdots,k$, belong.

4. The image class $\hat{i}$ that is assigned to the test image (i,j) is given by this majority class m':

$$\hat{i} = m'.$$

If there is no majority because more than one class are similarly well represented among these k training images, the class $\hat{i}$ whose training images in the k-NN sample have smallest average distance to the test image (i,j) is assigned to the test image (i,j).

Since the number k of nearest neighbors used in [185] is not stated for the results reported therein, we choose to set the number of nearest neighbors arbitrarily to $k = 1$. Hence, the estimated class $\hat{i}$ for the test image (i,j) is the class m of the training image $(m,\tilde{j})$ whose feature vector has smallest distance to the feature vector of the test image (i,j).

If for the test image (i,j) the assigned class $\hat{i} \equiv i$, it is correctly classified. The probability of correct classification for each class is defined as the fraction of test images in this class that are correctly classified. The probability of correct classification for the database is then given by the mean of the probabilities of correct classification of its classes.

Finally, note that probabilities of correct classification naturally depend on the choice of the training samples. Therefore, the estimated probabilities of correct classification are evaluated for randomly selected training samples, and the mean estimated probabilities of correct classifications are calculated over M such random selections.

13.2.3 Results

The mean estimated probabilities of correct classification (for $M = 50$) of the multifractal attribute based classification are reported in Fig. 13.3 (top), as a function of the number of training images T per class. Shown are mean estimated probabilities of correct classification for the best class, i.e., the class with largest mean estimated probability of correct classification among all classes in the database (left), for mean estimated probability of correct classification for all classes (center) and for the worst class (right), and for feature vectors composed of wavelet coefficient (blue circles) and wavelet Leader (red crosses) based multifractal attribute estimates, respectively. As a performance reference, we reproduce in Fig. 13.3 (bottom) the results reported in [185] (these figures are taken from

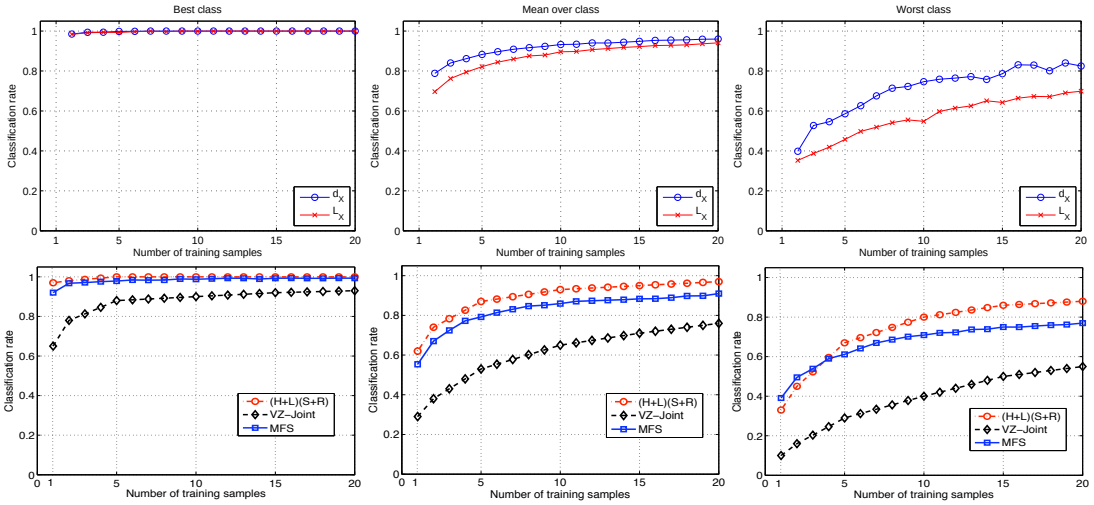


Figure 13.3: **Classification results.** Mean estimated probabilities of correct classification, obtained by k-nearest neighbor classification, for best class (left) all classes (center) and worst class (right) for the texture image database in [185], as a function of the number of training images per class T: Mean estimated probabilities of correct classification for wavelet coefficient (top, 'o') and Leaders (top, '$\times$') based features; Mean estimated probabilities of correct classification for MFS classification ('o' [185]), the (H+L)/(S+R) method in [112] (red circles), and the V-Z method in [167] ('$\diamond$'). The figures on the bottom row are taken from [185].

the article): Xu et al. chose to compare their MFS based classification (blue squares) to the (H+L)/(S+R) method proposed in [112] (red circles), and the V-Z method proposed in [167] (black diamonds). For details on these last two methods, we refer the reader to the corresponding references. The results indicate that classification based on multifractal attributes compares very favorably against the other methods proposed in [185] for this texture image database: Mean estimated probabilities of correct classification for d_X are similar or above those of the (H+L)/(S+R) method, for which [185] reports the best results among the three methods therein. Mean estimated probabilities of correct classification for L_X are slightly below those of d_X, but remain similar or better than those of the MFS method, and better than those obtained by the V-Z method.

13.3 Conclusions and Perspectives

The results reported above lead us to the conclusion that multifractal attribute estimates, as obtained by the methods defined in Section 2.6, are highly relevant for the characterization of texture images. Used as features for classification, they give rise to effective image classification schemes whose peformance compare very favorably against those of schemes previously proposed in the literature. What is more, the results have been obtained here without any specific feature selection or fine tuning, such as principal components analysis (PCA). The feature vectors may therefore contain redundant attributes, and results not reported here show that this is indeed the case: For instance, using estimates obtained for one wavelet (N_ψ) only does not significantly alter classification results. Also, some features may be numerically dominated by others and hence ineffec-

tive, though potentially discriminative. For example, the log-cumulants c_p of order $p \geq 2$ usually take on values that are relatively close to zero, as compared to other attributes, such as $D(q \approx 0)$. Furthermore, no optimization of the number k of nearest neighbors has been performed, which has been a priori and arbitrarily set to $k = 1$. Such issues demand further investigation and represent a large potential for further improving texture image probability of correct classifications for the proposed method.
The classification approach has reasonably low computational cost, and the processing of the 1000 high-resolution images takes approximately 5 hours on a standard PC. We note that computation time can be effectively halved by using feature vectors consisting of estimates for only one wavelet (N_ψ). Then, calculation of feature vectors takes roughly 10 seconds only per image.
Furthermore, the database made available to the authors by Xu et al. consists actually of 2000 images (50 classes instead of 25), of which only one half (25 classes) are used here and in [185]. We also processed this larger database. Results are not reported here, since we lack a reference for classification performance. Along the same line, preliminary analysis of other texture image databases have been performed.
Finally, it would be interesting to validate whether multifractal analysis enables to perform image classification at a higher semantic level, commonly referred to as scene recognition (i.e., discrimination of e.g. images of houses, landscapes, etc.). This is of importance in a large number of applications in, for instance, automatized image retrieval and computer vision. Such a procedure may be considered by combining the feature vectors proposed here with the image function space and uniform regularity estimates $\hat{\zeta}^d(q)$ and $\hat{h}_{min}$ defined in Section 5.1.

We would like to express our gratitude to Xu et al. – and in particular to Hui Ji – for making their database available to us, and for fruitful discussions.

Chapter 14

Conclusions and Perspectives

A short while before this work was undertaken, a new multifractal analysis tool had been introduced in the mathematical literature related to the theory of multifractal analysis by S. Jaffard [89, 92]: The wavelet Leaders. It had been proven to bring significant improvements in the precise analysis of the regularity of functions, hence mostly from a functional analysis perspective. Yet, at the outset of this thesis, most of the practical and statistical properties of this wavelet Leader based multifractal formalism remained poorly known. Now that we arrive at the end of this manuscript, let us take a moment to conclude on what results have been obtained in the present thesis work.

The detailed and systematic study of the behavior, practical use and statistical performance of this multifractal analysis tool constitutes one of the central contributions of this work. A first 1d implementation had been previously made available, and the practical effectiveness of the functional analysis theoretical results had been demonstrated. At present, we dispose of operational well controlled implementations of wavelet Leader based multifractal analysis procedures for both (1d) signal and (2d) image. Their statistical performance have been systematically investigated and are well understood and documented. Notably, they possess significantly better performance than previous formulations based either on discrete or on continuous wavelet transforms. This enhancement is particularly significant for the analysis and estimation of the scaling attributes that convey the key differences between self-similar processes and multiplicative cascades, i.e., between the two major classes of stochastic processes used to model scale invariance in applications. Therefore, this leads us to conclude that while wavelet coefficients perform satisfactorily in estimation problems for self-similar processes, wavelet Leaders are needed for a relevant analysis of multiplicative cascades and hence indispensable for an actual practical discrimination between these two categories of processes. Also, it has been shown here that the estimation of *advanced* multifractal attributes (e.g., the third log cumulant c_3) is feasible. Such conclusions had never been arrived at and have been validated clearly and unambiguously. However, the use of wavelet Leaders comes with restrictions: It can be applied directly only to a subclass of functions, namely, that of functions characterized only with positive Hölder exponents. Yet, for many real-life data, for both images and signals, these conditions are not necessarily fulfilled a priori. As a matter of fact, it is an important challenge to decide a priori whether real data fall in this class or not, as a direct application of wavelet Leaders to data outside this class yields results whose irrelevance may turn difficult to detect a posteriori for practitioners. Another central contribution of the present work consists of a clear formulation of this difficulty together with a detailed and comprehensive study of the solution, proposed earlier for other multifractal analysis tools, here tailored to wavelet Leaders and relying on (pseudo)-fractional integration. Also, these investigations lead us to conclude that wavelet coefficient based analyses should be conducted prior to wavelet Leader based one, and in a complementary manner, rather than in the usually envisioned competition perspective. An operational implementation of these completive analysis tools and of the corresponding modified wavelet Leader multifractal analysis tool is now incorporated in our standard toolbox.

At the time we started with this work, no statistical procedures for assessing confidence in estimates and for performing tests was available in multifractal analysis. In this thesis, we studied the potential of bootstrap to contribute to solutions for these issues. We have proposed two bootstrap approaches and considered them for the formulation of confidence intervals, hypothesis tests and time constancy tests of multifractal attributes. The

originality of our contribution lies in performing bootstrap directly in the time-scale plane, over blocks of wavelet coefficients or Leaders, with no attempt (nor necessity) to reconstruct a time domain signal (hence avoiding the burden of practically inverting the discrete wavelet transform). Also, blocks are shaped to match the intrinsically time-scale nature of the dependence structure of any multiresolution quantity. Operational implementations of the wavelet Leaders and bootstrap based multifractal analysis procedures are hence available for both (1d) signals and (2d) images. Numerical simulation results demonstrate that these bootstrap procedures for multifractal attributes have satisfactory performance. The procedures can therefore be practically and effectively used for statistical inference in multifractal analysis. It has been shown here that the wavelet Leaders conjoint and enhanced with bootstrap analyses enable – for the first time – an accurate and sharp practical multifractal characterization and discrimination of single observations of real data. Beyond, the theoretical analyses we have conducted on the dependence structure of multiresolution quantities for multifractal multiplicative cascades lead us to question the validity of bootstrap if applied independently at each scale. This opens room for further developments and research aiming at understanding why bootstrap procedures, when used in a multiresolution manner, i.e., combining scales, yield such satisfactory practical performance.
In addition, these theoretical analyses significantly renewed the understanding of the interplay between the *number of vanishing moments* (involved in the multiresolution analysis of data) and the dependence structure of multiresolution quantities (increments, wavelet coefficients, wavelet Leaders): Increasing the number of vanishing moments does decrease correlation range, but does not in general decrease dependence (or higher order correlation) range, as opposed to what happens for the specific Gaussian fractional Brownian motions. These conclusions have never been unequivocally validated before and overturn common comprehension respective to the specific role that multiresolution quantities play for scaling processes.

The main hurdle for advancing the theoretical understanding of estimation in multifractal analysis is related to the fact that multifractal processes constitute a large class, comprising processes with a rich variety of statistical properties that remain poorly studied. Therefore, we believe that solutions toward progress will need to concentrate on particular case studies for specific multifractal processes, backed up by numerical simulations, rather than aiming at the class of multifractal processes in its whole generality. In this spirit, we have conducted case studies of the linearization effect or of the dependence structures of increments for the particular example of compound Poisson motions. The results obtained on such specific cases are then expected to promote further insights on the properties of much larger classes of multifractal processes and can hence be used to formulate the core of conjectures that are assumed to hold within larger classes of processes. Numerical simulations can then be fruitfully used to assess the generality of these conjectures, by scanning systematically various classes of processes. In turn, numerical simulations often provide material for further theoretical analyses. We consider that numerical simulations and theoretical case studies are also to be envisioned in a complementary perspective rather than in different ways of conducting researches that exclude each other. These aspects make up for the main difficulties in this domain, and they also explain why multifractal analysis is such an extremely interesting and rich playground.

Though not constituting the core of the material presented in this manuscript, applica-

tions remain a central concern of the present work. In Chapters 12 and 13, we have illustrated that the proposed wavelet Leader and bootstrap based multifractal analysis procedures are effective for applications of very different natures. At present, we dispose of a wavelet Leader based multifractal analysis tool which is well documented in terms of performance and usage. It can be readily and systematically applied, including in situations where knowledge and control of the properties and limitations of the procedures are primary preconditions, such as in robotics or in automation. Also, it can by now be involved in applications where fundamental concern lies in inference from the data under analysis, as is the case in, for instance, most medical applications. What is more, the procedures can now be readily used for the characterization of (2d) images for detection, identification or classification tasks. Notable, voluminous databases of large size images can be efficiently scanned, given the very low computational load of the proposed tools. Therefore, multifractal analysis constitutes by now a novel tool for Image processing. The exploration of this potential is just about to start and promises exciting new perspectives.

Finally, this work lead to the development, implementation and validation of a number of tools for wavelet Leader and bootstrap based multifractal analysis of signals and images. They form the core of a MATLAB© based toolbox that will be published on the web page of the author.

Appendix A

A log cumulant expansion of the multifractal spectrum

We will show that the spectrum allows an expansion of the type $D(h) = f_{c_p}(h)$ around its maximum, where $f_{c_p}(h)$ is a function in h, parametrized by the log-cumulants c_p. For the first orders, the log-cumulants c_p will have a meaningful interpretation in this expansion. The starting point is the polynomial expansion[1] of $\zeta(q)$ Eq. (2.70) with log-cumulants c_p, for $q \simeq 0$ and on condition that $\mathbb{E}L_X^q < \infty$:

$$\zeta(q) = \sum_{p\geq 1} c_p \frac{q^p}{p!}, \tag{A.1}$$

and the expression of $D(h)$ given by the multifractal formalism Eq. (2.61):

$$D(h) = \min_{q\neq 0}(d + qh - \zeta(q)), \tag{A.2}$$

where d is the dimension in which the data live. Deriving Eq. (A.2) w.r.t. q and solving for the minimum gives the parametric forms:

$$h(q) = \zeta'(q) = \sum_{p\geq 1} c_p \frac{q^{(p-1)}}{(p-1)!} = c_1 + \sum_{p\geq 2} c_p \frac{q^{(p-1)}}{(p-1)!} \tag{A.3}$$

$$D(h(q)) = d + q\zeta'(q) - \zeta(q) = d + \sum_{p\geq 2} c_p (p-1)\frac{q^p}{p!} = d + \sum_{p\geq 2} c_p \frac{q^p}{(p-2)!\cdot p}. \tag{A.4}$$

With a slight abuse of notation, we will write $D(q)$ instead of $D(h(q))$ for convenience.

Expressions of the derivatives of $D(q)$ and $h(q)$. In order to find a development of $D(h)$ in h around its maximum, we need the successive derivatives in q of $D(q)$ and $h(q)$:

$$\frac{d^{(n)}D}{dq^{(n)}} = D^{(n)}(q) = \sum_{p\geq n} c_p (p-1)\frac{q^{(p-n)}}{(p-n)!}, \quad n \geq 1,$$

$$\frac{d^{(n)}h}{dq^{(n)}} = h^{(n)}(q) = \sum_{p\geq n+1} c_p \frac{q^{(p-n-1)}}{(p-n-1)!} = c_{n+1} + \sum_{p\geq n+2} c_p \frac{q^{(p-1)}}{(p-1)!}, \quad n \geq 0.$$

[1] Written either for wavelet Leaders, or for other multiresolution quantities such as increments, continuous wavelet coefficients or discrete wavelet coefficients as in [44], [60] or e.g. [180], respectively. For convenience, any superscript ^{L}, ^{d} or ^{I} is dropped since irrelevant for the derivation.

Maximum of $D(h)$. Let us assume that the maximum of $D(h)$ is d: $\max D(h) \equiv d$. Since $D(h)$ is a concave function, it only has one maximum. From Eq. (A.4), it is obvious that the maximum of $D(q)$ is taken at $q = 0$. Eq. (A.3) gives the corresponding value for h: $h(q = 0) = c_1$. Therefore, the maximum of $D(h) = D(h(q))$ is taken at $h = c_1$.

Development of $D(h)$ around its maximum at $h = c_1$. Noting that $D(h) = D(h(q))$, and denoting derivatives with respect to q by a $'$, and derivatives with respect to h by a *, we have:

$$\begin{aligned}
\frac{dD}{dq} &= \frac{dD}{dh}\frac{dh}{dq} = D^{*(1)}h'^{(1)} \\
\frac{d^2D}{dq^2} &= \frac{d}{dq}\left[\frac{dD}{dh}\frac{dh}{dq}\right] = \frac{dD}{dh}\frac{d^2h}{dq^2} + \left(\frac{dh}{dq}\right)^2\frac{d^2D}{dh^2} = D^{*(1)}h'^{(2)} + D^{*(2)}\left(h'^{(1)}\right)^2 \\
\frac{d^3D}{dq^3} &= \frac{d}{dq}\left[\frac{d^2D}{dq^2}\right] = \cdots = D^{*(1)}h'^{(3)} + 3h'^{(2)}h'^{(1)}D^{*(2)} + D^{*(3)}\left(h'^{(1)}\right)^3 \\
\frac{d^4D}{dq^4} &= \frac{d}{dq}\left[\frac{d^3D}{dq^3}\right] = \cdots \\
&= D^{*(1)}h'^{(4)} + \left[4h'^{(3)}h'^{(1)} + 3h'^{(2)}h'^{(2)}\right]D^{*(2)} + 6h'^{(2)}\left(h'^{(1)}\right)^2 D^{*(3)} + D^{*(4)}\left(h'^{(1)}\right)^4.
\end{aligned}$$

Substituting the values of the derivatives $h'^{(n)}$ and $D'^{(n)}$ at the maximum of the spectrum, i.e. at $q = 0$, in these equations:

$$\begin{aligned}
h(q=0) &= c_1 \\
h'^{(n)}(q=0) &= c_{n+1}, \quad n \geq 0 \\
D(q=0) &= d \\
D'^{(1)}(q=0) &= 0 \\
D'^{(n)}(q=0) &= (n-1)c_n, \quad n \geq 2
\end{aligned}$$

we can solve for the coefficients $D^{*(p)} = \frac{d^pD}{dh^p}\big|_{h=c_1}$ in a development of $D(h)$ around c_1,

$$D(h) = D(c_1) + \sum_{p\geq 1}\frac{(h-c_1)^p}{p!}\left.\frac{d^pD}{dh^p}\right|_{h=c_1}:$$

$$\begin{aligned}
&D'^{(1)}(0) = D^{*(1)}\cdot h'^{(1)} = 0\cdot c_2 = 0 && \Rightarrow \quad D^{*(1)} = 0 \\
&D'^{(2)}(0) = D^{*(1)}h'^{(2)} + D^{*(2)}\left(h'^{(1)}\right)^2 = 0\cdot c_2 + D^{*(2)}c_2^2 = c_2 && \Rightarrow \quad D^{*(2)} = \frac{1}{c_2} \\
&D'^{(3)}(0) = 0\cdot c_4 + 3c_3c_2\tfrac{1}{c_2} + D^{*(3)}c_2^3 = 2c_3 && \Rightarrow \quad D^{*(3)} = -\frac{c_3}{c_2^3} \\
&D'^{(4)}(0) = 0 + \tfrac{1}{c_2}[4c_4c_2 + 3c_3^2] - 6c_3c_2^2\tfrac{c_3}{c_2^3} + c_2^4D^{*(4)} = 3c_4 && \Rightarrow \quad D^{*(4)} = -\frac{c_4}{c_2^4} + \frac{3c_3^2}{c_2^5}.
\end{aligned}$$

The polynomial expansion of the spectrum around its maximum is therefore given by:

$$D(h) = d + \frac{c_2}{2!}\left(\frac{h-c_1}{c_2}\right)^2 + \frac{-c_3}{3!}\left(\frac{h-c_1}{c_2}\right)^3 + \frac{-c_4+3c_3^2/c_2}{4!}\left(\frac{h-c_1}{c_2}\right)^4 + \ldots \qquad \text{(A.5)}$$

Appendix B

Direct determination of the multifractal spectrum

Chhabra et al. [49] proposed a method for direct determination of the multifractal spectrum without explicit resort to the Legendre transform. We will give here a detailed derivation of this useful estimation procedure. It is based on Eq. (2.66) and the multifractal formalism Eq. (2.61):

$$\mathbb{E}L(j,k)^q = F_q(2^j)^{\zeta^L(q)}, \tag{B.1}$$

$$D(h) = \min_{q\neq 0}(d + qh - \zeta^L(q)), \tag{B.2}$$

where d is the dimension in which the data lives. Derivation of Eq. (B.2) with respect to q, and solving for the minimum of the right hand side gives:

$$h(q) = \frac{d}{dq}\zeta(q) = \zeta'(q), \tag{B.3}$$

$$D(q) = d + q\zeta'(q) - \zeta(q). \tag{B.4}$$

Solving for $h(q) = \zeta'(q)$. Taking the $\log_2$ in Eq. (B.1) gives:

$$\log_2 \mathbb{E}L(j,k)^q = \kappa_q + j\zeta(q), \quad \kappa_q = \log_2 F_q. \tag{B.5}$$

After reordering and taking the derivative with respect to q we have:

$$j\zeta'(q) = -\kappa'_q + (\log_2 \mathbb{E}L(j,k)^q)' \tag{B.6}$$

$$= \frac{1}{\ln 2}(\ln \mathbb{E}L(j,k)^q)' - \kappa'_q \tag{B.7}$$

$$= \frac{1}{\ln 2}\frac{(\mathbb{E}L(j,k)^q)'}{\mathbb{E}L(j,k)^q} - \kappa'_q \tag{B.8}$$

$$= \frac{1}{\ln 2}\frac{\mathbb{E}L(j,k)^q \ln L(j,k)}{\mathbb{E}L(j,k)^q} - \kappa'_q \tag{B.9}$$

$$= \frac{\mathbb{E}L(j,k)^q \log_2 L(j,k)}{\mathbb{E}L(j,k)^q} - \kappa'_q. \tag{B.10}$$

Finally, by replacing the expectations with sample means, we obtain the estimator Eq. (2.75) for $h(q) = \zeta'(q)$, given by the linear fit of Eq. B.10 vs. j:

$$\hat{h}(q) = \sum_{j=j_1}^{j_2} w_j \frac{\frac{1}{n_j}\sum_{k=1}^{n_j} L(j,k)^q \log_2 L(j,k)}{\frac{1}{n_j}\sum_{k=1}^{n_j} L(j,k)^q} - \underbrace{\sum_{j=j_1}^{j_2} w_j \kappa'_q}_{=0} \tag{B.11}$$

$$= \sum_{j=j_1}^{j_2} w_j \frac{\sum_{k=1}^{n_j} L(j,k)^q \log_2 L(j,k)}{\sum_{k=1}^{n_j} L(j,k)^q} \tag{B.12}$$

$$= \sum_{j=j_1}^{j_2} w_j \underbrace{\sum_{k=1}^{n_j} R^q(j,k) \log_2 L(j,k)}_{V(j,q)}, \quad R^q(j,k) = \frac{L(j,k)^q}{\sum_{k=1}^{n_j} L(j,k)^q} \tag{B.13}$$

$$= \sum_{j=j_1}^{j_2} w_j V(j,q). \tag{B.14}$$

Solving for $D(q) = 1 + q\zeta'(q) - \zeta(q)$. From Eqs. (B.5) and (B.10), we have:

$$j\zeta(q) = -\kappa_q + \log_2 \mathbb{E}L(j,k)^q, \tag{B.15}$$

$$jq\zeta'(q) = q\frac{\mathbb{E}L(j,k)^q \log_2 L(j,k)}{\mathbb{E}L(j,k)^q} \tag{B.16}$$

$$= \frac{\mathbb{E}L(j,k)^q \log_2 L(j,k)^q}{\mathbb{E}L(j,k)^q}. \tag{B.17}$$

Combining these with Eq. (B.4), we have:

$$j(D(q) - d) = \kappa_q + \frac{\mathbb{E}L(j,k)^q \log_2 L(j,k)^q}{\mathbb{E}L(j,k)^q} - \log_2 \mathbb{E}L(j,k)^q. \tag{B.18}$$

Finally, by replacing the expectations with sample means, the last equation gives the estimator for $D(q)$ Eq. (2.74) as the linear fit of Eq. (B.18) versus j:

$$\begin{aligned}
\hat{D}(q) &= d + \sum_{j=j_1}^{j_2} w_j \frac{\frac{1}{n_j}\sum_{k=1}^{n_j} L(j,k)^q \log_2 L(j,k)^q}{\frac{1}{n_j}\sum_{k=1}^{n_j} L(j,k)^q} - \log_2 \frac{1}{n_j}\sum_{k=1}^{n_j} L(j,k)^q \\
&= d + \sum_{j=j_1}^{j_2} w_j \frac{\frac{1}{n_j}\sum_{k=1}^{n_j} L(j,k)^q \log_2 L(j,k)^q - \frac{1}{n_j}(\sum_{k=1}^{n_j} L(j,k)^q)(\log_2 \sum_{k=1}^{n_j} L(j,k)^q)}{\frac{1}{n_j}\sum_{k=1}^{n_j} L(j,k)^q} + \\
&\quad + \log_2 n_j \\
&= d + \sum_{j=j_1}^{j_2} w_j \frac{\sum_{k=1}^{n_j} L(j,k)^q \log_2 L(j,k)^q - L(j,k)^q \log_2 \sum_{k=1}^{n_j} L(j,k)^q}{\sum_{k=1}^{n_j} L(j,k)^q} + \log_2 n_j \\
&= d + \sum_{j=j_1}^{j_2} w_j \sum_{k=1}^{n_j} R^q(j,k) \log_2 R^q(j,k) + \log_2 n_j \\
&= d + \sum_{j=j_1}^{j_2} w_j U(j,q) + \log_2 n_j.
\end{aligned}$$

Appendix C

Proofs of Theorem 11.1 and Theorem 11.2

Let us, for convenience of notation, write $A_\tau^q(t) = A_\tau^{(1)}(t)^q$ for the increments of order one.

The proofs of these results rely on a lemma of Bacry and Muzy for IDC motion [23]: Let us put $\varphi(\cdot) = \psi(-i\cdot)$ and let ω_r be defined by $Q_r = e^{\omega_r}$. For $t, t' \geq 0$, we define

$$\mathcal{C}_r(t, t') = \mathcal{C}_r(t) \cap \mathcal{C}_r(t'),$$

where $\mathcal{C}_r(t)$ is the cone in Eq. (2.100).

Lemma C.1 *Let $q \in \mathbb{N}^*$, $\overrightarrow{t_q} = (t_1, t_2, ..., t_q)$ with $t_1 \leq t_2 \leq ... \leq t_q$ and $\overrightarrow{p_q} = (p_1, p_2, ..., p_q)$. The characteristic function of the vector $\{w_r(t_m)\}_{1 \leq m \leq q}$ is*

$$\mathbb{E}\left(e^{\sum_{m=1}^q ip_m P(C_r(t_m))}\right) = e^{\sum_{j=1}^q \sum_{k=1}^j \alpha(j,k)\rho_r(t_k - t_j)}$$

where

$$\rho_r(t) = m(\mathcal{C}_r(0, t)),$$

and

$$\alpha(j, k) = \psi(r_{k,j}) + \psi(r_{k+1,j-1}) - \psi(r_{k,j-1}) - \psi(r_{k+1,j})$$

and

$$r_{k,j} = \begin{cases} \sum_{m=k}^j p_m, & \text{for } k \leq j, \\ 0 & \text{for } k > j. \end{cases}$$

Moreover

$$\sum_{j=1}^q \sum_{k=1}^j \alpha(j, k) = \psi\left(\sum_{k=1}^q p_k\right).$$

This result can be rewritten as

$$\mathbb{E}Q_r^{p_1}(t_1)Q_r^{p_2}(t_2)...Q_m^{p_m}(t_m) = e^{\sum_{j=1}^q \sum_{k=1}^j \beta(j,k)\rho_r(t_k - t_j)}$$

where

$$\beta(j,k) = \varphi(r_{k,j}) + \varphi(r_{k+1,j-1}) - \varphi(r_{k,j-1}) - \varphi(r_{k+1,j})$$

and

$$\sum_{j=1}^{q}\sum_{k=1}^{j}\beta(j,k) = \varphi\left(\sum_{k=1}^{q} p_k\right).$$

C.1 Proof of Theorem 11.1

Since $\mathbb{E}{A_\tau}^{2q} < \infty$, one has $\mathbb{E}{A_\tau}^q(t){A_\tau}^q(s) < \infty$ and, using the monotone convergence theorem for the 4th equality:

$$\begin{aligned}
\mathbb{E}{A_\tau}^q(t){A_\tau}^q(s) &= \mathbb{E}\left(\lim_{r_1\to 0}\int_t^{t+\tau} Q_{r_1}(x)dx\right)^q\left(\lim_{r_2\to 0}\int_s^{s+\tau} Q_{r_2}(y)dy\right)^q \\
&= \mathbb{E}\prod_{i=1}^{q}\lim_{r_{1,i}\to 0}\lim_{r_{2,i}\to 0}\int_t^{t+\tau} Q_{r_{1,i}}(x_i)dx_i\int_s^{s+\tau} Q_{r_{2,i}}(y_j)dy_j \\
&= \mathbb{E}\lim_{r\to 0}\prod_{i=1}^{q}\int_t^{t+\tau} Q_r(x_i)dx_i\int_s^{s+\tau} Q_r(y_i)dy_i \\
&= \lim_{r\to 0}\mathbb{E}\prod_{i=1}^{q}\int_t^{t+\tau} Q_r(x_i)dx_i\int_s^{s+\tau} Q_r(y_i)dy_i \\
&= \lim_{r\to 0}\int_{[t,t+\tau]^q}\int_{[s,s+\tau]^q}\mathbb{E}\prod_{i=1}^{q} Q_r(x_i)Q_r(y_i)d(x_1,...,x_q)d(y_1,...,y_q)
\end{aligned}$$

By symmetry, it comes:

$$\mathbb{E}{A_\tau}^q(t){A_\tau}^q(s) = (q!)^2\lim_{r\to 0}\int_{D_1}\int_{D_2}\mathbb{E}\prod_{i=1}^{q} Q_r(x_i)Q_r(y_i)d(x_1,...,x_q)d(y_1,...,y_q).$$

where $D_1 = \{t \le x_1 \le x_2, ... \le x_q \le t+\tau\}$ and $D_2 = \{s \le y_1 \le y_2 \le ... \le y_q \le s+\tau\}$.

Let us now fix $r < s - t - \tau$. We put $\Delta_r^+ = \{(t,z) \in \mathbb{R}^2; z > r\}$. We want to compute $I = \mathbb{E}\prod_{i=1}^{q} Q_r(x_i)Q_r(y_i)$ for $t \le x_1 \le x_2, ... \le x_q \le t+\tau$ and $s \le y_1 \le y_2 \le ... \le y_q \le s+\tau$. Using Lemma C.1, it comes:

$$I = e^{\sum_{j=1}^{q}\sum_{k=1}^{j}\beta(j,k)\rho_r(x_k - x_j)}e^{\sum_{j=1}^{q}\sum_{k=1}^{j}\beta(j+q,k+q)\rho_r(y_k - y_j)}e^{\sum_{j=1}^{q}\sum_{k=1}^{q}\beta(j,k+q)\rho_r(y_k - x_j)}. \tag{C.1}$$

Doing this, we have written I as the product of 3 terms. The estimation of the term with $\rho_r(y_k - x_j,)$ will give us the behaviour in $|t-s|^{\varphi(2q)-2\varphi(q)}$ while the computation of the other terms – which do not depend on $|t-s|$ – will give us the factor $|\tau|^{2(q+\varphi(q))}$.

More precisely, let us put:

$$J = e^{\sum_{j=1}^{q}\sum_{k=1}^{q}\beta(j+q,k)\rho_r(y_j - x_k)}$$

Since $r < s - t - \tau$, $\rho_r(y_k - x_j) = -\ln|y_k - x_j|$ with $s - t - \tau \leq |y_k - x_j| \leq s - t + \tau$. Besides,

$$\sum_{j=1}^{q}\sum_{k=1}^{q}\beta(j+q,k) = \sum_{j=1}^{2q}\sum_{k=1}^{j}\beta(j,k) - \sum_{j=1}^{q}\sum_{k=1}^{j}\beta(j,k) - \sum_{j=1}^{q}\sum_{k=1}^{j}\beta(j+q,k+q).$$

But, here, $p_m = 1$, $r_{k,j} = j - k + 1$ and $\beta(j,k)$ depends only on $k - j$. Finally, we obtain:

$$\sum_{j=1}^{q}\sum_{k=1}^{q}\beta(j+q,k) = \varphi(2q) - 2\varphi(q)$$

and

$$|s - t + \tau|^{\varphi(2q)-2\varphi(q)} \leq J \leq |s - t - \tau|^{\varphi(2q)-2\varphi(q)}.$$

Hence,

$$|s - t + \tau|^{\varphi(2q)-2\varphi(q)} \lim_{r\to 0} L(r,\tau) \leq \mathbb{E}{A_\tau}^q(t){A_\tau}^q(s) \leq |s - t - \tau|^{\varphi(2q)-2\varphi(q)} \lim_{r\to 0} L(r,\tau)$$

with:

$$L(r,\tau) = (q!)^2 \int_{D_1}\int_{D_2} e^{\sum_{j=1}^{q}\sum_{k=1}^{j}\beta(j,k)\rho_r(x_k - x_j)} e^{\sum_{j=1}^{q}\sum_{k=1}^{j}\beta(j+q,k+q)\rho_r(y_j - y_k)}.$$

Note that $\lim_{r\to 0} L(r,\tau)$ is well defined, since $L(\cdot,\tau)$ is an increasing function of r bounded by $\frac{\mathbb{E}{A_\tau}^q(t){A_\tau}^q(s)}{|s-t-\tau|^{c(\varphi(2q)-2\varphi(q))}(q!)^2}$.

Lemma C.2

$$\lim_{r\to 0} L(r,\tau) = (\mathbb{E}{A_\tau}^q(t))^2 = C(q)|\tau|^{2(q+\varphi(q))}.$$

PROOF. It is known that $\mathbb{E}{A_\tau}^q(t) = \tilde{C}(q)|\tau|^{q+\varphi(q)}$. But, again, we can write

$$\begin{aligned}\mathbb{E}{A_\tau}^q(t) &= \lim_{r\to 0}\int_{t\leq x_1,\ldots,x_q} \mathbb{E}\prod_{i=1}^{q} Q_r(x_i)d(x_1,\ldots,x_q)\\ &= (q!)\lim_{r\to 0}\int_{D_1} \mathbb{E}\prod_{i=1}^{q} Q_r(x_i)d(x_1,\ldots,x_q)\end{aligned}$$

which gives $\mathbb{E}{A_\tau}^q(t)\mathbb{E}{A_\tau}^q(s) = \lim_{r\to 0} L(r,\tau)$.

Replacing $\lim_{r\to 0} L(r,\tau)$ with $C(q)|\tau|^{2(q+\varphi(q))}$, we obtain:

$$|s - t + \tau|^{\varphi(2q)-2\varphi(q)} \leq \frac{{A_\tau}^q(t){A_\tau}^q(s)}{C(q)|\tau|^{2(q+\varphi(q))}} \leq |s - t - \tau|^{\varphi(2q)-2\varphi(q)}.$$

C.2 Proof of Theorem 11.2

We consider the case $q > 1$, $q \notin \mathbb{N}$. Let us put $q = m + \varepsilon$ with $m = [q]$ and $0 < \varepsilon < 1$. One can write:

$$\mathbb{E}{A_\tau}^q(s){A_\tau}^q(t) \quad = \quad \mathbb{E}{A_\tau}^{m-1}(s){A_\tau}^{m-1}(t){A_\tau}^{1+\varepsilon}(s){A_\tau}^{1+\varepsilon}(t).$$

Again ${A_\tau}^{m-1}(s){A_\tau}^{m-1}(t)$ can be written as a multiple integral. Also, a classical Hölder inequality yields to:

$${A_\tau}^{1+\varepsilon}(t) = \left(\int_t^{t+\tau} Q_r(x)dx\right)^{1+\varepsilon} \leq \tau^\varepsilon \int_t^{t+\tau} Q_r^{1+\varepsilon}(x)dx.$$

Hence, one gets:

$$\mathbb{E}{A_\tau}^q(s){A_\tau}^q(t) \leq \lim_{r\to 0} \tau^{2\epsilon} \int_D E \prod_{i=1}^{m-1} Q_r(x_i)Q_r(y_i)Q_r(x_m)^{1+\varepsilon}Q_r(y_m)^{1+\varepsilon}d(x_1,...,x_m)d(y_1,...,y_m) \tag{C.2}$$

where $D = [t, t+\tau]^m \times [s, s+\tau]^m$. From Lemma C.1, one can write again:

$$\mathbb{E}\prod_{i=1}^{m-1} Q_r(x_i)Q_r(y_i)Q_r(x_m)^{1+\varepsilon}Q_r(y_m)^{1+\varepsilon}$$

as the product of three terms. The term:

$$J = e^{\sum_{j=1}^q \sum_{k=1}^q \beta(j+q,k)\rho_r(y_j - x_k)}$$

is majorated by $|t-s+\tau|^{\varphi(2q)-2\varphi(q)}$ and the integral on D of the other ones by $\tau^{2\varphi(q)+2m}$. Finally, one gets:

$$\mathbb{E}{A_\tau}^q(s){A_\tau}^q(t) \leq \tau^{2\varepsilon}\tau^{2\varphi(q)+2m}|t-s+\tau|^{\varphi(2q)-2\varphi(q)} \leq \tau^{2q+2\varphi(2q)}|t-s+\tau|^{\varphi(2q)-2\varphi(q)}.$$

To obtain the minoration, we write:

$$\mathbb{E}{A_\tau}^q(s){A_\tau}^q(t) = \mathbb{E}{A_\tau}^m(s){A_\tau}^m(t){A_\tau}^\varepsilon(s){A_\tau}^\varepsilon(t).$$

with

$${A_\tau}^\varepsilon(t) = \left(\int_t^{t+\tau} Q_r(x)dx\right)^\varepsilon \geq \tau^{\varepsilon-1}\int_t^{t+\tau} Q_r^\varepsilon(x)dx.$$

With the same arguments as before, we get:

$$\mathbb{E}{A_\tau}^q(s){A_\tau}^q(t) \geq C(q)\tau^{2q+2\varphi(q)}|t-s+\tau|^{\varphi(2q)-2\varphi(q)}$$

where $C(q) > 0$ depends only on q.

Appendix D

List of Publications

Refereed Journal Articles

[A] Herwig Wendt, Patrice Abry, and Stéphane Jaffard, "Bootstrap for empirical multifractal analysis," *IEEE Signal Processing Mag.*, vol. 24, no. 4, pp. 38–48, 2007.

[B] Herwig Wendt and Patrice Abry, "Multifractality tests using bootstrapped wavelet leaders," *IEEE Trans. Signal Processing*, vol. 55, no. 10, pp. 4811–4820, 2007.

[C] Philippe Ciuciu, Patrice Abry, Cécile Rabrait and Herwig Wendt, "Log Wavelet Leaders Cumulant based Multifractal Analysis of EVI fMRI time series: evidence of scaling in ongoing and evoked brain activity," *IEEE J. of Selected Topics in Signal Processing*, to appear, 2008.

[D] Herwig Wendt, Stéphane G. Roux, Patrice Abry, and Stéphane Jaffard, "Wavelet leaders and bootstrap for multifractal analysis of images," *Signal Processing*, under review, 2008.

[E] Herwig Wendt, Patrice Abry, Stéphane G. Roux, Stéphane Jaffard and Béatrice Vedel, "Analyse multifractale d'image: l'apport des coefficients dominants," *Traitement du Signal, Special issue GRETSI 2007*, under review, 2008.

[F] Stéphane Jaffard, Patrice Abry, Herwig Wendt, Stéphane G. Roux and Béatrice Vedel, "The contribution of wavelets in multifractal analysis," *Proc. ISFMA Zhuhai Summer school on wavelets and applications*, to appear, 2008.

[G] Béatrice Vedel, Herwig Wendt, Patrice Abry and Stéphane Jaffard "Dependence structure of the increments of compound Poisson motion," *Lecture Notes in Statistics, Special issue on Statistical Dependence*, in preparation, 2008.

[H] Vladas Pipiras, Béatrice Vedel, Herwig Wendt and Patrice Abry, "The Rosenblatt process and dependence structure of wavelet coefficients," in preparation, 2008.

Conference Proceedings

[I] Patrice Abry, Stéphane Jaffard, Stéphane G. Roux, Béatrice Vedel and Herwig Wendt, "Wavelet decomposition of measures: Application to multifractal analysis of images," in *Proc. NATO-ASI Conf. on Unexploded Ordnance Detection and Mitigation*, to appear, Springer, 2008.

[J] Herwig Wendt and Patrice Abry, "Bootstrap tests for the time constancy of multifractal attributes," in *33rd IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Las Vegas, USA, 2008.

[K] Patrice Abry, Vladas Pipiras, and Herwig Wendt, "Extreme values, heavy tails and linearization effect: A contribution to empirical multifractal analysis," in *21st GRETSI Symposium on Signal and Image Processing*, Troyes, France, 2007.

[L] Herwig Wendt, Patrice Abry, and Stéphane Jaffard, "Analyse multifractale d'image: L'apport des coefficients dominants," in *21st GRETSI Symposium on Signal and Image Processing*, Troyes, France, 2007.

[M] Herwig Wendt and Patrice Abry, "Time-scale block bootstrap tests for non gaussian finite variance self-similar processes with stationary increments," in *IEEE Workshop on Statistical Signal Processing (SSP'07)*, Madison, Wisconsin, USA, 2007.

[N] Herwig Wendt, Stéphane G. Roux, and Patrice Abry, "Impact of data quantization on empirical multifractal analysis," in *32nd IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Honolulu, USA, 2007.

[O] Philippe Ciuciu, Patrice Abry, Cécile Rabrait, Herwig Wendt, and Alexis Roche, "Leader-based multifractal analysis for evi fmri time series: Ongoing versus task-related brain activity," in *IEEE International Symposium on Biomedical Imaging (ISBI)*, Washington DC, USA, 2007.

[P] Herwig Wendt, Stéphane G. Roux, and Patrice Abry, "Bootstrap for log wavelet leaders cumulant based multifractal analysis," in *14th European Signal Processing Conference (EUSIPCO)*, Florence, Italy, 2006.

[Q] Herwig Wendt and Patrice Abry, "Bootstrap for multifractal analysis," in *31st IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Toulouse, France, 2006.

Bibliography

[1] P. Abry. *Ondelettes et Turbulence - Multirésolutions, algorithmes de décomposition, invariance d'échelle et signaux de pression*. Diderot, Editeurs des sciences et des arts, Paris, France, 1997.

[2] P. Abry, R. Baraniuk, P. Flandrin, R. Riedi, and D. Veitch. Multiscale nature of network traffic. *IEEE Signal Proc. Mag.*, 19(3):28–46, 2002.

[3] P. Abry and P. Flandrin. On the initialization of the discrete wavelet transform. *IEEE Signal Proc. Let.*, 1(2):32–34, 1994.

[4] 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, editors, *Self-similar Network Traffic and Performance Evaluation*, pages 39–88, New York, 2000. Wiley.

[5] P. Abry, P. Gonçalvès, and P. Flandrin. Wavelet-based spectral analysis of $1/f$ processes. In *Proc. IEEE Int. Conf. Acoust., Speech, and Signal Processing (ICASSP)*, volume III, pages 237–240, Minneapolis, U.S.A, 1993.

[6] P. Abry, P. Gonçalvès, and P. Flandrin. Wavelets, spectrum analysis and $1/f$ processes. In A. Antoniadis and G. Oppenheim, editors, *Lecture Notes in Statistics: Wavelets and Statistics*, pages 15–29. Springer-Verlag, New York, 1995.

[7] P. Abry, P. Gonçalvès, and J. Lévy-Véhel, editors. *Lois d'échelle, fractale et ondelettes, Vol. 1 & 2*. Hermès Scientific Publications, 2002.

[8] P. Abry, S. Jaffard, and B. Lashermes. Revisiting scaling, multifractal, and multiplicative cascades with the wavelet leader lens. In *SPIE, Optic East, Wavelet Applications in Industrial Applications II*, volume 5607, pages 103–117, Philadelphia, USA, 2004.

[9] P. Abry, S. Jaffard, S.G. Roux, B. Vedel, and H. Wendt. Wavelet decomposition of measures: Application to multifractal analysis of images. In *Proc. NATO-ASI Conf. on Unexploded Ordnance Detection and Mitigation*. Springer, 2008. To appear.

[10] P. Abry and V. Pipiras. Wavelet-based synthesis of the Rosenblatt process. *Signal Process.*, 86(9):2326–2339, 2006.

[11] P. Abry, V. Pipiras, and H. Wendt. Extreme values, heavy tails and linearization effect: A contribution to empirical multifractal analysis. In *Proc. GRETSI Symposium Signal and Image Processing*, Troyes, France, 2007.

[12] P. Abry and D. Veitch. Wavelet analysis of long-range dependent traffic. *IEEE Trans. on Info. Theory*, 44(1):2–15, 1998.

[13] C. Angelini, D. Cava, G. Katul, and B. Vidakovic. Resampling hierarchical processes in the wavelet domain: A case study using atmospheric turbulence. *Physica D*, 207:24–40, 2005.

[14] J.-P. Antoine, R. Murenzi, P. Vandergheynst, and S. Twareque Ali. *Two-Dimensional Wavelets and their Relatives*. Cambridge University Press, Cambridge, 2004.

[15] A. Arneodo, E. Bacry, and J.F. Muzy. Random cascades on wavelet dyadic trees. *J. Math. Phys.*, 39(8):4142–4164, 1998.

[16] A. Arneodo, N. Decoster, P. Kestener, and S.G. Roux. A wavelet-based method for multifractal image analysis: from theoretical concepts to experimental applications. In P.W. Hawkes, B. Kazan, and T. Mulvey, editors, *Advances in Imaging and Electron Physics*, volume 126, pages 1–98. Academic Press, 2003.

[17] A. Arneodo, N. Decoster, and S.G. Roux. A wavelet-based method for multifractal image analysis. I. Methodology and test applications on isotropic and anisotropic random rough surfaces. *Eur. Phys. J. B*, 15(3):567–600, 2000.

[18] A. Arneodo, S. Manneville, and J.-F. Muzy. Towards log-normal statistics in high Reynolds number turbulence. *Eur. Phys. J. B*, 1:129–140, 1998.

[19] A. Arneodo, J.F. Muzy, and S. Roux. Experimental analysis of self-similarity and random cascade processes: application to fully developed turbulence data. *J. Phys. II*, 7:363–370, 1997.

[20] J.-M. Aubry and S. Jaffard. Random wavelet series. *Commun. Math. Phys.*, 227:483–514, 2002.

[21] E. Bacry, J. Delour, and J.F. Muzy. Multifractal random walk. *Phys. Rev. E*, 64(2):026103, 2001.

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

[23] E. Bacry and J.F. Muzy. Log-infinitely divisible multifractal processes. *Commun. Math. Phys.*, 236:449–475, 2003.

[24] E. Bacry, J.F. Muzy, and A. Arneodo. Singularity spectrum of fractal signals from wavelet analysis: Exact results. *J. Stat. Phys.*, 70(3-4):635–674, 1993.

[25] P. Barbe and P. Bertail. *The Weighted Bootstrap*, volume 98 of *Lecture Notes in Statistics*. Springer, New York, 1995.

[26] J.-M. Bardet, G. Lang, G. Oppenheim, A. Philippe, S. Stoev, and M.S. Taqqu. Semi-parametric estimation of the long-range dependence parameter: A survey. In P. Doukhan, G. Oppenheim, and M. S. Taqqu, editors, *Theory and applications of Long-range dependence*, pages 557–577, Boston, 2003. Birkhäuser.

[27] J.-M. Bardet, G. Lang, G. Oppenheim, A. Philippe, and M.S. Taqqu. Generators of long range dependent processes: a survey. In P. Doukhan, G. Oppenheim, and M. S. Taqqu, editors, *Theory and applications of Long-range dependence*, pages 579–624, Boston, 2003. Birkhäuser.

[28] V. Bareikis and R. Katilius, editors. *Proc. Int. Conf. Noise in Physical Systems and* $1/f$ *Fluctuations*, Palanga, Lithuania, 1995. World Scientific.

[29] J. Barral. Continuity of the multiftactal spectrum of a random statistically self-similar measure. *J. Theor. Probab.*, 13(4):1027–1060, 2000.

[30] J. Barral and B. Mandelbrot. Multiplicative products of cylindrical pulses. *Probab. Theory Rel.*, 124:409–430, 2002.

[31] A. Benassi, S. Cohen, J. Istas, and S. Jaffard. Identification of elliptic gaussian random processes. In J. Lévy-Véhel and C. Tricot, editors, *Fractals and engineering*, pages 115–123. Springer-Verlag, 1997.

[32] A. Benassi, S. Jaffard, and D. Roux. Gaussian processes and pseudodifferential elliptic operators. *Revista Mathematica Iberoamericano*, 13(1):19–89, 1997.

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

[34] R. Beran. Simulated power functions. *Ann. Stat.*, 14(1):151–173, 1986.

[35] R. Beran. Prepivoting to reduce level error of confidence sets. *Biometrika*, 74(3):457–468, 1987.

[36] R. Beran. Prepivoting test statistics: A bootstrap view of asymptotic refinements. *J. Am. Stat. Assoc.*, 83(403):687–697, 1988.

[37] R.N. Bhattacharya and J.K. Ghosh. On the validity of the formal edgeworth expansion. *Ann. Stat.*, 7:434–451, 1978.

[38] P.J. Bickel and D.A. Freedman. Some asymptotic theory for the bootstrap. *Ann. Stat.*, 9(6):1196–1217, 1981.

[39] P. Borgnat, P.-O. Amblard, and P. Flandrin. Scale invariances and Lamperti transformations for stochastic processes. *J. Phys. A : Math. Gen.*, 38:2081–2101, 2005.

[40] M. Breakspear, M. Brammer, and P.A. Robinson. Construction of multivariate surrogate stets from nonlinear data using the wavelet transform. *Physica D*, 182(1–2):1–22, 2003.

[41] E. Bullmore, C. Long, J. Suckling, J. Fadili, G. Calvert, F. Zelaya, T.A. Carpenter, and M. Brammer. Colored noise and computational inference in neurophysiological (fMRI) time series analysis: Resampling methods in time and wavelet domains. *Hum. Brain Mapp*, 12(2):61–78, 2001.

[42] E. Carlstein. The use of subseries methods for estimating the variance of a general statistic from a stationary time series. *Ann. Stat.*, 14:1171–1179, 1986.

[43] R. Carmona, W.-L. Hwang, and B. Torresani. *Practical Time-Frequency Analysis - Gabor and Wavelet Transforms with an Implementation in S*. Academic Press, 1998.

[44] B. Castaing, Y. Gagne, and M. Marchand. Log-similarity for turbulent flows. *Physica D*, 68(3-4):387–400, 1993.

[45] P. Chainais. *Cascades log-infiniment divisibles et analyse multirésolution. Application à l'étude des intermittences en turbulence.* PhD thesis, Ecole Normale Supérieure de Lyon, 2001.

[46] P. Chainais, R. Riedi, and P. Abry. On non scale invariant infinitely divisible cascades. *IEEE Trans. Inform. Theory*, 51(3), March 2005.

[47] P. Chainais, R. Riedi, and P. Abry. Warped infinitely divisible cascades: beyond scale invariance. *Traitement du Signal*, 22(1), 2005.

[48] M.R. Chernick. *Bootstrap Methods: A Practitioner's Guide*. Wiley, 1999.

[49] A. Chhabra, C. Meneveau, R.V. Jensen, and K. R. Sreenivasan. Direct determination of the singularity spectrum and its application to fully developed turbulence. *Phys. Rev. A*, 40(9):5284 – 5294, 1989.

[50] P. Ciuciu, P. Abry, C. Rabrait, and H. Wendt. Log wavelet leaders cumulant based multifractal analysis of evi fmri time series: evidence of scaling in ongoing and evoked brain activity. *IEEE J. of Selected Topics in Signal Process.*, 2008. To appear.

[51] P. Ciuciu, P. Abry, C. Rabrait, H. Wendt, and A. Roche. Leader-based multifractal analysis for evi fmri time series: Ongoing versus task-related brain activity. In *Proc. IEEE Int. Symposium Biomedical Imaging (ISBI)*, Washington DC, 2007.

[52] J.-F. Coeurjolly. Estimating the parameters of a fractional Brownian motion by discrete variations of its sample paths. *Statistical Inference for Stochastic Processes*, 4(2):199–227, 2001.

[53] A. Conci and L.H. Monteiro. Multifractal characterization of texture-based segmentation. In *Proc. IEEE Int. Conf. Image Processing (ICIP)*, volume 1, pages 792–795, Vancouver, BC, Canada, 2000.

[54] I. Daubechies. Otrthonormal bases of compactly supported wavelets. *Comm. Pure and App. Math.*, 41:909–996, 1988.

[55] I. Daubechies. *Ten Lectures on Wavelets*. SIAM, New York, 1992.

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

[57] N. Decoster, S.G. Roux, and A. Arneodo. A wavelet-based method for multifractal image analysis. II. Applications to synthetic multifractal rough surfaces. *Eur. Phys. J. B*, 15(4):739–764, 2000.

[58] L. Delbeke. *Wavelet based estimators for the Hurst parameter of a self-similar process*. PhD thesis, K.U. Leuven, Department of Science, 1998.

[59] L. Delbeke and W. Van Assche. A wavelet based estimator for the parameter of self-similarity of fractional brownian motion. In *Int. Conf. on Approximation and Optimization in the Caribbean location, Puebla, 1995. Aportaciones Matematicas: Communicaciones, vol. 24, Sociedad Matematica Mexicana*, pages 65–76, Mexico, 1998.

[60] J. Delour, J.F. Muzy, and A. Arneodo. Intermittency of 1d velocity spatial profiles in turbulence: A magnitude cumulant analysis. *Eur. Phys. J. B*, 23(2):243–248, 2001.

[61] T. DiCiccio and B. Efron. Bootstrap confidence intervals. *Stat. Sci.*, 11(3):189–212, 1996.

[62] R.L. Dobrushin and P. Major. Non-central limit theorems for non-linear functionals of Gaussian random fields. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 50:27–52, 1979.

[63] P. Doukhan. *Mixing: Properties and Examples*, volume 85 of *Lecture Notes in Statistics*. Springer, New York, 1994.

[64] B. Efron. Bootstrap methods: Another look at the jackknife. *Ann. Stat.*, 7:1–26, 1979.

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

[66] B. Efron. Better bootstrap confidence intervals. *J. Am. Stat. Assoc.*, 82(397):171–185, 1987.

[67] B. Efron and R. Tibshirani. Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. *Stat. Sci.*, 1:54–77, 1986.

[68] B. Efron and R. Tibshirani. *An Introduction to the Bootstrap*. Chapman and Hall, London, 1993.

[69] P. Embrechts, C. Klüppelberg, and T. Mikosch. *Modelling extremal events*. Springer-Verlag, Berlin, 1997.

[70] P. Flandrin. Wavelet analysis and synthesis of fractional Brownian motions. *IEEE Trans. Inform. Theory*, 38:910–917, 1992.

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

[72] P. Frankhauser. L'approche fractale : un nouvel outil dans l'analyse spatiale des agglomerations urbaines. *Population*, 4:1005–1040, 1997.

[73] D.A. Freedman. Bootstrapping regression models. *Ann. Stat.*, 9:1218–1228, 1981.

[74] D.A. Freedman. On bootstrapping two-stage least-squares estimates in stationary linear models. *Ann. Stat.*, 12:827–842, 1984.

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

[76] U. Frisch and G. Parisi. Fully developed turbulence and intermittency. In *Proc. Int. Summer School Turbulence and Predictability in Geophysical Fluid Dynamics and Climate Dynamics*, pages 84–88, Amsterdam, 1985.

[77] P. Gonçalvès and R. Riedi. Wavelet analysis of fractional Brownian motion in multifractal time. In *Proc. GRETSI Symposium Signal and Image Processing*, Vannes, France, 1999.

[78] R.M. Gray and D.L. Neuhoff. Quantization. *IEEE Trans. Inform. Theory*, 44(6):2325–2383, 1998.

[79] P. Hall. Resampling a coverage pattern. *Stochastic Processes and Their Application*, 20:231–246, 1985.

[80] P. Hall. Theoretical comparison of bootstrap confidence intervals. *Ann. Stat.*, 16(3):927–953, 1988.

[81] P. Hall. *The Bootstrap and Edgeworth Expansion*. Springer, New York, 1992.

[82] P. Hall and S.R. Wilson. Two guidelines for bootstrap hypothesis testing. *Biometrics*, 47(2):757–762, 1991.

[83] D.R. Iskander, W. Alkhaldi, and A.M. Zoubir. On the computer intensive methods in model selection. In *Proc. IEEE Int. Conf. Acoust., Speech, and Signal Processing (ICASSP)*, Las Vegas, USA, 2008.

[84] J. Istas and G. Lang. Quadratic variations and estimation of the local Hölder index of a Gaussian process. *Annales de l'institut Henri Poincaré (B) Probabilités et Statistiques*, 33(4):407–436, 1997.

[85] S. Jaffard. Private communication. 2008.

[86] S. Jaffard. Multifractal formalism for functions part i: Results valid for all functions. *SIAM J. Math. Anal.*, 28(4):944–970, 1997.

[87] S. Jaffard. Multifractal formalism for functions part ii: Self-similar functions. *SIAM J. Math. Anal.*, 28(4):971–998, 1997.

[88] S. Jaffard. The multifractal nature of Lévy processes. *Probab. Theory Rel.*, 114(2):207–227, 1999.

[89] S. Jaffard. Wavelet techniques in multifractal analysis. In M. Lapidus and M. van Frankenhuijsen, editors, *Fractal Geometry and Applications: A Jubilee of Benoît Mandelbrot, M. Lapidus and M. van Frankenhuijsen Eds., Proceedings of Symposia in Pure Mathematics*, volume 72(2), pages 91–152. AMS, 2004.

[90] S. Jaffard, P. Abry, S.G. Roux, B. Vedel, and H. Wendt. The contribution of wavelets in multifractal analysis. In *Proc. ISFMA Zhuhai Summer school Wavelets and Applications*, 2008. To appear.

[91] S. Jaffard, P. Abry, H. Wendt, S.G. Roux, and B. Vedel. Wavelet analysis of multifractal measures. *In preparation.*, 2008.

[92] S. Jaffard, B. Lashermes, and P. Abry. Wavelet leaders in multifractal analysis. In T. Qian, M.I. Vai, and X. Yuesheng, editors, *Wavelet Analysis and Applications*, pages 219–264, Basel, Switzerland, 2006. Birkhäuser Verlag.

[93] R. Jennane, W. J. Ohley, S. Majumdar, and G. Lemineur. Fractal analysis of bone x-ray computed microscopy projections. *IEEE Trans. Med. Imaging*, 20(5):443–449, 2001.

[94] H. Kahalerras, Y. Malécot, Y. Gagne, and B. Castaing. Intermittency and Reynolds number. *Physics of Fluids*, 10(4):910–921, 1998.

[95] J.-P. Kahane and J. Peyrière. Sur certaines martingales de Benoit Mandelbrot. *Adv. Math.*, 22:131–145, 1976.

[96] L. M. Kaplan. Extended fractal analysis for texture classification and segmentation. *IEEE Trans. Pattern Anal.*, 8(11):1572–1585, 1999.

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

[98] M. Keshner. $1/f$ noise. *Proc. IEEE*, 70(3):212–218, 1982.

[99] A. N. Kolmogorov. A refinement hypothesis concerning the local structure of turbulence in a viscous incompressible fluid at high Reynolds number. *J. Fluid Mech.*, 13:82–85, 1962.

[100] A.N. Kolmogorov. a) dissipation of energy in the locally isotropic turbulence. b) the local structure of turbulence in incompressible viscous fluid for very large Reynolds number. c) on degeneration of isotropic turbulence in an incompressible viscous liquid. In S.K. Friedlander and L. Topper, editors, *Turbulence, Classic papers on statistical theory*, pages 151–161. Interscience publishers, 1941.

[101] J.P. Kreiss and J. Franke. Bootstrapping stationary autoregressive moving-average models. *Journal of Time Series Analysis*, 13:297–317, 1992.

[102] H.R. Künsch. The jackknife and the bootstrap for general stationary observations. *Ann. Stat.*, 17:1217–1261, 1989.

[103] S.N. Lahiri. On the moving block bootstrap under long range dependence. *Statistics and Probability Letters*, 18:405–413, 1993.

[104] S.N. Lahiri. *Resampling Methods for Dependent Data*. Springer, New York, 2003.

[105] J. Lamperti. Semi-stable stochastic processes. *Trans. Am. Math. Soc.*, 104:62–78, 1962.

[106] M. Lapidus and M. van Frankhuysen. *Fractal Geometry and Number Theory*. Birkhäuser, 2000.

[107] B. Lashermes. *Analyse multifractale pratique : coefficients dominants et ordres critiques. Applications à la turbulence pleinement développée. Effets de nombre de Reynolds fini.* PhD thesis, Ecole Normale Supérieure de Lyon, 2005.

[108] B. Lashermes, P. Abry, and P. Chainais. New insight in the estimation of scaling exponents. *Int. J. Wavelets Multi.*, 2(4):497–523, 2004.

[109] B. Lashermes, P. Abry, and P. Chainais. Scaling exponents estimation for multiscaling processes. In *Proc. IEEE Int. Conf. Acoust., Speech, and Signal Processing (ICASSP)*, Montreal, Canada, 2004.

[110] B. Lashermes, S. Jaffard, and P. Abry. Wavelet leader based multifractal analysis. In *Proc. IEEE Int. Conf. Acoust., Speech, and Signal Processing (ICASSP)*, Philadelphia, USA, 2005.

[111] B. Lashermes, S. Roux, P. P. Abry, and S. Jaffard. Comprehensive multifractal analysis of turbulent velocity using the wavelet leaders. *Eur. Phys. J. B*, 61(2):201–215, 2008.

[112] S. Lazebnik, C. Schmid, and J. Ponce. A sparse texture representation using affine-invariant regions. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 8(27):1265–1278, 2005.

[113] E.L. Lehmann. *Testing Statistical Hypotheses*. Wiley, New York, 1959.

[114] L.S. Liebovitch and A.T. Todorov. Fractal dynamics of human gait: stability of long-range correlations in stride interval fluctuations. *J. Appl. Physiol.*, 80(5):1446–1447, 1996.

[115] R.Y. Liu and K. Singh. Moving blocks jackknife and bootstrap capture weak dependence. In R. Lepage and L. Billards, editors, *Exploring the Limits of the Bootstrap*, pages 225–248, New York, 1992. Wiley.

[116] S.B. Lowen and M.C. Teich. *Fractal-Based Point Processes*. Wiley, Hoboken, NJ, 2005.

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

[118] N. Mallick, P.P. Cortet, S. Santucci, S.G. Roux, and L. Vanel. Discrepancy between subcritical and fast rupture roughness: A cumulant analysis. *Phys. Rev. Lett.*, 98:255502, 2007.

[119] E. Mammen. *When Does Bootstrap Work? Asymptotic Results and Simulations*, volume 77 of *Lecture Notes in Statistics*. Springer, New York, 1992.

[120] E. Mammen. Bootstrap and wild bootstrap for high dimensional linear models. *Ann. Stat.*, 21:255–285, 1993.

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

[122] B.B. Mandelbrot. Limit lognormal multifractal measures. In E. A. Gotsman, Y. Ne'eman, and A. Voronel, editors, *Frontiers of Physics, Proc. Landau Memorial Conf., Tel Aviv, 1988*, pages 309–340, Oxford, 1990. Pergamon Press.

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

[124] B.B. Mandelbrot. A multifractal walk down Wall Street. *Sci. Am.*, 280(2):70–73, 1999.

[125] Y. Meyer. *Oscillating patterns in image processing and nonlinear evolution equations*, volume 22 of *University Lecture Series*. American Mathematical Society, Providence, RI, 2001. The fifteenth Dean Jacqueline B. Lewis memorial lectures.

[126] S.J. Mitchell. Discontinuities in self-affine functions lead to multiaffinity. *Phys. Rev. E*, 72:4065103(R), 2005.

[127] G.M. Molchan. Scaling exponents and multifractal dimensions for independent random cascades. *Commun. Math. Phys.*, 179(3):681–702, 1996.

[128] A.S. Monin and A.M. Yaglom. *Statistical Fluid Mechanics*. M.I.T. Press, Cambridge, MA, 1975.

[129] J.F. 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*, 47(2):875–884, 1993.

[130] J.F. Muzy, E. Bacry, and A. Arneodo. The multifractal formalism revisited with wavelets. *Int. J. Bifurcat. Chaos*, 4(2):245–301, 1994.

[131] J.F. Muzy, E. Bacry, and A. Kozhemyak. Extreme values and fat tails of multifractal fluctuations. *Phys. Rev. E*, 73:066114, 2006.

[132] S. Nicolay. *Analyse des séquences d'ADN par la transformée en ondelettes : extraction d'informations structurelles, dynamiques et fonctionnelles*. PhD thesis, University of Liège, Belgium, 2006.

[133] A.M. Obukhov. Some specific features of atmospheric turbulence. *J. Fluid Mech.*, 13:77–81, 1962.

[134] M. Ossiander and E.C. Waymire. Statistical estimation for multiplicative cascades. *Ann. Stat.*, 28(6):1533–1560, 2000.

[135] K. Park and W. Willinger. Self-similar network traffic: An overview. In K. Park and W. Willinger, editors, *Self-Similar Network Traffic and Performance Evaluation*, pages 1–38, New York, 2000. Wiley.

[136] S. Peleg, J. Naor, R. Hartley, and D. Avnir. Multiple resolution texture analysis and classification. *IEEE Trans. Pattern Anal.*, 6(4):518–523, 1984.

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

[138] V. Pipiras, P. Chainais, L. Coutin, and P. Abry. Multifractal random walks as fractional wiener integrals. *Information Theory*, 2008. Under review.

[139] V. Pipiras, B. Vedel, H. Wendt, and P. Abry. The Rosenblatt process and dependence structure of wavelet coefficients. *In preparation*, 2008.

[140] D.N. Politis and J.P. Romano. A circular block resampling procedure for stationary data. In R. Lepage and L. Billard, editors, *Exploring the Limits of the Bootstrap*, pages 263–270, New York, 1992. Wiley.

[141] D.N. Politis and J.P. Romano. The stationary bootstrap. *Journal of the American Statistical Association*, 89:1303–1313, 1994.

[142] L. Ponson, D. Bonamy, H. Auradou, G. Mourot, S. Morel, E. Bouchaud, C. Guillot, and J.P. Hulin. Anisotropic self-affine properties of experimental fracture surface. *Int. J. Fracture*, 140(1-4):27–36, 2006.

[143] V. J. Ribeiro, R. H. Riedi, and R. G. Baraniuk. Wavelets and multifractals for network traffic modeling and inference. In *Proc. IEEE Int. Conf. Acoust., Speech, and Signal Processing (ICASSP)*, Salt Lake City, Utah, USA, 2001.

[144] L.F. Richardson. *Weather prediction by numerical process*. Cambridge University Press, Cambridge, 1922.

[145] R.H. Riedi. Multifractal processes. In P. Doukhan, G. Oppenheim, and M.S. Taqqu, editors, *Theory and applications of long range dependence*, pages 625–717. Birkhäuser, 2003.

[146] M. Rosenblatt. Independence and dependence. In *Proc. Berkeley Symposium Mathematics, Statistics and Probability*, volume 2, pages 431–443, Berkeley, CA, 1961.

[147] S.G. Roux, A. Arneodo, and N. Decoster. A wavelet-based method for multifractal image analysis. III. Applications to high-resolution satellite images of cloud structure. *Eur. Phys. J. B*, 15(4):765–786, 2000.

[148] S.G. Roux, B. Lashermes, P. Abry, and S. Jaffard. Contributions à l'étude des performances statistiques des estimateurs multifractals. In *Proc. GRETSI Symposium Signal and Image Processing*, Louvain-la-Neuve, Belgium, 2005.

[149] L. Rudin and S. Osher. Nonlinear total variation based noise removal algorithms. *Physica D*, 60:259–268, 1992.

[150] G. Ruiz Chavarria, C. Baudet, and S. Ciliberto. Hierarchy of the energy dissipation moments in fully developed turbulence. *Phys. Rev. Lett.*, 74:1986–1989, 1995.

[151] A.M. Sabatini. Wavelet-based estimation of $1/f$-type signal parameters: Confidence intervals using the bootstrap. *IEEE Trans. Signal Process.*, 47(12):3406–3409, 1999.

[152] A.M. Sabatini. A wavelet-based bootstrap method applied to inertial sensor stochastic error modelling using Allan variance. *Meas. Sci. Technol.*, 17:2980–2988, 2006.

[153] G. Samorodnitsky and M. Taqqu. *Stable non-Gaussian random processes*. Chapman and Hall, New York, 1994.

[154] S. Santucci, K.-J. Måløy, A. Delaplace, J. Mathiesen, A. Hansen, J.Ø. Haavig Bakke, J. Schmittbuhl, L. Vanel, and P. Ray. Statistics of fracture surfaces. *Phys. Rev. E*, 75:016104, 2007.

[155] D. Schertzer, S. Lovejoy, F. Schmitt, Y. Ghigisinskaya, and D. Marsan. Multifractal cascade dynamics and turbulent intermittency. *Fractals*, 5(3):427–471, 1997.

[156] F. Schmitt and D. Marsan. Stochastic equations generating continuous multiplicative cascades. *Eur. Phys. J. B*, 20(1):3–6, 2001.

[157] J. Shao and D. Tu. *The Jackknife and Bootstrap*. Springer, New York, 1995.

[158] V. Sharifi-Salamantian, B. Pesquet-Popescu, J. Simoni-Lafontaine, and J. P. Rigaut. Index for spatial heterogeneity in breast cancer. *J. Microsc.-Oxford*, 216(2):110–122, 2004.

[159] Z.S. She and E. Lévêque. Universal scaling laws in fully developed turbulence. *Phys. Rev. Lett.*, 72(3):336–339, 1994.

[160] K. Sigman. Appendix: a primer on heavy-tailed distributions. *Queueing Syst.*, 33(1-3):261–275, 1999.

[161] K. Singh. On the asymptotic accuracy of Efron's bootstrap. *Ann. Stat.*, 9(6):1187–1195, 1981.

[162] J.-L. Starck, F. Murtagh, and A. Bijaoui. *Image Processing and Data Analysis: The Multiscale Approach*. Cambridge University Press, Cambridge, 1998.

[163] M.L. Stein. Fast and exact simulation of fractional Brownian surfaces. *J. Comput. Graph. Stat.*, 11(3):587–599, 2002.

[164] M.S. Taqqu. Weak convergence to fractional Brownian motion and to the Rosenblatt process. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 31:287–302, 1975.

[165] M.S. Taqqu. Convergence of integrated processes of arbitrary hermite rank. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 50:53–83, 1979.

[166] A.H. Tewfik and M. Kim. Correlation structure of the discrete wavelet coefficients of fractional Brownian motions. *IEEE Trans. Inform. Theory*, IT-38(2):904–909, 1992.

[167] M. Varma and A. Zissermann. Classifying images of materials: Achieving viewpoint and illumination independence. *ECCV*, 3:255–271, 2002.

[168] B. Vedel, H. Wendt, P. Abry, and S. Jaffard. Dependence structure of the increments of compound Poisson motion. *Lecture Notes in Statistics, Special issue on Statistical Dependence*, 2008. In preparation.

[169] D. Veitch and P. Abry. A wavelet-based joint estimator of the parameters of long-range dependence. *IEEE Trans. Inform. Theory*, 45(3):878–897, 1999.

[170] D. Veitch and P. Abry. A statistical test for the time constancy of scaling exponents. *IEEE Trans. Signal Proces.*, 49(10):2325–2334, 2001.

[171] D. Veitch, P. Abry, and M. Taqqu. On the automatic selection of the onset of scaling. *Fractals*, 11(4):377–390, 2003.

[172] D. Veitch, M. Taqqu, and P. Abry. Meaningful MRA initialization for discrete time series. *Signal Process.*, 80(9):1971–1983, 2000.

[173] H. Wendt and P. Abry. Bootstrap for multifractal analysis. In *Proc. IEEE Int. Conf. Acoust., Speech, and Signal Processing (ICASSP)*, Toulouse, France, 2006.

[174] H. Wendt and P. Abry. Multifractality tests using bootstrapped wavelet leaders. *IEEE Trans. Signal Proces.*, 55(10):4811–4820, 2007.

[175] H. Wendt and P. Abry. Time-scale block bootstrap tests for non gaussian finite variance self-similar processes with stationary increments. In *Proc. IEEE Workshop Statistical Signal Processing (SSP)*, Madison, Wisconsin, USA, 2007.

[176] H. Wendt and P. Abry. Bootstrap tests for the time constancy of multifractal attributes. In *Proc. IEEE Int. Conf. Acoust., Speech, and Signal Processing (ICASSP)*, Las Vegas, USA, 2008.

[177] H. Wendt, P. Abry, and S. Jaffard. Analyse multifractale d'image: L'apport des coefficients dominants. In *Proc. GRETSI Symposium Signal and Image Processing*, Troyes, France, 2007.

[178] H. Wendt, P. Abry, S.G. Roux, S. Jaffard, and B. Vedel. Analyse multifractale d'images : l'apport des coefficients dominants. *Traitement du Signal*, 2008. Submitted.

[179] H. Wendt, P. Abry, and S.Jaffard. Bootstrap for empirical multifractal analysis. *IEEE Signal Proc. Mag.*, 24(4):38–48, 2007.

[180] H. Wendt, S.G. Roux, and P. Abry. Bootstrap for log wavelet leaders cumulant based multifractal analysis. In *Proc. European Signal Processing Conf. (EUSIPCO)*, Florence, Italy, 2006.

[181] H. Wendt, S.G. Roux, and P. Abry. Impact of data quantization on empirical multifractal analysis. In *Proc. IEEE Int. Conf. Acoust., Speech, and Signal Processing (ICASSP)*, Honolulu, USA, 2007.

[182] H Wendt, S.G. Roux, P. Abry, and S. Jaffard. Wavelet leaders and bootstrap for multifractal analysis of images. *Signal Process.*, 2008. Under review.

[183] B. Whitcher. Wavelet-based bootstrap of spatial patterns on a finite lattice. *Computational Stat. and Data Analysis*, 50:2399–2421, 2006.

[184] B. Widrow, I. Kollár, and M.-C. Liu. Statistical theory of quantization. *IEEE. Trans. Instrument. Measurement*, 45(2):353–361, 1996.

[185] Y. Xu, J. Hui, and C. Fermüller. A projective invariant for texture. In *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, pages 1932–1939, New York, USA, 2006.

[186] A.M. Yaglom. Effect of fluctuations in energy dissipation rate on the form of turbulence characteristics in the inertial subrange. *Dokl. Akad. Nauk. SSR*, 166:49–52, 1966.

[187] A.M. Zoubir. Bootstrap: theory and applications. In F.T. Luk, editor, *Proc. SPIE, Advanced Signal Processing Algorithms, Architectures, and Implementations IV*, volume 2027, pages 216–235, San Diego, CA, USA, 1993.

[188] A.M. Zoubir. Multiple bootstrap tests and their application. In *Proc. IEEE Int. Conf. Acoust., Speech, and Signal Processing (ICASSP)*, Adelaide, Australia, 1994.

[189] A.M. Zoubir. The bootstrap and its application in signal processing. *IEEE Signal Proc. Mag.*, 15(1):56–76, 1998.

[190] A.M. Zoubir. Bootstrap based spectral analysis. In *Proc. IEEE Workshop Statistical Signal Processing (SSP)*, Bordeaux, France, 2005.

[191] A.M. Zoubir. On confidence intervals for the coherence function. In *Proc. IEEE Int. Conf. Acoust., Speech, and Signal Processing (ICASSP)*, Philadelphia, USA, 2005.

[192] A.M. Zoubir. Bootstrap based confidence intervals for the conditional coherence. In *Proc. IEEE Int. Conf. Acoust., Speech, and Signal Processing (ICASSP)*, Honolulu, USA, 2007.

[193] A.M. Zoubir and D.R. Iskander. Bootstrapping bispectra: An application to testing departure from gaussianity of stationary signals. *IEEE Trans. Signal Proces.*, 47(3):880–884, 1999.

[194] A.M. Zoubir and D.R. Iskander. Bootstrap modeling of a class of nonstationary signals. *IEEE Trans. Signal Proces.*, 48(2):399–408, 2000.

[195] A.M. Zoubir and D.R. Iskander. *Bootstrap Techniques for Signal Processing*. Cambridge University Press, 2004.

[196] A.M. Zoubir and D.R. Iskander. Bootstrap methods and applications. *IEEE Signal Proc. Mag.*, 24:10–19, 2007.

Résumé : L'invariance d'échelle constitue un paradigme souvent avancé pour l'analyse et la modélisation de données expérimentales issues d'applications de natures différentes. L'analyse multifractale fournit un cadre conceptuel pour ses études théorique et pratique. Dans ce contexte, l'objectif de cette thèse réside dans l'apport à l'analyse multifractale, de l'utilisation des coefficients d'ondelettes dominants, d'une part, et des techniques statistiques de type bootstrap, d'autre part. Dans la première partie de ce travail, les propriétés et performances statistiques de procédures d'analyse multifractale construites à partir de coefficients dominants sont étudiées et caracterisées. Il est notamment montré qu'elles se comparent favorablement à celles obtenues à partir de coefficients d'ondelettes. De plus, une extension aux signaux bidimensionnels (images) est proposée et validée. En complément sont étudiées plusieurs difficultés théoriques, d'importance cruciale pour une réelle mise en oeuvre pratique de l'analyse multifractale : régularité minimale et espaces fonctionnels, effet de linearisation, robustesse vis-à-vis d'éventuelles quantifications des données. La deuxième partie de ce travail de thèse s'intéresse à la construction, pour les attributs multifractals, d'intervalles de confiance et de tests d'hypothèse, à partir de techniques 'bootstrap'. L'originalité de notre approche réside dans la mise en œuvre du bootstrap par construction de blocs temps-échelle dans le plan des coefficients d'ondelettes. Cette procédure, validée par simulations numériques, permet d'obtenir des intervalles de confiance et d'effectuer des tests d'hypothèses à partir d'une seule observation des données, de longueur finie. Une étude précise des structures de dépendance des coefficients d'ondelettes et coefficients dominants complète ce travail. Elle montre notamment que l'augmentation du nombre de moments nuls de l'ondelette d'analyse, qui, pour le mouvement brownien fractionnaire, permet de réduire la portée de la structure de dépendance de longue à courte, est inopérante pour les cascades multiplicatives multifractales : si l'augmentation du nombre de moments nuls décorrèle effectivement les coefficients d'ondelette, elle échoue à faire disparaître la dépendance longue. Enfin, les procédures d'analyse multifractale par coefficients dominants et bootstrap sont illustrées sur deux applications : la turbulence hydrodynamique et la classification de texture d'images.

Mots clefs : *Analyse multifractale, Invariance d'échelle, coefficients d'ondelettes dominants, Analyse multirésolution, Bootstrap, Intervalles de confiance, Tests d'hypothèse, Turbulence hydrodynamique, Traitement d'image.*

Abstract: Scale invariance constitutes a paradigm that is frequently used for the analysis and modeling of empirical data in various applications of very different natures. Multifractal analysis provides a conceptual framework for its theoretical and practical studies. The aim of this thesis is to investigate the benefits of the use of wavelet Leaders, on one hand, and bootstrap methods, for practical multifractal analysis. In the first part of this work, the statistical properties and performance of wavelet Leader based multifractal analysis procedures are studied. It is shown that they compare very favorably to those obtained by wavelet coefficient based ones. Moreover, a practical extension to two dimensional signals (images) is validated. In addition, a number of theoretical questions of fundamental practical importance in applications are investigated: Function space embedding models and minimum regularity, linearization effect, robustness with respect to quantization of the data. The second part of this thesis proposes bootstrap based procedures for statistical inference in multifractal analysis. These procedures are validated by numerical simulations and permit the construction of confidence intervals and hypothesis tests for multifractal attributes, from one single finite length observation of data. This is achieved by an original time-scale block bootstrap approach in the wavelet domain. This work is further completed by the detailed study of the dependence structures of wavelet coefficients and wavelet Leaders. Notably, it is shown that the number of vanishing moments of the analyzing wavelet, which permits to convert long range to weak dependence for fractional Brownian motion, is ineffective for multifractal multiplicative cascades: Increasing the number of vanishing moments still controls the correlation of wavelet coefficients, but has no effect on their long range dependence structure. Finally, the wavelet Leader and bootstrap based multifractal analysis tools are applied to hydrodynamic turbulence data, and to texture image classification.

Keywords: *Multifractal analysis, Scale invariance, Wavelet Leaders, Multiresolution analysis, Bootstrap, Confidence intervals, Hypothesis tests, Hydrodynamic turbulence, Image processing.*