A bridge between geometric measure theory and signal processing: Mutifractal analysis

P. Abry[*], S. Jaffard[†], H. Wendt[‡]

June 25, 2013

Abstract:
We describe the main features of wavelet techniques in multifractal analysis, using wavelet bases both as a tool for analysis, and for synthesis. We focus on two promising developments: On the analysis side, we introduce the quantile leader method, which allows to put into light nonconcave multifractal spectra; on the synthesis side, we study some extensions of random wavelet series which allow to test multifractal techniques fitted to functions that are not locally bounded but only belong to an L^q space (determination of the q-spectrum).

Keywords: *Scaling, Scale Invariance, Fractal, Multifractal, Fractional dimensions, Hölder regularity, Oscillations, Wavelet, Wavelet Leader, Multifractal Spectrum, p-exponent, p-variation*

[*]Signal, Systems and Physics, Physics Dept., CNRS UMR 5672, Ecole Normale Supérieure de Lyon, Lyon, France patrice.abry@ens-lyon.fr

[†]Address: Université Paris Est, Laboratoire d'Analyse et de Mathématiques Appliquées, CNRS, UMR 8050, UPEC, Créteil, France jaffard@u-pec.fr

[‡]IRIT - ENSEEIH, 2 rue Camichel, B.P. 7122, 31071 Toulouse cedex France, herwig.wendt@irit.fr

Contents

1 Introduction

On the mathematical side, fractal geometry has two different origins: One is the quest for non-smooth functions, rising from the following key question that motivated a large part of the progresses in analysis during the nineteenth century: How irregular can a continuous function be? And, more precisely, does a continuous function necessarily have points of differentiability? The first (negative) answer to this question was supplied by B. Bolzano in the first half of the 19th century: He constructed a counter-example, which actually turned out to be, historically, the first example of a *multifractal function*. This example, however, had no direct influence, because it remained unpublished. On the contrary, the next counterexamples, namely the *Weierstrass functions*

$$\mathcal{W}_{a,H}(x) = \sum_{n=0}^{+\infty} \frac{\sin(a^n x)}{a^{Hn}} \qquad \text{for} \quad a > 1 \quad \text{and} \quad H \in (0,1) \tag{1}$$

had a deep impact on the developments of analysis in the 19th century. The fact that these functions are continuous and nowhere differentiable can be sharpened in a way which requires the notion of *pointwise Hölder regularity*.

Definition 1 *Let $f : \mathbb{R}^d \to \mathbb{R}$ be a locally bounded function, $x_0 \in \mathbb{R}^d$ and let $\gamma \geq 0$; f belongs to $C^\gamma(x_0)$ if there exist $C > 0$, $R > 0$ and a polynomial P of degree less than γ such that*

$$\text{if } |x - x_0| \leq R, \quad \text{then} \qquad |f(x) - P(x - x_0)| \leq C|x - x_0|^\gamma.$$

The Hölder exponent of f at x_0 is

$$h_f(x_0) = \sup\{\gamma : \quad f \quad \text{is} \quad C^\gamma(x_0)\}.$$

The polynomial P is clearly unique, and will be referred to in the following as the *Taylor polynomial* of f at x_0. Note that differentiability at x_0 implies that $h_f(x_0) \geq 1$.

The Hölder exponent of $\mathcal{W}_{a,H}$ is a constant function, which is equal to H at every point; since $H < 1$ we thus recover the fact that $\mathcal{W}_{a,H}$ is nowhere differentiable, but the sharper notion of Hölder exponent allows to draw a difference between each of the Weierstrass functions, and classify them using a regularity parameter.

The connection with fractal geometry follows from the fact that the graphs of these functions supply important examples of fractal sets that still motivate research. In order to make this point explicit, we need to recall the notion of *box dimension* which is commonly used to classify fractal sets.

Definition 2 *Let A be a bounded subset of $\mathbb{R}^d$; if $\varepsilon > 0$, let $N_\varepsilon(A)$ be the smallest number such that there exists a covering of A by $N_\varepsilon(A)$ balls of radius ε.*

The upper and lower box dimension of A are respectively given by

$$\overline{dim}_B(A) = \limsup_{\varepsilon \to 0} \frac{\log N_\varepsilon(A)}{-\log \varepsilon}, \qquad \text{and} \qquad \underline{dim}_B(A) = \liminf_{\varepsilon \to 0} \frac{\log N_\varepsilon(A)}{-\log \varepsilon}. \tag{2}$$

When both limits coincide (as it is the case for the graphs of the Weierstrass functions), they are referred to as the *box dimension* of the set A:

$$dim_B(A) = \lim_{\varepsilon \to 0} \frac{\log N_\varepsilon(A)}{-\log \varepsilon}. \tag{3}$$

The box dimension of the graph of $\mathcal{W}_{a,H}$ is $2 - H$; note that the determination of its Hausdorff dimension (see Definition 6 below) still is an open problem (it is however conjectured to coincide with the box dimension). Other examples of functions with fractal graphs were introduced in the nineteenth century: The proper mathematical definition of Brownian motion was obtained by Louis Bachelier in his thesis in 1900, for the purpose of modeling in finance, and its sample paths now supply some of the simplest examples of fractal sets: The box and Hausdorff dimensions of their graph a. s. is $3/2$. Furthermore Brownian motion displays the same qualitative property as the Weierstrass functions: Its Hölder exponent is constant, and related with the box dimension of its graph by the same relationship.

The other origin of fractal geometry came from the challenge of defining the notions of "length" of a curve or "area" of a surface through a definition that would not require the notion of differentiability: This requirement led to the notion of *Hausdorff outer measure* (see Definition 6 below) first in the integer dimensional case, and then in non-integer cases, and thus provided a mathematical tool fitted to the geometry of sets such as the triadic Cantor set.

However, the use of a single parameter (e.g. the fractional dimension of the graph) is too reductive as a classification tool in many situations that are met in applications. Let us now review some additional parameters which have been used.

Consider for example the case of the triadic Cantor set: The fact that it is made of two parts which are identical to the whole set scaled down by a factor of 3 leads to a *similarity dimension* of $\log 2 / \log 3$. But one easily checks that its box and Hausdorff dimensions also coincide with this number.

Let us now consider the example of *fractional Brownian motion* (hereafter referred to as fBm), a family of stochastic processes introduced by Kolmogorov [35], and whose importance for the modeling of scale invariance and fractal properties in data was made explicit by Mandelbrot and Van Ness in [44]. This family is indexed by a parameter $H \in (0, 1)$, and generalizes Brownian motion (which corresponds to the case $H = 1/2$); fBm of index H is the only centered Gaussian random process B^H satisfying

$$\forall x, y \geq 0 \qquad E(|B^H(x) - B^H(y)|^2) = |x - y|^{2H}.$$

The key role played by fBms in signal processing comes from the fact that they supply the most simple one parameter family of stochastic processes with stationary increments, and therefore are widely used in modeling.

Weierstrass-Mandelbrot functions and fBm sample paths have everywhere the same and constant Hölder exponent

$$\forall x:\ h_f(x) = H.$$

Furthermore, for these two families, the pointwise regularity exponent also coincides with a *uniform regularity exponent* which can be defined as follows.

The Lipschitz spaces $C^s(\mathbb{R}^d)$ are defined for $0 < s < 1$ by the conditions : $f \in L^\infty$ and

$$\exists C, N, \quad \forall x, y \in \mathbb{R}^d, \qquad |f(x) - f(y)| \leq C|x - y|^s.$$

If $s > 1$, they are then defined by recursion on $[s]$ by the condition: $f \in C^s(\mathbb{R}^d)$ if $f \in L^\infty$ and if all its partial derivatives (taken in the sense of distributions) $\partial f/\partial x_i$ (for $i = 1, \cdots d$) belong to $C^{s-1}(\mathbb{R}^d)$. If $s < 0$, then the C^s spaces are composed of distributions, also defined by recursion on $[s]$ as follows: $f \in C^s(\mathbb{R}^d)$ if f is a finite sum of partial derivatives (in the sense of distributions) of order 1 of elements of $C^{s+1}(\mathbb{R}^d)$. This allows to define the C^s spaces for any $s \notin \mathbb{Z}$ (note that a consistent definition using the Zygmund classes can also be supplied for $s \in \mathbb{Z}$, see [45], however we will not need to consider these specific values in the following). A distribution f belongs to $f \in C^s_{loc}$ if $f\varphi \in C^s$ for any C^∞ compactly supported function φ.

Definition 3 *The uniform Hölder exponent of a tempered distribution f is*

$$H_f^{min} = \sup\{s :\ f \in C^s_{loc}(\mathbb{R}^d)\}. \tag{4}$$

This definition does not make any a priori assumption on f: The uniform Hölder exponent is defined for any tempered distribution, and it can be positive and negative. More precisely:

- If $H_f^{min} > 0$, then f is a locally bounded function,
- if $H_f^{min} < 0$, then f is not a locally bounded function.

Other parameters can also be considered in the case of (deterministic or random) functions, which mimic, in a functional setting, the geometric property that we mentioned for the triadic Cantor set, and allow to encapsulate the intuitive idea that the graph of f "looks" the same after proper rescalings. The Weierstrass functions satisfy a deterministic selfsimilarity relationship

$$\forall x \in \mathbb{R}, \qquad \mathcal{W}_{a,H}(ax) = a^H \mathcal{W}_{a,H}(x) + g(x) \tag{5}$$

(where g is a C^∞ function) thus yielding a *selfsimilarity exponent* equal to H. On the other hand, fBm satisfies a *stochastic selfsimilarity*; this probabilistic notion means that, $\forall a > 0$, the (random) functions $a^H f(ax)$ do not coincide sample path by sample path, but share the same *statistical laws*. Recall that two vectors of $\mathbb{R}^l$: $(X_1, \cdots X_l)$ and $(Y_1, \cdots Y_l)$ share the same law if, for any Borel set $A \subset \mathbb{R}^l$, $\mathbb{P}(\{X \in A\}) = \mathbb{P}(\{Y \in A\})$. Similarly, two processes X_t and Y_t share the same law if, $\forall l \geq 1$, for any finite set of time points $t_1, \cdots t_l$,

the vectors of $\mathbb{R}^l$ $(X_{t_1}, \cdots X_{t_l})$ and$(Y_{t_1}, \cdots Y_{t_l})$ share the same law. A stochastic process X_t is said to be *selfsimilar*, with selfsimilarity exponent H, iff

$$\forall a > 0, \quad \{X_{at},\ t \in \mathbb{R}\} \stackrel{\mathcal{L}}{=} \{a^H X_t,\ t \in \mathbb{R}\} \tag{6}$$

One can show that fBm is selfsimilar, and H also is the selfsimilarity exponent of fBm.

The fact that so many notions coincide (in the deterministic setting) for Weierstrass functions and (in the random setting) for fBm, is indeed remarkable, but can be seen as a drawback for modeling, in complex situations where this coincidence does not necessarily exist, and where several parameters would be needed for classification. Concrete examples are supplied by *Lévy processes*, which are another extension of Brownian motion: Indeed, a Lévy process is a random process with independent and stationary increments (therefore Brownian motion can be viewed as the only Lévy process with continuous sample paths), see a sample path in Fig. 5. Lévy processes nowadays play an important role in modeling, in situations for instance where Gaussianity is proved not to hold (such as in finance, see [43, 12] for instance), or where modeling via probability laws with fat tails is mandatory. Dropping the continuity assumption has drastic consequences on the regularity properties of these processes and most Lévy processes display a Hölder exponent, which, far from being constant, becomes an extremely erratic, nowhere continuous function, see [24]. This observation has strong implications in modeling, showing that the parameters which are used to characterize such models cannot be derived on real life data, and thus must be reconsidered. Before following this idea, we start by working out two simple mathematical examples where this situation occurs. We will first follow the intuition of Paul Lévy, who proposed such functions as simple deterministic "toy-examples" displaying some of the key properties of Lévy processes. We will also work out another example, in the setting of measures: We will consider the simplest possible of *multiplicative cascades* which also display such irregularity properties. Note that multiplicative cascades were introduced as turbulence models, but now have a wide range of applications in modeling (see e.g. in fragmentation theory [6]). In order to deal with this second example, we now introduce a notion of pointwise regularity adapted to measures.

Definition 4 *Let μ be a positive Radon measure defined on $\mathbb{R}^d$. Let $x_0 \in \mathbb{R}^d$ and let $\alpha \geq 0$. Let x_0 belong to the support of μ. The local dimension of μ at x_0 is*

$$h_\mu(x_0) = \sup\{\alpha : \mu \in h^\alpha(x_0)\} = \liminf_{r \to 0^+} \frac{\log \mu(B(x_0, r))}{\log\ r}. \tag{7}$$

Note that we use the same notation as for the Hölder exponent, which will lead to no ambiguity in the following. In Section 2 two examples (one function and one measure) of the situation usually met in multifractal analysis are developed: Their pointwise regularity exponent is shown to be extremely erratic, thus pointing to the necessity of other tools to characterize and classify such behaviors. In Section 3, we show how to deal with such situations by describing tools coming both from mathematics and signal processing: They allow to associate to such examples several *multifractal spectra*, which present robustness

properties, both from a mathematical and computational point of view. In Section 4, we reformulate some of the analysis tools introduced in the previous section in terms of wavelet coefficients, and introduce alternative ones that are based on wavelet leaders (i.e. local suprema of wavelet coefficients) and are specific to the wavelet setting. In Section 5, we show how to blend ideas coming from statistics (quantiles) and wavelet leaders to derive new spectra that allow to put into light multifractal spectra that are not concave (which, by construction, cannot be reached using Legendre transform techniques). One drawback of the wavelet leader method is that it can be used only for locally bounded data, and the mathematical results that back it even require some uniform smoothness which excludes discontinuities. In Section 6, we show how to extend the multifractal framework to the L^q setting; examples of random fields where these extensions are required are worked out.

The notions that we introduce are illustrated by selected applications where the method described is shown at work on toy examples, and on real-life data; these illustrations are collected at the end of the paper.

2 Pontwise regularity: Some examples

Our purpose in this section is to show that simply defined functions and measures can display an extremely irregular pointwise regularity exponent. These examples will motivate the introduction of specific tools which are now used in multifractal analysis, and which will be developed in Section 3; they are also typical of the two large classes of multifractal objects which have been considered: The first one falls in the additive setting (which also includes Lévy processes, and random wavelet series), and the second one falls in the multiplicative setting (which contains multiplicative cascades and their generalizations).

2.1 Lévy functions

Let

$$\{x\} = x - [x] - \frac{1}{2},$$

where $[x]$ denotes the integer part of the real number x; $\{x\}$ is the 1-periodic "sawtooth" function, which is nothing but the fractional part of x (recentered, so that its mean vanishes). The Lévy functions, which depend on a parameter $\beta > 0$, are defined by

$$L_\beta = \sum_{j=1}^{\infty} \frac{\{2^j x\}}{2^{\beta j}}. \tag{8}$$

Such functions were proposed by P. Lévy as a toy example of deterministic functions which display some characteristics of Lévy processes; they have a dense set of discontinuities, and are constructed through the accumulation of "compensated jumps": Indeed, if $\beta < 1$, all jumps are negative and the sum of their amplitudes is infinite; the compensation is performed through the linear part of $\{2^n x\}$ which makes the series converge.

Since $|\{x\}| \leq 1$, the series (8) is uniformly convergent towards a 1-periodic function. Since the functions $\{2^j x\}$ are continuous except at dyadic points, L_β is also continuous

except perhaps at dyadic points. Consider now such a point $x_0 = \frac{K}{2^J}$; clearly L_β has right and left limits there and the amplitude of the jump of L_β at x_0 (difference between the right and left limits) is

$$\Delta(x_0) = \sum_{j \geq J} 2^{-\beta j} = C\, 2^{-\beta J}. \tag{9}$$

In order to determine the pointwise regularity of these functions, we will use a general result on functions with a dense set of discontinuities, which yields an upper bound on their Hölder exponent. The *jump of f at a point s* is

$$\Delta_f(s) = \limsup_{x \to s} f(x) - \liminf_{x \to s} f(x).$$

Lemma 1 *Let $f : \mathbb{R}^d \to \mathbb{R}$ be a locally bounded function and let $x_0 \in \mathbb{R}^d$; then*

$$h_f(x_0) \leq \liminf_{s \to x_0} \frac{\log(\Delta_f(s))}{\log(|x_0 - s|)}. \tag{10}$$

Proof: Let

$$L = \liminf_{s \to x_0} \frac{\log(\Delta_f(s))}{\log(|x_0 - s|)},$$

and let P be the Taylor polynomial of f at x_0. Let s be a discontinuity point of f, which we can assume to differ from x_0. Thus $\Delta_f(s) > 0$. Let $\varepsilon = \Delta_f(s)/10$. By definition of $\Delta_f(s)$, there exist x_1 and x_2 which can be chosen arbitrarily close to s, and are such that

$$|f(x_1) - f(x_2)| \geq \Delta_f(s) - \varepsilon$$

and, since P is continuous,

$$|P(x_1 - x_0) - P(x_2 - x_0)| \leq \varepsilon;$$

so that

$$|f(x_1) - P(x_1 - x_0) - (f(x_2) - P(x_2 - x_0))| \geq \Delta_f(s) - 2\varepsilon;$$

therefore, one of the points x_1 or x_2, which we will now denote by $x(s)$, satisfies

$$|f(x(s)) - P(x(s) - x_0)| \geq \Delta_f(s)/3, \tag{11}$$

and, since x_1 and x_2 are arbitrarily close to s, we can assume that

$$\frac{1}{2}|s - x_0| \leq |x(s) - x_0| \leq 2|s - x_0|. \tag{12}$$

Let s_n be a sequence for which the liminf is reached in (10). We obtain a sequence $x(s_n)$ such that

$$\frac{\log(|f(x(s_n)) - P(x(s_n) - x_0)|)}{\log(|x(s_n) - x_0|)} \longrightarrow L. \tag{13}$$

But, for any $\delta > 0$, and for x close enough to x_0,

$$|f(x) - P(x - x_0)| \leq |x - x_0|^{h_f(x_0) - \delta}$$

so that

$$\frac{\log(|f(x)-P(x-x_0)|)}{\log(|x-x_0|)} \geq h_f(x_0)-\delta;$$

and therefore it follows from (13) that $L \geq h_f(x_0)$.

The pointwise regularity of the Lévy functions can be expressed in terms of the *dyadic approximation properties* of the point considered. Let us start by recalling this notion.

Definition 5 *Let $x_0 \in \mathbb{R}$ and $\alpha \geq 1$; x_0 is α-approximable by dyadics if there exists a sequence $(k_n, j_n) \in \mathbb{Z} \times \mathbb{N}$ such that*

$$\left| x_0 - \frac{k_n}{2^{j_n}} \right| \leq \frac{1}{2^{\alpha j_n}}. \tag{14}$$

the dyadic exponent of x_0 is the supremum of all αs such that x_0 is α-approximable by dyadics. This exponent will be denoted by $\alpha(x_0)$.

We will denote by D_α the set of points that are α-approximable by dyadics, and by Δ_α the set of points where the dyadic exponent is exactly α. Note that $D_1 = \mathbb{R}$ and, if $\alpha > \alpha'$, then $D_\alpha \subset D_{\alpha'}$. Furthermore,

$$\Delta_\alpha = \bigcap_{\alpha'<\alpha} D_{\alpha'} - \bigcup_{\alpha'>\alpha} D_{\alpha'}. \tag{15}$$

Recall that the jump of f_β at the points $\frac{K}{2^J}$ (where k is odd) is $b_J = C \cdot 2^{-\beta J}$ (see (9)). It follows then from Definition 5 and Lemma 1 that,

$$\text{if} \quad x_0 \in D_\alpha, \qquad \text{then} \qquad h_f(x_0) \leq \frac{\beta}{\alpha}. \tag{16}$$

Let us now prove regularity outside of D_α.

Lemma 2 *If $x_0 \notin D_\alpha$, then $f_\beta \in C^{\beta/\alpha}(x_0)$.*

Proof: Suppose that $0 < \beta < 1$; let

$$I^\alpha_{j,k} = \left[\frac{k}{2^j} - \frac{1}{2^{\alpha j}}, \frac{k}{2^j} + \frac{1}{2^{\alpha j}} \right].$$

By hypothesis, there exists j_0 such that, if $j \geq j_0$, then x_0 does not belong to any of the intervals $I^\alpha_{j,k}$. The function

$$\sum_{j<j_0} 2^{-\beta j}\{2^j x\}$$

is C^∞ at x_0 so that is does not play a role in the determination of the regularity of f; consider now the remaining part

$$g(x) = \sum_{j \geq j_0} 2^{-\beta j}\{2^j x\}.$$

Let $h > 0$ and j_1 be the integer such that

$$\frac{1}{2^{\alpha(j_1+1)}} \leq h < \frac{1}{2^{\alpha j_1}}.$$

Let us estimate

$$g(x_0+h) - g(x_0) = \sum_{j \geq j_0} 2^{-\beta j} \left(\{2^j(x_0+h)\} - \{2^j x_0\} \right).$$

If $j < j_1$, the function $\{2^j x\}$ has no jump in the interval $[x_0, x_0+h]$, so that

$$2^{-\beta j}\{2^j(x_0+h)\} - 2^{-\beta j}\{2^j x_0\} = 2^{-\beta j} 2^j h;$$

if $j \geq j_1$ we bound the increment $\{2^j(x_0+h)\} - \{2^j x_0\}$ by 2; we finally get

$$|g(x_0+h) - g(x_0)| \leq \sum_{j<j_1} 2^{(1-\beta)j} h + \sum_{j \geq j_1} 2^{-\beta j} \leq C 2^{(1-\beta)j_1} h + C 2^{-\beta j_1}$$

which, using $h \sim 2^{-\alpha j_1}$ and $\alpha \geq 1$, is bounded by $Ch^{\beta/\alpha}$.

The case $h < 0$ is similar; and if $\beta > 1$, we proceed as above, but subtract first from f_β the "linear part"

$$\sum_{j=0}^{\infty} 2^{-\beta j} 2^j x,$$

which is the sum of a convergent series (even if β is large, we need not subtract higher order terms in the Taylor expansion at x_0).

The following result follows from (16), Lemma 2 and (15).

Proposition 1 *The set Δ_α is exactly the set of points where the Hölder exponent of L_β takes the value $H = \beta/\alpha$.*

Note that the sets Δ_α are everywhere dense; indeed, by example, the point

$$x_\alpha = \sum_{j=1}^{\infty} 2^{-j^2} + 2^{-[\alpha j^2]}$$

clearly belongs to Δ_α, which is thus not empty; Δ_α is therefore dense, since it clearly is invariant under any dyadic shift. It follows that, on any arbitrarily small interval, the Hölder exponent of L_β takes all possible values between 0 and β. In particular, it is an everywhere discontinuous function.

2.2 Binomial measures

We now consider another construction, in a multiplicative setting, which also leads to an extremely irregular pointwise regularity function: Binomial measures supported by a Cantor-like fractal set. We first define this support K in a recursive way.

Let l_0 and l_1 be two real numbers such that

$$l_0 > 0, \qquad l_1 > 0 \qquad \text{and} \qquad l_0 + l_1 < 1.$$

If $I = [a, b]$ is an interval, its two children I_0 and I_1 are defined by

$$I_0 = [a, a + (b - a)l_0] \qquad \text{and} \qquad I_1 = [b - (b - a)l_1, b].$$

We start the construction from $I_\emptyset = [0, 1]$. The fundamental intervals of first generation are its two children $I_{(0)} = (I_\emptyset)_0$ and $I_{(1)} = (I_\emptyset)_1$. We define thus iteratively intervals of each generation using the same recipe: Each fundamental interval $I_{(i_1,\cdot,i_k)}$ of generation k has two children $I_{(i_1,\cdots,i_k,0)} = (I_{(i_1,\cdots,i_k)})_0$ and $I_{(i_1,\cdots,i_k,1)} = (I_{(i_1,\cdots,i_k)})_1$. Thus fundamental intervals are indexed by the path that allowed to generate them in the tree.

We now construct the measure. Let m_0 and m_1 be two real numbers such that

$$m_0 > 0, \qquad m_1 > 0 \qquad \text{and} \qquad m_0 + m_1 = 1.$$

We define the measure m on the fundamental intervals by

$$m(I_{(i_1,\cdots,i_k)}) = m_{i_1} \cdots m_{i_k}.$$

It is clearly additive on the fundamental intervals, and it is then extended to all intervals as follows: If I is an open interval of $\mathbb{R}$, we define $m(I) = \sum m(I_l)$, where the sum is taken on all maximal fundamental intervals included in I. One easily checks that m can be extended into a Borel measure on $\mathbb{R}$ supported by the Cantor-type set

$$K = supp(m) = \bigcap_{k\geq 0} \bigcup_{Gen(I_l)=k} I_l, \tag{17}$$

where $Gen(I_l)$ denotes the "generation" of the interval I; the union is thus taken on all fundamental intervals of generation k. It is clear that K is a compact set (as intersection of a decreasing sequence of compacts) with empty interior and without isolated points, i.e. is a *perfect set*. Any point $x \in K$ can be indexed by an infinite sequence $(i_1(x), \cdots, i_k(x), \cdots)$ such that $I_{(i_1(x),\cdots,i_k(x))}$ is the unique fundamental intervals of generation k which contains x; let us denote it by $\mathcal{I}_k(x)$. Thus

$$m(\mathcal{I}_k(x)) = m_{i_1(x)} \cdots m_{i_k(x)} \qquad \text{and} \qquad |\mathcal{I}_k(x)| = l_{i_1(x)} \cdots l_{i_k(x)}. \tag{18}$$

Let us now determine the pointwise regularity exponent of the measure m according to Definition 4. It follows from (18) that an interval of diameter $2|i_k(x)|$ centered at x has measure at least $m(\mathcal{I}_k(x))$. Therefore

$$\forall x \in K \qquad h_m(x) \geq \liminf \frac{\log(m_{i_1(x)} \cdots m_{i_k(x)})}{\log(l_{i_1(x)} \cdots l_{i_k(x)})}. \tag{19}$$

The converse inequality follows from the condition $l_0 + l_1 < 1$ which implies a separation property: If intervals of the same generation are "close" for the euclidean distance, then they are also "close" for the tree distance; the following *doubling property* is a consequence of this remark: If I is a fundamental interval, and if CI denotes the interval of same center as I and C times wider, then

$$\forall C \geq 1, \quad \exists C' > 0 \quad \text{such that, } \forall I \text{ fundamental interval,} \qquad m(CI) \leq C' m(I).$$

The upper bound in (19) follows from this estimate, together with (18) and we therefore obtain the following result.

Proposition 2 *The pointwise regularity exponent of the measure m satisfies*

$$\forall x \in K \qquad h_m(x) = \liminf \frac{\log(m_{i_1(x)} \cdots m_{i_k(x)})}{\log(l_{i_1(x)} \cdots l_{i_k(x)})}. \tag{20}$$

In particular, we note that, if x is such that

$$\frac{1}{k} \sum_{l=1}^{k} i_k(x) \longrightarrow p \qquad \text{when} \qquad k \to \infty,$$

then

$$h_m(x) = \frac{p \log(m_1) + (1-p) \log(m_0)}{p \log(l_1) + (1-p) \log(l_0)}. \tag{21}$$

Therefore, the pointwise exponent of m can take any value in the interval of bounds $(\log m_0)/(\log l_0)$ and $(\log m_1)/(\log l_1)$. The index $p \in [0,1]$ can be used to parametrize the collection of sets of points E_m^H which share the same exponent H; clearly, each of these sets is dense in K.

3 Mathematical notions pertinent in multifractal analysis

The two examples that we worked out in the previous section are by no means exceptional: Many (deterministic or random) mathematical functions or measures share the same property of having an everywhere discontinuous pointwise regularity. In the applications that use such models, it is clear that a direct determination of pointwise regularity exponents would lead to totally unstable algorithms, and therefore be meaningless as a way to derive classification parameters. If one expects such mathematical examples to be relevant in modeling, then, one should associate to them new parameters, which would involve some global, numerically stable, quantities. We now describe the notions which have proved pertinent when dealing with the functions and measures that we considered in the previous section.

3.1 Tools derived from geometric measure theory

In the two examples that we considered, the pointwise Hölder exponent h_f takes all possible values in an interval $[H_{min}, H_{max}]$. Rather than determining the exact behavior of h_f

which, in such cases, is extremely erratic, a more realistic option consists in deriving some quantitative information about h_f, namely, about the size of the sets where f has a given regularity. We therefore consider the *iso-Hölder sets* defined as

$$E_f^H = \{x : \ h_f(x) = H\}.$$

A first problem is to find a way to draw differences between these sets by using a notion of "size" that would, in general, take different values for such sets. The examples supplied by the Lévy functions and the binomial measure allow to test efficiently several natural candidates:

One could first think of the notion of size supplied by the Lebesgue measure. However, in the binomial measure case, since the support of the measure has a vanishing Lebesgue measure, it follows that the size of all sets E_m^H would vanish. More generally, in any situation where the function h_f takes all values in an interval, the Lebesque measure of almost all sets sets E_f^H will necessarily vanish: Indeed, as a consequence of the countable additivity property of the Lebesgue measure, and since the sets E_f^H are disjoint, it follows that the values of H for which $meas(E_f^H) > 0$ (where *meas* denotes the Lebesgue measure) necessarily form a countable set. (Note that the situation usually is even more extreme: One (or none) of the sets E_f^H has full Lebesgue measure, and all others have a vanishing one).

One should therefore use a tool that allows to draw differences between sets of vanishing Lebesgue measure; pertinent notions in such settings are supplied by the different variants of *fractional dimensions*. In the examples that we considered, the sets E_f^H are everywhere dense, either in $\mathbb{R}$ (in the case of the Lévy functions), or in K (in the case of the binomial measure); it follows that the box dimensions of these sets all coincide; therefore the box dimensions will not allow to draw differences between them. The notion which is commonly used in such settings is the *Hausdorff dimension* which we now recall.

Definition 6 *Let $A \subset \mathbb{R}^d$. If $\varepsilon > 0$ and $\delta \in [0, d]$, let*

$$M_\varepsilon^\delta = \inf_R \left(\sum_i |A_i|^\delta \right),$$

where R is an ε-covering of A, i.e. a covering of A by bounded sets $\{A_i\}_{i\in\mathbb{N}}$ of diameters $|A_i| \leq \varepsilon$. (The infimum is therefore taken on all ε-coverings.)

For any $\delta \in [0, d]$, the δ-dimensional outer Hausdorff measure of A is

$$mes_\delta(A) = \lim_{\varepsilon \to 0} M_\varepsilon^\delta.$$

One easily checks that there exists $\delta_0 \in [0, d]$ such that

$$\begin{array}{ll} \forall \delta < \delta_0, & mes_\delta(A) = +\infty \\ \forall \delta > \delta_0, & mes_\delta(A) = 0. \end{array}$$

This critical δ_0 is called the Hausdorff dimension of A, and is denoted by $\dim(A)$.

An important convention, in view of the use of these dimensions in the context supplied by the multifractal formalism, is that, if A is empty, then $\dim(A) = \dim(\emptyset) = -\infty$.

As a typical example of the way to compute Hausdorff dimensions, let us determine $\dim(K)$, where K is defined in (17). This is done by obtaining first an upper bound, and then a lower bound. For the upper bound, we use as ε-covering the set of all intervals at generation n (if n is chosen large enough, the supremum of their lengths can clearly be chosen arbitrarily small). This covering consists in intervals of length $(l_0)^k (l_1)^{n-k}$, and, for a given k, there are $\binom{n}{k}$ of them. Therefore, the quantity $\sum_i |A_i|^\delta$, introduced in Definition 6, takes the value

$$\sum_{k=0}^{2^n} \binom{n}{k} \left((l_0)^k (l_1)^{n-k}\right)^\delta = \left((l_0)^\delta + (l_1)^\delta\right)^n.$$

It follows that, if δ is the unique solution of the equation

$$(l_0)^\delta + (l_1)^\delta = 1, \tag{22}$$

then

$$\dim K \leq \delta.$$

In order to obtain lower bounds for the Hausdorff dimension of a set, it is sufficient to consider, as above, a particular sequence of ε-coverings. On the other hand, deriving lower bounds directly from the definition usually is unfeasible, because it requires to take into account all possible ε-coverings. The following principle (the so called *mass distribution principle*) efficiently replaces the study of all possible ε-coverings by the construction of a particular measure.

Proposition 3 *Let μ be a probability measure supported by a set $A \subset \mathbb{R}^d$. If there exists $\delta \in [0, d]$, $C > 0$ and $\varepsilon > 0$ such that, for any ball B of diameter at most ε,*

$$\mu(B) \leq C|B|^\delta.$$

Then $mes_\delta(A) \geq 1/C$.

Proof: Let $\{B_i\}$ be an arbitrary ε-covering of A. We have

$$1 = \mu(A) = \mu\left(\bigcup B_i\right) \leq \sum \mu(B_i) \leq C \sum |B_i|^\delta.$$

The result follows by passing to the limit when $\varepsilon \to 0$.

The lower bound for the dimension of K follows from Proposition 3 using for μ one of the binomial measures m that we constructed: We pick m_0 and m_1 such that

$$m_0 = (l_0)^\delta \quad \text{and} \quad m_1 = (l_1)^\delta, \tag{23}$$

so that $m_0 + m_1 = 1$. For these particular values, we have exactly, for fundamental intervals

$$m(I) = |I|^\delta,$$

and, a fortiori, for any interval I, $m(I) \leq |I|^\delta$. The mass distribution principle therefore implies then that $\dim K \geq \delta$, so that we have obtained the following result.

Proposition 4 *Let K be the Cantor set K defined by (17). Then*

$$\dim K = \delta,$$

where δ is the solution of (22).

The considerations developed at the beginning of Section 3 motivate the introduction of the following general notion, which plays a central role in multifractal analysis.

Definition 7 *Let f be a function, or a measure, and h_f be its pointwise regularity exponent. The multifractal spectrum of f is the function $d_f(H)$ defined by*

$$d_f(H) = \dim(E_f^H).$$

Note that, in the case of the Hölder exponent, one often refers to $d_f(H)$ as the *Hölder spectrum*, or the *spectrum of singularities* of f. This notion is sometimes used in other contexts: We will see the example of the q-exponent in Section 6.2. It can also be used for exponents of different nature, see for instance [3, 41] where the exponent considered is the size of ergodic averages, or [4, 19] where it is the rate of divergence of Fourier series. In Section 6.1 we will give an easy example in such an alternative setting: We will consider the rate of divergence of the wavelet series of a function in a given Sobolev or Besov space.

It is remarkable that, though Lévy functions or binomial measures have an extremely irregular Hölder exponent, their multifractal spectra are smooth functions, from which the corresponding parameters can be recovered. Indeed, the Hausdorff dimensions of the level sets are obtained by standard computations (see [26] for instance), and one obtains the following results.

Theorem 1 *The multifractal spectrum of the Lévy functions is given by*

$$d(H) = \begin{cases} \beta H & \text{if} \quad H \in [0, 1/\beta] \\ -\infty & \text{else.} \end{cases}$$

Since we determined the Hölder exponent of the Lévy functions at every point, and showed that they only depend on the rate of dyadic approximation, this theorem follows directly from the determination of the Hausdorff dimensions of these sets, which can be found for instance in [24].

Similarly, the multifractal spectrum of the binomial measures can also be deduced from its Hölder exponent and one finds a bell-shaped spectrum, which is a real-analytic concave function supported by the interval whose ends are located at $H^{min} = \log(m_0)/\log(l_0)$ and $H^{max} = \log(m_1)/\log(l_1)$ (assuming that $\log(m_0)/\log(l_0) \leq \log(m_1)/\log(l_1)$). Let us just give a brief idea of the proof: An upper bound for the spectrum is obtained using the multifractal formalism (we give a version of it adapted to the function setting at the end of Section 4.3, but it was first derived in the measure setting, see[18, 49]); lower bounds can be obtained using the mass distribution principle applied to another measure $\tilde{m}$ constructed as m but with different values for m_0 and m_1: One fixes such a couple arbitrarily (but

satisfying $\tilde{m}_0 + \tilde{m}_1 = 1$, and one uses the law of large numbers, which yields that at $\tilde{m}$-almost every point x satisfies that the right-hand side of (20) is a limit, therefore yielding a lower bound of the spectrum for a particular value of H. Changing the values of $\tilde{m}_0$ and $\tilde{m}_1$ makes H take all values in the interval $[H^{min}, H^{max}]$ see [15, 16] for details.

Note however that these results do not suggest a practical way to derive the spectrum, since its definition is based on the Hölder exponent, i.e. on quantities that are ultimately unstable to compute. Therefore, the derivation of parameters used in signal processing follow a different path.

3.2 Tools derived from physics and signal processing

Let us start by a few words concerning the seminal work of Kolmogorov in fully developed turbulence. The stream-wise component of turbulent flow velocity spatial field exhibits very irregular fluctuations over a large range of scales, whose statistical moments furthermore behave, within the so-called inertial scale range, like power laws with respect to the scale h; this velocity measured at a given point is therefore a function of time only, which we denote by $v(t)$. This power-law behavior is written

$$\int |v(t+h) - v(t)|^p dt \quad \sim \quad h^{\eta(p)}. \tag{24}$$

This statement means that the function $\eta(p)$ can be determined as a limit when $h \to 0$ on a log-log plot; it is called the *scaling function* of the velocity v. Characterization and understanding of the observed scaling properties play a central role in the theoretical description of turbulence, and Kolmogorov in 1941 expected a linear scaling function for turbulent flows [36]: $\eta(p) = p/3$. This prediction has been refined by Obukhov and Kolmogorov in 1962 who predicted a quadratic behavior of the scaling exponents [38, 48]. The non-linear behavior of $\eta(p)$ was confirmed by various experimental results and other models have been proposed leading to different scaling functions $\eta(p)$.

Definition 8 *Let $f : \mathbb{R}^d \to \mathbb{R}$. The scaling function of f (see [37]) is the function $\eta_f(p)$ defined by*

$$\forall p \geq 1, \qquad \eta_f(p) = \liminf_{|h| \to 0} \frac{\log \left(\int |f(x+h) - f(x)|^p dx \right)}{\log |h|}. \tag{25}$$

Note that, if data are smooth (i.e., if one obtains that $\eta_f(p) \geq p$), then one has to use differences of order 2 (or more) in (25) in order to define correctly the scaling function.

In applications, multifractal analysis consists in the determination of scaling functions (variants to the original proposition of Kolmogorov will be considered later). Such scaling functions can then be involved into classification or model selection procedures.

An obvious advantage of the use of the scaling function $\eta_f(p)$ is that its dependence in p can take a large variety of forms, hence providing versatility in adjustment of models to data. Therefore multifractal analysis, being based on a whole function rather than on a single exponent, yields much richer tools for classification or model selection. The

scaling function however satisfies a few constraints, for example, it has to be a concave non-decreasing function (cf. e.g., [32, 51]).

Later refinements and extensions of the wavelet scaling function were an indirect consequence of its interpretation in terms of fractal dimensions of Hölder singularities, proposed by G. Parisi and U. Frisch in their seminal paper [49], and which relates this point of view with the one developed above in Section 3.1.

The formula proposed by Parisi and Frisch is the following relationship between the scaling function and the multifractal spectrum

$$\dim(E_f^H) = \inf_p \left(d + Hp - \eta_f(p)\right), \tag{26}$$

see [49]. Though the remarkable intuition which lies behind this formula proved extremely fruitful (see for instance [49] for a heuristic justification based on statistical physics ideas) we now know that it needs to be pushed further in order to be completely effective; indeed many natural processes used in signal or image modeling, such as Brownian motion, or fBm, are counterexamples, see [40]; additionally, the only mathematical result relating the spectrum of singularities and the scaling function in all generality is very partial, see [22, 26].

Theorem 2 *Let $f : \mathbb{R}^d \to \mathbb{R}$ be such that $H_f^{min} > 0$. Define p_0 by the condition:*

$$\eta_f(p_0) = dp_0;$$

then

$$\dim(E_f^H) \leq \inf_{p>p_0} (d + Hp - \eta_f(p)). \tag{27}$$

The motivation for introducing new scaling functions has been to obtain alternative ones for which (26) would hold with some generality, and also for which the upper bound stated in Theorem 2 would be sharper. An additional side advantage is that such new scaling functions also yield new classification and model selection parameters. The best results following this benchmark have been obtained through the use of *wavelet based scaling functions* in the construction of such scaling functions. We now turn towards this more recent approach.

4 Wavelet based scaling functions

4.1 Wavelet bases

Orthonormal wavelet bases are a privileged tool to study multifractal functions for several reasons that will be made explicit. In this subsection, we only recall the properties of orthonormal wavelet bases that will be useful in the sequel. We refer the reader for instance to [13, 42, 45] for detailed expositions.

Orthonormal wavelet bases on $\mathbb{R}^d$ are of the following form. There exists a function $\varphi(x)$ and $2^d - 1$ functions $\psi^{(i)}$ with the properties: The functions $\varphi(x-k)$ $(k \in \mathbb{Z}^d)$ and the $2^{dj/2}\psi^{(i)}(2^j x - k)$ $(k \in \mathbb{Z}^d,\ j \geq 0)$ form an orthonormal wavelet basis of $L^2(\mathbb{R}^d)$. This

basis is r-smooth if φ and the $\psi^{(i)}$ are C^r and if the $\partial^\alpha\varphi$, and the $\partial^\alpha\psi^{(i)}$, for $|\alpha| \leq r$, have fast decay.

Therefore,

$$\forall f \in L^2, \qquad f(x) = \sum_{k\in\mathbb{Z}^d} c_k^{(0)}\varphi(x-k) + \sum_{j=0}^{\infty}\sum_{k\in\mathbb{Z}^d}\sum_i c_{j,k}^i\psi^{(i)}(2^j x - k); \tag{28}$$

the $c_{j,k}^i$ and $c_k^{(0)}$ are called the wavelet coefficients of f and given by

$$c_{j,k}^i = 2^{dj}\int_{\mathbb{R}^d} f(x)\psi^{(i)}(2^j x - k)dx, \quad \text{and} \quad c_k^{(0)} = \int_{\mathbb{R}^d} f(x)\varphi(x-k)dx.$$

Note that the computation of these coefficients makes sense with very little assumption on f (a wide mathematical setting is supplied by tempered distributions). A natural setting for functions is given by the space L^1 *with slow growth*, which is defined as follows.

Definition 9 *Let f be a locally integrable function defined over $\mathbb{R}^d$; f belongs to $L^1_{SG}(\mathbb{R}^d)$ if*

$$\exists C, N > 0 \quad \textit{such that} \qquad \int_{\mathbb{R}^d} |f(x)|(1+|x|)^{-N}dx \leq C.$$

The wavelet expansion of a function $f \in L^1_{SG}(\mathbb{R}^d)$ converges a.e.; in particular at Lebesgue points, it converges towards the Lebesgue value

$$\lim_{r\to 0}\frac{1}{Vol(B(x_0,r))}\int_{B(x_0,r)} f(x)dx.$$

Furthermore, let $C_{SG}(\mathbb{R}^d)$ be the set of locally bounded and continuous functions which satisfy

$$\exists C, N > 0 : \qquad |f(x)| \leq C(1+|x|)^N.$$

Then, if $f \in C_{SG}(\mathbb{R}^d)$, its wavelet expansion converges uniformly on compact sets.

We will use more compact notations for indexing wavelets:

- Instead of using the three indices (i,j,k), we will use dyadic cubes. Since i takes 2^d-1 values, we can assume that it takes values in $\{0,1\}^d - (0,\ldots,0)$; we introduce

$$\lambda\ (=\lambda(i,j,k)) \ = \frac{k}{2^j} + \frac{i}{2^{j+1}} + \left[0, \frac{1}{2^{j+1}}\right)^d,$$

  and, accordingly: $c_\lambda = c_{j,k}^i$ and $\psi_\lambda(x) = \psi^{(i)}(2^j x - k)$. Indexing by dyadic cubes will be useful in the sequel because the cube λ indicates the localization of the corresponding wavelet. Note that this indexing is one to one: If $(i,j,k) \neq (i',j',k')$, then $\lambda(i,j,k) \neq \lambda(i',j',k')$.

- In order to have a common notation for wavelets and functions φ, when $j = 0$, we note $\psi_\lambda(x)$ the function $\varphi(x - k)$ (where λ is, in this case, the unit cube shifted by k), and by c_λ the corresponding coefficient.

- Finally, Λ_j will denote the set of dyadic cubes λ which index a wavelet of scale j, i.e., wavelets of the form $\psi_\lambda(x) = \psi^{(i)}(2^j x - k)$ (note that Λ_j is a subset of the dyadic cubes of side 2^{j+1}).

Note that the wavelet ψ_λ is essentially localized near the cube λ; more precisely, when the wavelets are compactly supported, then, $\exists C > 0$ such that when $\psi^{(i)} \subset [-C/2, C/2]^d$ then $\psi_\lambda \subset 2^{-j}k + 2^{-j}[-C/2, C/2] \subset 2C\lambda$.

A key property of wavelets, which plays a central role in applications, is that they are *universal bases* of many function spaces. Let us explain this notion. Recall that a quasi-norm satisfies the requirements of a norm except for the triangular inequality which is replaced by the weaker condition

$$\exists C > 0, \ \ \forall x, y \in E, \ \ \| \, x + y \, \| \leq C(\| \, x \, \| + \| \, y \, \|).$$

A quasi-Banach space is a complete topological vector space endowed with a quasi-norm. Typical examples are the real Hardy spaces H^p, and the Besov spaces $B^{s,p}_p$ with $p \in (0, 1)$ or $q \in (0, 1)$.

Definition 10 *Let E be either a Banach space, or a quasi-Banach space; a sequence e_n is an unconditional basis of E if:*

- *Each vector $f \in E$ has a unique expansion $f = \sum a_n e_n$, where the series converges in E.*

- *$\exists C \ \ \forall (a_n), \ \ \forall (\varepsilon_n)$ such that $|\varepsilon_n| \leq 1$, $\| \sum \varepsilon_n a_n e_n \| \leq C \, \| \sum a_n e_n \|$.*

Note that one can slightly weaken the first condition in cases where E is not separable but is the dual of a separable space F, in which case one replaces strong convergence to f by a weak* convergence; this is typically the case for the spaces C^α. The second condition has the following consequence: The norm (or semi-norm) of a function (or a distribution) $f \in E$ is equivalent to a condition on the sequence $|a_n|$; we denote this condition by $\mathcal{C}_E$. Suppose now that the sequence (e_n) also is an orthonormal (or a bi-orthogonal) basis of L^2. Then the coefficients a_n are defined by $a_n = \langle f | e_n \rangle$ (or $a_n = \langle f | f_n \rangle$ in the bi-orthogonal case, in which case (f_n) will be the dual basis), and these coefficients can be well defined even if f does not belong to L^2. It is typically the case for wavelet bases: If the wavelet basis belongs to the Schwartz class, then the coefficients are well defined as soon as f is a tempered distribution (and if the wavelets have a limited regularity, it will still be the case if f is a distribution of limited order). Assume now that f is a tempered distribution. In many situations, one would like to have a criterium which allows to decide whether $f \in E$. Note however that the fact that the coefficients of f satisfy the condition $\mathcal{C}_E$ is, in general,

not sufficient to imply that $f \in E$: A simple counter-example is supplied by wavelet bases where one picks for basis the collection of functions

$$2^{dj/2}\psi^{(i)}(2^j x - k) \quad \text{with} \qquad k \in \mathbb{Z}^d, \;\; j \in \mathbb{Z},$$

which also forms an orthonormal basis of L^2. However the constant function $f(x) = 1$ has all vanishing coefficients in this basis (because $\forall i$, $\int \psi^{(i)}(x)\, dx = 0$) and does not belong to L^2. We thus see that a stronger requirement on the basis is needed; hence the following definition.

Definition 11 *Let E be either a Banach space, or a quasi-Banach space of tempered distributions defined on $\mathbb{R}^d$. Let (e_n) be an unconditional basis of E, whose elements belong to the Schwartz class $\mathcal{S}$, and which also is an orthonormal basis of E. The sequence (e_n) is a universal basis of E if the following property holds: $\forall f \in \mathcal{S}'$, if the sequence of coefficients $a_n = \langle f | e_n \rangle$ (defined in the sense of the $(\mathcal{S}, \mathcal{S}')$ duality)satisfies $\mathcal{C}_E$, then $f \in E$.*

This definition easily extends to the bi-orthogonal case and to settings of bases of limited regularity (in which case one only deals with distributions of limited order). An important result is the fact that *wavelet bases, as defined in the expansion (28), are universal bases* of many function spaces, such as the Sobolev spaces $L^{p,s}$ for $p \in (1, \infty)$ or the Besov spaces $B_p^{s,q}$. This can be verified using the fact that they are unconditional bases of these spaces (see [45]) and then using a localization argument: One picks a C^∞ compactly supported function ω which is equal to 1 on $B(0, R)$ and one considers $f\omega$; its wavelet coefficients are locally uniformly close to those of f, because of the uniform localization of the elements of the basis (which does not hold for wavelets with $j \in \mathbb{Z}$). The result then follows by passing to the limit when $R \to +\infty$.

4.2 The wavelet scaling function

An important property of wavelet expansions is that many function spaces have a simple characterization by conditions bearing on wavelet coefficients. This property has a direct consequence on the practical determination of the scaling function.

The function space interpretation of the scaling function $\eta_f(p)$ can be obtained through the use of the Sobolev spaces $L^{p,s}(\mathbb{R}^d)$. Let $p \geq 1$ and $s \in \mathbb{R}$; recall that a tempered distribution f belongs to $L^{p,s}(\mathbb{R}^d)$ if its Fourier transform $\hat{f}$ is a function satisfying the following property: If $\hat{g}(\xi) = (1 + |\xi|^2)^{s/2}\hat{g}(\xi)$, then $g \in L^p$. The different variants in the definition of Sobolev spaces and the embeddings between them imply that

$$\eta_f(p) = p \, \sup \left\{ \, s : \quad f \in L^{p,s}_{loc} \, \right\}. \tag{29}$$

Let

$$S_f(p, j) = 2^{-dj} \sum_{\lambda \in \Lambda_j} |c_\lambda|^p. \tag{30}$$

The wavelet characterization of Sobolev spaces implies that the Kolmogorov scaling function can be re-expressed as (cf. [22])

$$\forall p \geq 1, \qquad \eta_f(p) = \liminf_{j \to +\infty} \frac{\log\left(S_f(p,j)\right)}{\log(2^{-j})}. \tag{31}$$

Note that this result only holds if the wavelets used are smooth enough. The rule of thumb is that wavelets should be smoother and have more vanishing moments than the regularity index appearing in the definition of the function space. In the following we will never specify the required smoothness, and always assume that *smooth enough wavelets* are used (the minimal regularity required being always easy to infer).

This characterization, which, again, yields the scaling function through linear regressions in log-log plots, has several advantages when compared to the earlier version (25). First, (31) allows to extend the scaling function to the range $0 < p \leq 1$ (in which case the function space interpretation requires the use of Besov spaces, see [27, 30] and references therein). We will call this extension supplied by (31) the *wavelet scaling function*, and we will keep the same notation. The wavelet scaling function can be used for classification, but it also meets specific purposes: Indeed, its function space interpretation implies that the values it takes for particular values of p carry a key information in several circumstances. Let us give a few examples.

We start by an application motivated by image processing. A function f belongs to the space BV, i.e., has *bounded variation*, if its gradient, taken in the sense of distributions, is a finite (signed) measure. A standard assumption in image processing is that real-world images can be modeled as the sum of a function $u \in BV$ which models the *cartoon part*, and another term v which accounts for the noise and texture parts (for instance, the first "$u+v$ model", introduced by Rudin, Osher and Fatemi in 1992 ([39]) assume that $v \in L^2$). The BV model is motivated by the fact that if an image is composed of smooth parts separated by contours which are piecewise smooth curves, then its gradient will be the sum of a smooth function (the gradient of the image inside the smooth parts) and Dirac masses along the edges, which are typical finite measures. On the opposite, characteristic functions of domains with fractal boundaries usually do not belong to BV, see Fig. 2 for an illustration. Therefore, a natural question in order to validate such models is to determine whether an image (or a portion of an image) actually belongs to the space BV, or to the space L^2, or not. This question can be given a sharp answer using the wavelet scaling function. Indeed, the values taken by the wavelet scaling function at $p = 1$ and $p = 2$ allow practitioners to determine if data belong to BV or L^2:

- If $\eta_f(1) > 1$, then $f \in BV$, and if $\eta_f(1) < 1$, then $f \notin BV$
- If $\eta_f(2) > 0$, then $f \in L^2$ and if $\eta_f(2) < 0$, then $f \notin L^2$.

Thus wavelet techniques allow to check if the assumptions which are made, in certain denoising algorithms relying on then $u + v$ model, are valid.

Examples of synthetic images are shown in Fig. 2, together with the corresponding measures of $\eta_f(1)$ and $\eta_f(2)$. The image consisting of a simple discontinuity along a circle and

no texture, (i.e., a typical *cartoon* part of the image in the $u+v$ decomposition) is in BV. This is in accordance with the value found for $\eta_f(1)$ which is close to 1. The image of textures or discontinuities existing on a complicated support (such as the Von Koch snowflake) are not in BV., and the function η is found numerically in very good accordance with the theoretical value, and yields that this characteristic function does not belong to BV.; Y. Gousseau and J.-M. Morel were the first authors to raise the question of finding statistical tests to verify if natural images belong to BV [17]. We finally show the scaling function of the characteristic function of the Mandelbrot set, which is not known theoretically. An interesting question would be to relate the values of the scaling function with some geometric properties of this set, and see if the values computed can help to confirm some conjectures concerning this set. Note that an advantage of the use of the scaling function is that the answer is not given by a yes/no procedure, but tells how far the data are from belonging to the space considered, allowing for the possible use of error bars.

Another illustration that we show is the numerical determination of the scaling function of the uniform measure on the Sierpinski gasket, see Fig. 3. Let us sketch how it can be derived mathematically: Each triangle of width 2^{-j} has a measure 3^{-j}. Therefore, if the support of a wavelet intersects Sierpinski gasket, the corresponding wavelet coefficient will be of size $\sim 2^{2j}3^{-j}$, and there are $\sim 3^j$ such wavelet coefficients. It follows that

$$S_f(p,j) \sim 2^{-2j}3^j \left(2^{2j}3^{-j}\right)^p,$$

so that

$$\eta_m(p) = (p-1)(\delta-2) \qquad \text{where} \qquad \delta = \frac{\log 3}{\log 2}.$$

Another motivation for function space interpretations of scaling functions will be provided in Section 4.4, for the estimation of the p-variation.

The computation of the wavelet scaling function also is a prerequisite in some variants of multifractal analysis. This will be exposed in Section 6.2, where variants of multifractal analysis are investigated: They requires that $f \in L^q_{loc}$, a requirement that can be verified by checking that $\eta_f(q) > 0$.

4.3 Wavelet leaders

At the end of Section 3, we mentioned the importance of looking for an "improved" scaling function, i.e., one such that (26) would have a wider range of validity, and for which the upper bound supplied by Theorem 2 would be sharper. This led to the construction of the *wavelet leader scaling function*, which we now recall. The "basic ingredients" in this formula are no more wavelet coefficients, but wavelet leaders, i.e., local suprema of wavelet coefficients. The reason is that pointwise smoothness can be expressed much more simply in terms of wavelet leaders than in terms of wavelet coefficients.

Definition 12 *Let $f \in C_{SG}(\mathbb{R}^d)$, and let λ be a dyadic cube; 3λ will denote the cube of same center and three times wider. If f is a bounded function, the wavelet leaders of f are*

the quantities

$$d_\lambda = \sup_{\lambda' \subset 3\lambda} |c_{\lambda'}|$$

Note that it is important to require f to belong to $C_{SG}(\mathbb{R}^d)$; otherwise, the wavelet leaders of f can be infinite; therefore checking that $H_f^{min} > 0$ is a prerequisite for the whole method.

The reason for introducing wavelet leaders is that they give an information on the pointwise Hölder regularity of the function. Indeed, let $x_0 \in \mathbb{R}^d$, and denote by $\lambda_j(x_0)$ the dyadic cube of width 2^{-j} which contains x_0. If $H_f^{min} > 0$, then

$$h_f(x_0) = \liminf_{j \to +\infty} \frac{\log\left(d_{\lambda_j(x_0)}\right)}{\log(2^{-j})}. \tag{32}$$

(see [26] and references therein). Therefore, constructing a scaling function with the help of wavelet leaders is a way to incorporate pointwise smoothness information. It is therefore natural to expect that (27) will be improved when using such a scaling function instead of $\eta_f(p)$.

The *leader scaling function* is defined by

$$\forall p \in \mathbb{R}, \qquad \zeta_f(p) = \liminf_{j \to +\infty} \frac{\log\left(2^{-dj} \sum_{\lambda \in \Lambda_j} (d_\lambda)^p\right)}{\log(2^{-j})}. \tag{33}$$

Here again, this mathematical definition should be interpreted as meaning that

$$\sum_{\lambda \in \Lambda_j} (d_\lambda)^p \sim 2^{-\zeta_f(p) j}.$$

An important property of the leader scaling function is that it is "well defined" for all values of p. By "well defined", we mean that it has the following robustness properties if the wavelets belong to the Schwartz class (partial results still hold otherwise), see [31, 26]:

- ζ_f is independent of the (smooth enough) wavelet basis.
- ζ_f is invariant under the addition of a C^∞ perturbation.
- ζ_f is invariant under a C^∞ change of variable.

The *leader spectrum* of f is defined through a Legendre transform of the leader scaling function as follows

$$L_f(H) = \inf_{p \in \mathbb{R}} \left(d + Hp - \zeta_f(p)\right). \tag{34}$$

The following result of [26] shows the improvement obtained when using wavelet leaders.

Theorem 3 *If $H_f^{min} > 0$, then, $\forall H$, $\quad d_f(H) \leq L_f(H)$.*

Not that, if compared with (27), the upper bound is sharpened since one can show that, on one hand $\eta_f(p) = \zeta_f(p)$ if $p > p_0$, and, on the other hand, the infimum in (34) is taken for all $p \in \mathbb{R}$. Furthermore, equality holds for large classes of models used in signal and image processing, such as fBm, lacunary and random wavelet series, cascade models,... see [5, 6, 32, 33, 31] and references therein; when this is the case, we will say that the *multifractal formalism based on wavelet leaders holds*.

Note that ζ_f is a concave function by construction (a direct consequence of Hölder inequality, see [31]), so that there is no loss of information when considering its Legendre transform. However, by definition, this Legendre transform is always concave, so that, if $d_f(H)$ is not a concave function (and it has no reason to be in general), then, the multifractal formalism won't hold. In such cases, we can expect however a weaker result, namely that the concave hull of $d_f(H)$ is recovered by (34) (recall that the concave hull of a function f is the smallest concave function g which is everywhere larger than f). We will then say that the *weak multifractal formalism* holds.

4.4 Estimation of the p-oscillation and p-variation

Another motivation for the practical computation of scaling functions is supplied by the *finite quadratic variation* hypothesis in finance. There exist several slightly different formulations of such conditions depending on the notion of p-variation that is used. We start by recalling these notions and their relationships.

Let A be a convex subset of $\mathbb{R}^d$; the (first order) oscillation of f on A is

$$osc(f, A) = \sup_{x \in A} f(x) - \inf_{x \in A} f(x) \tag{35}$$

The p-oscillation of f at scale j is defined by

$$Osc_p(f, j) = \sum_{\lambda \in \Lambda_j} (osc(f, 3\lambda))^p .$$

For analyzing smooth functions, the definition of the p-oscillation using (35), which takes into account only first order differences, has to be modified; one uses instead n-th order differences, which are defined by induction as follows: At step 1,

$$\Delta_f^1(t, h) = f(t + h) - f(t)$$

and, for $n \geq 2$,

$$\Delta_f^n(t, h) = \Delta_f^{n-1}(t + h, h) - \Delta_f^{n-1}(t, h);$$

and, in the definition of the p-oscillation, (35) is replaced by

$$osc(f, A) = \sup_{[t, t+nh] \subset A} |\Delta_f^n(t, h)|. \tag{36}$$

The p-variation spaces $V^{s,p}(\mathbb{R}^d)$ are defined by the condition

$$\exists C \; \forall j \geq 0, \qquad \left(2^{-dj} \sum_{\lambda \in \Lambda_j} (osc(f, 3\lambda))^p \right)^{1/p} \leq C 2^{-js} \tag{37}$$

(one chooses here in the definition of the oscillation a fixed n larger than $[s]+1$), see [30]. This scale of function spaces allows to introduce the *p-oscillation exponent*

$$\omega_p(f) = \liminf_{j\to\infty} \frac{\log\left(Osc_p(f,j)\right)}{\log(2^{-j})}.$$

Of course $\omega_p(f) = \sup\{s : f \in V^{s/p,p}\}$. The following result of [23, 28] allows to relate this quantity with the leader scaling function.

Theorem 4 *Let f be such that $H_f^{min} > 0$. Then*

$$\forall p \geq 1, \qquad \omega_p(f) = \zeta_f(p).$$

This theorem yields a practical criterium in order to determine whether the p-oscillation of a function is bounded, which can be used with $p = 2$ (and $d = 1$) in the context of finance data:

- If $\zeta_f(p) > d$, then f has a finite p-oscillation.
- If $\zeta_f(p) < d$, then the p-oscillation of f is unbounded.

Note that the leader scaling function also has another role for $p = 1$: It yields the upper box dimension of the graph of f, see [23]: Let $f : \mathbb{R}^d \to \mathbb{R}$ be a compactly supported function such that $H_f^{min} > 0$; then

$$\overline{dim_b}(Graph(f)) = \sup\left(d, d+1-\zeta_f(1)\right).$$

In one variable, these notions are closely related with the p-variation. Recall that a function $f : [0,1] \to \mathbb{R}$ has a bounded p-variation (we write $f \in V^p$) if the following condition holds: There exists $C > 0$ such that, for any arbitrary subdivision of $[0,1]$ $0 \leq t_1 \leq \cdots \leq t_n \leq 1$,

$$\sum_{i=1}^{n-1} |f(t_{i+1}) - f(t_i)|^p \leq C. \tag{38}$$

Clearly, if f belongs to V^p,

$$A_j = \sum_k |\sup_{\lambda_{j,k}} f(x) - \inf_{\lambda_{j,k}} f(x)|^p \leq C \tag{39}$$

(take for subdivision the points where the supremum and the infimum are attained in each dyadic interval of length 2^{-j}). Thus, if f belongs to V_p, then $f \in V^{0,p}$. Conversely, suppose that there exists a positive ε such that $f \in V^{\varepsilon,p}$. It follows that for any j, $A_j \leq C\, 2^{-\varepsilon j}$; let $(t_i)_{i=1,\ldots,n}$ be a finite subdivision; let J be such that each interval of length 2^{-J} contains at most two points of the subdivision. The sum of all indexes i such that t_{i+1} and t_i belong to the same dyadic interval of length 2^{-J} is bounded by A_J. The sum of all indexes i such that t_{i+1} and t_i belong to the same dyadic interval of length 2^{-J-1} but do not belong to

the same dyadic interval of length 2^{-J} is bounded by $A_{J-1},...$ thus, since $A_j \leq C\ 2^{-\varepsilon j}$, $f \in V_p$. Hence the following imbeddings

$$\forall \varepsilon > 0 \quad V^{\varepsilon,p} \hookrightarrow V_p \hookrightarrow V^{0,p}, \tag{40}$$

which also allow to estimate the p-variation from the knowledge of the leader scaling function.

In one space variable, the following alternative definition is also used for the p-variation.

Definition 13 *Let $f_a(x) = f(x-a)$. The function f has a finite regular p-variation if*

$$\exists C,\ \ \forall a, h \in]0,1], \qquad \sum_n |f_a((n+1)h) - f_a(nh)|^p \leq C.$$

Comparing the different notions of oscillation and variation allows to derive the following practical criterium, see [23, 28, 30]:

- If $H_f^{min} < 0$ or if $\eta_f(p) < 1$, then the regular p-variation of f is not bounded.
- If $\eta_f(p) > 1$, then $\zeta_f(p) = \eta_f(p)$ and the p-oscillation of f is bounded (and its regular p-variation is therefore also bounded)

These theoretical results have found practical applications in [2], where it is shown that the US Dollar vs. Euro, considered on 11 consecutive years can be considered as a function of bounded quadratic variations.

5 The curse of concavity

Many deterministic functions and stochastic processes yield examples of non-concave multifractal spectra. It is easy to construct toy-examples of such functions, typically by superposing, or concatenating, functions with known spectra. Note that, in the case of concatenation of two functions: f_1 supported by an interval I_1 and f_2 supported by an interval I_2, these intervals being such that $I_2 \cap I_2 = \emptyset$, then the whole spectrum of $f = f_1 + f_2$ on $I_2 \cup I_2$ will be the supremum of the two spectra, hence, in general, not a concave spectrum. In such situations, if the multifractal formalism holds separately for f_1 and f_2, yielding the spectra D_1 and D_2, respectively, then, when applied to f, we expect the multifractal formalism to yield the concave hull of $\sup(D_1, D_2)$. Note however that in such situations one could solve the problem and obtain the correct spectrum by localizing the analysis, see [7] for a corresponding mathematical development (see also [8] for an example of a Markov process whose multifractal spectrum changes with time). We now start by reviewing a few instructive examples of situations naturally leading to non-concave spectra where where such a localization is not feasible.

5.1 Examples of non-concave spectra

A first example is supplied by the first function which was proved to be a multifractal function, i.e. *Riemann's non-differentiable function*

$$\mathcal{R}(x) = \sum_{1}^{\infty} \frac{\sin(\pi n^2 x)}{n^2},$$

see [21] and the top of Fig. 1 for its graph. The multifractal spectrum of $\mathcal{R}$ is given by

$$\begin{array}{lll} D(H) & = 4H - 2 & \text{if } H \in [\frac{1}{2}, \frac{3}{4}] \\ & = 0 & \text{if } H = 3/2 \\ & = -\infty & \text{else,} \end{array} \tag{41}$$

see bottom of Fig. 1 for this spectrum. The point H = 3/2 in the spectrum corresponds to the differentiability points found by J. Gerver: They are the rational numbers of the form $(2p+1)/(2q+1)$ (and form a set of vanishing Hausdorff dimension). However, since these points are dense, it is clear that one could not "separate" the two parts of the spectrum corresponding to $H \leq 3/4$ and $H = 3/2$ by using a localized multifractal analysis. It is actually an easy consequence of the computations performed in [21] that the Riemann function is an *homogeneous multifractal*, i.e. on any interval of non-empty interior, the spectrum is given by (41). The Riemann function is plotted in Figure 1, together with its theoretical multifractal spectrum and its numerically computed Legendre spectrum, which, predictably is the concave hull of the theoretical multifractal spectrum. This illustration puts in light an ambiguity of the Legendre spectrum: In the case of the Riemann function, it yields two segments of straight lines, and one cannot decide whether the actual multifractal spectrum is this whole straight line (as it is the case of the left hand-side) or only the two end-points (as it it the case for the right hand side)... or some intermediate behavior! The purpose of the introduction of the *Quantile Leader Spectrum* in Section 5.3 is to propose a numerically stable method in order to lift this type of ambiguities.

Other examples of homogeneous multifractal functions are supplied by an important subclass of the Lévy processes. Recall that a *pure jump function* is a function whose derivative can be written as the sum of a constant term and a series of Dirac masses. An important classification result of P. Lévy states that any Lévy process X can be decomposed as the sum of a (possibly vanishing) Brownian part, and an independent pure jump process. This pure jump process can itself be decomposed as a series of compound compensated Poisson processes. We now exclude the case where the sum of the series is itself a compound compensated Poisson process (therefore assuming that the Lévy measure associated with X is infinite); then X will have a dense set of discontinuities. The multifractal properties of the sample paths of X are governed by an index $\beta \in [0, 2]$, the *Blumenthal and Getoor lower index* which describes how the Lévy measure diverges at the origin, and also by the presence or absence of a Brownian component. If X has no Brownian component, then with probability one, its multifractal spectrum is given by

$$d_X(H) = \begin{cases} \beta H & \text{if } H \in [0, 1/\beta] \\ -\infty & \text{else.} \end{cases}$$

On the other hand, if X has a Brownian component (which can be achieved by adding an independent Brownian motion to the previous process), then with probability one, its multifractal spectrum is given by

$$d_X(H) = \begin{cases} \beta H & \text{if } H \in [0, 1/2) \\ 1 & \text{if } H = 1/2 \\ -\infty & \text{else} \end{cases}$$

(see [24], and also [14] for extensions to random fields). Note that, in this case too, the restriction of this process to any interval of non-empty interior yields the same spectrum, so that the spectrum cannot be "split" into several concave ones by localization. The sample path of such a Lévy process with Brownian component is showed at the top of Fig. 5; below, we show its theoretical multifractal spectrum together with its Legendre spectrum numerically computed form sample paths. Here again, the Legendre spectrum is the concave hull of the theoretical multifractal spectrum, thus missing to show the non-concavity in the increasing part of the spectrum.

A slightly different situation is supplied by the square of a fBm as considered in [2]. As already mentioned, the sample paths of fBm $B_H(t)$, are not multifractal: The Hölder exponent is everywhere equal to H. Let us now consider its square $Y_H(t) = (B_H(t))^2$: On one hand, at points where the sample path of fBm does not vanish, the action of the mapping $x \to x^2$ locally acts as a C^∞ diffeomorphism, and the pointwise regularity is therefore preserved. On other hand, consider now the (random) set A of points where fBm vanishes. The uniform modulus of continuity of fBm implies that a.s., for s small enough,

$$\sup_t |B_H(t+s) - B_H(t)| \leq C|s|^H \sqrt{\log(1/|s|)}.$$

Therefore, if B_H vanishes at t, then $B_H(t+s)^2 \leq C|s|^{2H} \log(1/|s|)$, so that $h_Y(t) \geq 2H$. The converse estimate follows from the fact that, for every t,

$$\limsup_{s \to 0} \frac{|B_H(t+s) - B_H(t)|}{|s|^H} \geq 1,$$

so that, if $B_H(t) = 0$, then

$$\limsup_{s \to 0} \frac{(B_H(t+s))^2}{|s|^{2H}} \geq 1,$$

so that $h_Y(t) \leq 2H$. Thus, at vanishing points of B_H, the action of the square is to shift the Hölder exponent from $h = H$ to $h = 2H$. This set of points has been the subject of investigations by probabilists (cf. the flourishing literature on the *local time* either of Brownian motion, or of fBm); in particular, it is known to be a fractal set of dimension $1 - H$, cf [47]. It follows that the multifractal spectrum of $Y_H(t)$ is given by

$$D_{Y_H}(h) = \begin{cases} 1 & \text{if} \quad h = H, \\ 1 - H & \text{if} \quad h = 2H, \\ -\infty & \text{elsewhere.} \end{cases} \tag{42}$$

Because of Theorem 3, the multifractal formalism will yield for Legendre spectrum a function above the segment of endpoints $(H, 1)$ and $(2H, 1-H)$ thus suggesting to practitioners that $Y_H(t)$ is a *fully* multifractal function whose spectrum is supported (at least) by the interval $[H, 2H]$. This is illustrated in Figure 4, where a smaple path is shown (top), and the thoretical multifarctal spectrum is also shown (bottow) together with the Legendre spectrum numerically computed from sample paths.

Note also that the sample paths of fBm are *homogeneous* (i.e., the spectrum measured from a restriction of the sample path to any interval (a, b) on the real line is the same as that corresponding to the whole one); this is no longer true for $(B_H)^2$: The spectrum measured on a restricted interval (a, b) will vary depending on whether the interval includes or not a point where B_H vanishes. However, since the set A is of empty interior, it follows that we cannot localize $(B_H)^2$ in regions where the exponent $2H$ only would be present.

5.2 The Large Deviation Leader Spectrum

After this review of a few characteristic examples, we now come back to the general situation supplied by non-concave spectra. In such situations, as already mentioned, one can expect, at best, the weak formalism to hold (see Figures 4 and 5 where it is the case), and it is clear that one should introduce new non-concave quantities in order to hope to put into light non-concave spectra. The first line of ideas which has been developed in this respect is to replace scaling functions (which are necessarily concave by construction) by an *Increment Large Deviation Spectrum*, which, roughly speaking, indicate (on a log-log scale) the number of increments of the function of size $\sim 2^{-Hj}$ after a discretization of the function on a time step 2^{-j}. Let us be more precise: Starting with a function defined on (say) $[0, 1]$, we consider its 2^j increments

$$\delta_{j,k} = \left| f\left(\frac{k+1}{2^j}\right) - f\left(\frac{k}{2^j}\right)\right| \qquad k = 0 \cdots 2^j - 1. \tag{43}$$

Let

$$L_j(\alpha, \beta) = Card\left\{\delta_{j,k} : \ \ 2^{-\beta j} \leq \delta_{j,k} \leq 2^{-\alpha j}\right\}$$

Then the *large deviation increment spectrum* is

$$\nu_f(H) = \lim_{\varepsilon \to 0} \limsup_{j \to +\infty} \frac{\log(L_j(H-\varepsilon, H+\varepsilon))}{\log(2^j)}. \tag{44}$$

Considering this spectrum leads however to several problems, either on the mathematical side (non-invariance with respect to translations, or to small smooth perturbations of f), but also it is very hard to compute in a numerically stable way because of the double limit in the definition, which in practice has to be computed as a single limit. Recall that shifting from the Kolmogorov scaling function to the Leader scaling function was an important improvement on the mathematical and practical side; following the same line of ideas, one can replace increments of the function by wavelet leaders in (43), and define a *Large Deviation Leaders Spectrum* as follows.

Definition 14 *Let f be such that $H_f^{min} > 0$; for $0 \leq \alpha < \beta$, let*

$$M_j(\alpha, \beta) = Card\left\{(d_\lambda)_{\lambda \in \Lambda_j} : \ 2^{-\beta j} \leq d_\lambda \leq 2^{-\alpha j}\right\}.$$

The Large Deviation Leader Spectrum (LDLS) is of f is

$$\mu_f(H) = \lim_{\varepsilon \to 0} \limsup_{j \to +\infty} \frac{\log(M_j(H-\varepsilon, H+\varepsilon))}{\log(2^j)} \tag{45}$$

The heuristic behind this definition is that, at scale j, there are $\sim 2^{\mu_f(H)j}$ leaders of size $\sim 2^{-Hj}$. Note that an important advantage of working with quantities that are defined through quantiles is that they are not sensitive to instabilities due to the possible presence of *fat tails* in the probability distributions of the quantities observed (whether they be increments, wavelet coefficients or leaders); their presence often makes high moments diverge theoretically, and creates numerical instabilities in practical computations, whereas quantiles remain well defined, and insensitive to these phenomena. The following result shows that, if $\mu_f(H)$ is not concave, then estimating the multifractal spectrum with the help of the LDLS leads to a sharper estimation than with the Legendre spectrum.

Theorem 5 *If $H_f^{min} > 0$, then,*

$$\forall H, \quad d_f(H) \leq \mu_f(H) \leq L_f(H),$$

and $L_f(H)$ is the concave hull of $\mu_f(H)$.

Proof: We first prove that $d_f(H) \leq \mu_f(H)$. Let H be fixed; if $x_0 \in E_H^f$, then it follows from (32) that $\exists j_n \to \infty$ such that

$$2^{(-H-\varepsilon)j_n} \leq d_{j_n}(x_0) \leq 2^{(-H+\varepsilon)j_n}; \tag{46}$$

let

$$D_j = \{\lambda : \quad 2^{(-H-\varepsilon)j} \leq d_\lambda \leq 2^{(-H+\varepsilon)j}\}$$

It follows that

$$E_H^f \subset \limsup D_j.$$

Let now $\delta > 0$ be fixed. We pick ε such that

$$\left| \limsup_{j \to +\infty} \frac{\log(M_j(H-\varepsilon, H+\varepsilon))}{\log(2^j)} - \mu_f(H) \right| \leq \delta;$$

then, for j large enough,

$$\frac{\log(M_j(H-\varepsilon, H+\varepsilon))}{\log(2^j)} \leq \mu_f(H) + 2\delta,$$

so that

$$M_j(H-\varepsilon, H+\varepsilon) \leq 2^{(\mu_f(H)+2\delta)j}.$$

Therefore D_j is covered by $2^{(\mu_f(H)+2\delta)j}$ dyadic cubes of generation j, and it follows that

$$\dim(\limsup D_j) \leq \mu_f(H) + 2\delta.$$

Since this estimate holds for any $\delta > 0$, the first upper bound is proved. We now turn to the second one.

We first assume that the support of μ_f is a closed interval of the form $[H_{min}, H_{max}]$. Let $\varepsilon > 0$ be given; $\forall H \in [H_{min}, H_{max}]$, $\exists \delta(H) > 0$ such that

$$\forall j \qquad M_j(H - \delta(H), H + \delta(H) \leq 2^{(\mu_f(H)+\varepsilon)j};$$

the union of the intervals $(H - \delta(H), H + \delta(H))$ covers $[H_{min}, H_{max}]$. By compacity, we extract a finite covering thus obtaining $H_1, \cdots, H_n$ such that

$$2^{-dj} \sum_{\lambda \in \Lambda_j} (d_\lambda)^p \leq 2^{-dj} \sum_{i=1}^{n} 2^{(\mu_f(H_i)+\varepsilon)j} 2^{(-H_i p + \delta|p|)j},$$

so that, $\forall p$

$$\zeta_f(p) \geq \inf_{i=1,\cdots,n} d + H_i p - \delta|p| - \mu_f(H - i) - \varepsilon.$$

Since this is true for ε and δ arbitrarily small, the second bound follows.

If the support of μ_f is not bounded from above, then L_f is an increasing function.

We now prove the last assertion of the theorem. Let $\varepsilon > 0$ be fixed; by definition of $\mu_f(H)$, $\forall \delta > 0$, $\exists j_n \to \infty$ such that

$$M_{j_n}(H - \delta, H + \delta) \geq 2^{(\mu_f(H)-\delta)j_n}.$$

Therefore

$$\forall p > 0, \qquad \zeta_f(p) \leq d + p(H + \delta) - \mu_f(H) + \varepsilon$$

and

$$\forall p < 0, \qquad \zeta_f(p) \leq d + p(H - \delta) - \mu_f(H) + \varepsilon$$

Since this estimate holds for ε and δ arbitrarily small, the result follows from the definition of $L_f(H)$.

5.3 The Quantile Leader Spectrum

Note that some of the drawbacks attached with the Increment Large Deviation Spectrum remain with its leader variant, and, in particular the numerical instabilities due to the definition through the double limit in (44). Therefore, we now develop and intermediate method, which yields a spectrum defined as a single limit, but still allows put into light non-concave spectra. It is based on quantities which are in the spirit of the definitions of

quantiles in statistics, and therefore, we call the quantity that it yields the *Quantile Leader Spectrum*.

We now assume that the Large Deviation Leader Spectrum has a unique maximum at a point $H = H_{med}$ for which

$$\mu_f(H_{med}) = d \tag{47}$$

(note that this condition is equivalent to the fact that the Legendre spectrum has the same property). Note that H_{med} can then be computed using the median of the (d_λ) at scale j, which we denote by $Med_j(f)$. Indeed,

$$H_{med} = \lim_{j \to +\infty} \frac{\log(Med_j(f))}{\log(2^{-j})}. \tag{48}$$

Proof of (48): Because of the uniqueness of the maximum, $\forall \varepsilon > 0, \quad \exists \delta > 0$, such that, for j large enough,

$$\text{Card}\left\{\lambda : \quad d_\lambda \leq 2^{-(H_{med}+\varepsilon)j}\right\} \leq 2^{(d-\delta)j}$$

and

$$\text{Card}\left\{\lambda : \quad d_\lambda \geq 2^{-(H_{med}-\varepsilon)j}\right\} \leq 2^{(d-\delta)j};$$

therefore,

$$\text{Card}\left\{\lambda : \quad 2^{-(H_{med}+\varepsilon)j} \leq d_\lambda \leq 2^{-(H_{med}-\varepsilon)j}\right\} \leq 2^{dj} - 2 \cdot 2^{(d-\delta)j},$$

so that, for j large enough, the median of the d_λ belongs to $\left[2^{-(H_{med}+\varepsilon)j}, 2^{-(H_{med}-\varepsilon)j}\right]$. Since ε can be chosen arbitrarily small, (48) follows.

Definition 15 *The Quantile Leader Spectrum (QLS) $\mathcal{Q}_f(H)$ is an increasing function for $H \leq H_{med}$ and decreasing for $H \geq H_{med}$, defined by:*

- *If $H < H_{med}$, let*

  $$M_j(H) = Card\left\{(d_\lambda)_{\lambda \in \Lambda_j} : \quad d_\lambda \geq 2^{-Hj}\right\},$$

  Then

  $$\mathcal{Q}_f(H) = \limsup_{j \to +\infty} \frac{\log(M_j(H))}{\log(2^j)}.$$

- *If $H > H_{med}$, let*

  $$M_j(H) = Card\left\{(d_\lambda)_{\lambda \in \Lambda_j} : \quad d_\lambda \leq 2^{-Hj}\right\},$$

  Then

  $$\mathcal{Q}_f(H) = \limsup_{j \to +\infty} \frac{\log(M_j(H))}{\log(2^j)}.$$

The following result shows that, if $\mathcal{Q}_f(H)$ is not concave, then estimating the multifractal spectrum with the help of the QLS leads to a sharper estimation than with the Legendre spectrum. Recall that the increasing hull of a function f is the smallest increasing function g which is everywhere larger than f (the definition is similar for the decreasing hull).

Theorem 6 *Let f be a function such that $H_f^{min} > 0$. The Quantile Leader Spectrum $\mathcal{Q}_f(H)$ is the increasing hull of $\mu_f(H)$ for $H \leq H_{med}$ and its decreasing hull for $H \geq H_{med}$. Therefore, the following inequalities hold*

$$\forall H, \quad d_f(H) \leq \mu_f(H) \leq \mathcal{Q}_f(H) \leq L_f(H).$$

The proof follows follows directly from the definition of the QLS.

In Figures 1, 4 and 5, we show several examples of functions and random processes with non-concave spectra, and the additional information which is supplied by the quantile spectrum. The example supplied by the Riemann function (Fig. 1) is particularly instructive: The Legendre spectrum shows two straight lines in the spectrum, which may potentially both be signatures of non-concavities in the multifractal spectrum; the analysis through leader quantiles allows to settle these ambiguities and put into light two different behaviors: The increasing part is indeed a straight line whereas the decreasing part is the concave hull of a non-concave part of the spectrum. A similar conclusion can be drawn for the example of Lévy processes with Brownian component: The nonconcavity of the increasing part of the spectrum is revealed by the quantile method. The third example (Fig. 5,) however shows the limitations of the method: The quantile method does not allow the reveal the non-concavity of the multifractal spectrum. One possible explanation is supplied by finite size effects: In theory the only two pointwise exponents that can appear are H and $2H$, yielding wavelet leders of size $\sim 2^{-Hj}$ and $\sim 2^{-2Hj}$; however, regions where the fBm is close to vanish but does not do so are expected to generate wavelet leaders of intermediate size; ans this phenomenon can happen at all scales, because of the selfsimilarity of the underlying fBm. Hence the presence of a spurious numerical spectrum that is responsible for a "phase transition" between H and $2H$, and is also present in the quantile spectrum.

Note that the quantile procedure will allow to recover large deviation leader spectra that are increasing for $H < H_{med}$ and then decreasing for $H > H_{med}$; however, it only yields partial results in situations where this spectrum has several local maxima. If such is the case, one can however often reduce to the one maximum case, while still avoiding the double limit problem inherent with Definition 14, by using the following strategy: One first performs a quantile analysis, which will indicate the positions of some local maxima at the extremity of the flat regions. This allows for a first guess of the locations of the local minima (taking for instance the middle points of these flat regions). A localization of the spectrum is then performed by picking two consecutive "guessed" local minima H_1 and H_2, and then considering the "localized histogram" obtained by only keeping the wavelet leaders that satisfy $2^{-H_2 j} \leq d_\lambda \leq 2^{-H_1 j}$, and then performing a new quantile analysis on these reduced data. Note that, if the "region" selected in the LDLS does not contain H_{med}, then (47) will not hold for this set of coefficients. However, assuming that the LDLS for these coefficients still displays a unique local maximum for a value $H = H_3$ with $\nu_f(H_3) = \delta < d$, then the same arguments as above can be reproduced, indeed, the number of leaders selected will be $\sim 2^{\delta j}$ and the median of this new set will be located close to $2^{\delta j}$; this second step allows to reveal new local maxima of the LDS which become global maxima of the reduced spectrum, an it also allows to reveal new parts of the LDS that were "hidden" below the flat parts of the first quantile spectrum. One can iterate this procedure allowing to sharpen the position

of the local minima, and to multiply them, until one reaches a situation where the accuracy of the procedure is not sufficient to proceed further.

Note that an alternative procedure in order to deal with non-concave spectra has been proposed by C. Beck and H. Touchette, which has the effect of adding a parabola (which can be arbitrarily tailored) to the spectrum, therefore allowing to "dig into the holes" of the spectrum, while still using a Legendre-type method, see [50]. A comparison (both theoretical and numerical) between these methods still needs to be performed.

6 Multifractal analysis of non-locally bounded functions

6.1 Convergence and divergence rates for wavelet series

In this section, we investigate how the concepts of multifractal analysis can be developed in a different framework: The convergence and divergence rates of series. We start by recalling some results concerning the multifractal analysis of the divergence of Fourier series. We denote by $S_n f$ the partial sums of the Fourier series of a 2π periodic function f:

$$S_n f(x) = \sum_{k=-n}^{n} c_k e^{ikx} \qquad \text{where} \qquad c_k = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)e^{-ikt}dt.$$

Let

$$E_f^\beta = \left\{ x : \quad \limsup_{n\to\infty}\ n^{-\beta}|S_n f(x)| > 0 \right\}.$$

J.-M. Aubry proved that, if $f \in L^p([-\pi,\pi])$, and if $\beta > 0$, then $\dim(E_f^\beta) \leq 1 - \beta p$, and he showed the optimality of this result, see [4]. This was later extended and refined by F. Bayart and Y. Heurteaux, who, in particular, showed that optimality holds for generic functions (in the sense supplied by prevalence), see [19].

We now consider wavelet series. Recall that, if $f \in L^p$, then its wavelet series converges almost everywhere; however this leaves open the problem of improving this result if f is smoother (typically $f \in L^{p,s}$ for an $s > 0$) or, in the opposite direction, determining how fast the wavelet series diverges, when it does. The first result of this type stated that, if $f \in L^{p,s}(\mathbb{R}^d)$ for a $s > 0$, then the wavelet series of f converges outside of a set of dimension at most $d - sp$, see [20].

The consideration of sets of divergence such as E_f^β can also be done in the wavelet setting. For simplicity, we assume that the wavelets used are compactly supported. We rewrite the series (28) under the form

$$f(x) = \sum_{j=-1}^{\infty} \Delta_j f(x), \qquad \text{where} \qquad \Delta_j f(x) = \sum_{i,k} c_{j,k}^i \psi^{(i)}(2^j x - k).$$

Note that, for a given x, $\Delta_j f(x)$ only contains a finite number of terms, bounded by a constant which depends only on the size of the support of the wavelets and on the space

dimension. Let $\beta > 0$ and let F_f^β denote the set of points where the partial sums are not bounded by $2^{\beta j}$. J.-M. Aubry put into light an interesting similarity between Fourier series and wavelet series, showing that, if $f \in L^p(\mathbb{R})$, then $\dim(F_f^\beta) \leq 1 - \beta p$, see [4].

Our purpose now is to show some simple extensions of these results; we will prove their optimality through a simple use of random wavelet series in Section 6.3.

Definition 16 *Let $\beta > 0$.*

- *The wavelet series of a distribution f converges at rate β at x if*

$$\exists C > 0 : \qquad |\Delta_j f(x)| \leq C 2^{-\beta j};$$

  we denote by F_f^β the complement of this set.

- *The wavelet series of a distribution f diverges at rate β at x if*

$$\exists C > 0 : \qquad \limsup_{j \to +\infty} 2^{-\beta j} |\Delta_j f(x)| > 0.$$

we denote this set by D_f^β.

Proposition 5 *Let $f \in L^{p,s}$ where $s < d/p$, and let β be such that*

$$-s < \beta < \frac{d}{p} - s;$$

then

$$\dim\left(D_f^\beta\right) \leq d - sp - \beta p.$$

Let $f \in L^{p,s}$ with $s \in \mathbb{R}$, and let $\beta > 0$; then

$$\dim\left(F_f^\beta\right) \leq d - sp + \beta p.$$

Remark: If $s \geq d/p$, then $f \in C^{s-d/p}$ so that its wavelet coefficients are bounded by $C 2^{-(s-d/p)j}$ and therefore the wavelet series of f converges everywhere (and at rate at least $s - d/p$). This explains why we make the assumption $s < d/p$ in the first part of the proposition.

Proof: Let $f \in L^{p,s}$. Using the classical embedding between Sobolev and Besov spaces $L^{p,s} \subset B_p^{s,\infty}$, we obtain that

$$2^{-dj} \sum_{\lambda \in \Lambda_j} |c_\lambda|^p \leq C \cdot 2^{-spj} \tag{49}$$

Let $\alpha \in \mathbb{R}$, and

$$E_{j,\alpha} = \{\lambda : \quad |c_\lambda| \geq 2^{-\alpha j}\}.$$

Let A be an odd integer larger than the support of the wavelets, and

$$F_{j,\alpha} = \{A \cdot \lambda\}_{\lambda \in E_{j,\alpha}}.$$

It follows from (49) that $Card(E_{j,\alpha}) \leq C \cdot 2^{(d-sp-\alpha p)j}$, so that

$$Card(F_{j,\alpha}) \leq C \cdot 2^{(d-sp-\alpha p)j}.$$

Let $F_\alpha = \limsup F_{j,\alpha}$. Since $F_{j,\alpha}$ is composed of at most $2^{(d-sp-\alpha p)j}$ dyadic cubes of width 2^{-j}, using these cubes for $j \geq J$ as a covering of F_α, we obtain that

$$\dim F_\alpha \leq d - sp - \alpha p.$$

If $x \notin F_\alpha$, then $|c_\lambda \psi_\lambda(x)| \leq C 2^{-\alpha j}$, and the localization of the wavelets implies that $|\Delta_j(f)(x)| \leq C 2^{-\alpha j}$. We apply this result with either $\beta = \alpha$ in the first part of the proposition, or $\beta = -\alpha$ in the second part.

The optimality of Proposition 5 will be proved in Section 6.3 by considering random wavelet series.

6.2 q-leaders

The construction of new scaling functions beyond (33) was motivated by the following restriction: In order to be used, the wavelet leader method requires the data to be locally bounded. We saw a practical procedure in order to decide if this assumption is valid, namely the determination of the uniform Hölder exponent H_f^{min}. Experimental investigations showed that H_f^{min} is negative for large classes of natural "texture type" images, see [1, 2, 30, 31, 52, 53], and therefore the method cannot be used as such.

In order to circumvent this problem, one can replace the wavelet leaders by alternative quantities, which measure pointwise regularity (for another definition of regularity) and make sense even if the data are no more locally bounded.

We will use the following extension of pointwise smoothness, which was introduced by Calderón and Zygmund in 1961, see [11].

Definition 17 *Let $B(x_0, r)$ denote the open ball centered at x_0 and of radius r; let $q \in [1, +\infty)$ and $\alpha > -d/q$. Let f be function which locally belongs to $L^q(\mathbb{R}^d)$. Then f belongs to $T_\alpha^q(x_0)$ if there exist $C, R > 0$ and a polynomial P such that*

$$\forall r \leq R, \quad \left(\frac{1}{r^d} \int_{B(x_0,r)} |f(x) - P(x - x_0)|^q dx \right)^{1/q} \leq C r^\alpha. \tag{50}$$

The q-exponent of f at x_0 is

$$h_f^q(x_0) = \sup\{\alpha : f \in T_\alpha^q(x_0)\}.$$

Note that the Hölder exponent corresponds to the case $q = +\infty$. This definition is a natural substitute for pointwise Hölder regularity when dealing with functions which are not locally bounded, but locally belong to L^q. In particular, the q-exponent can take values down to $-d/q$, and therefore it allows to model behaviors which locally are of the form $1/|x - x_0|^\alpha$ for $\alpha < d/q$, i.e., to deal with negative regularity exponents. Before going further, let us give a practical criterium in order to determine if, indeed, data locally belong to L^q.

The function space interpretation of the wavelet scaling function (29) implies that

- If $\eta_f(q) > 0$ then $f \in L^q_{loc}$
- If $\eta_f(q) < 0$ then $f \notin L^q_{loc}$.

Therefore the value taken by the wavelet scaling function at q allows to determine if, indeed, data locally belong to L^q. We see here another use of the wavelet scaling function, as a preliminary quantity which is required to be computed. Therefore it plays a similar role as the computation of H_f^{min} when dealing with the multifractal analysis based on wavelet leaders. The verification of the criterium $\eta_f(q) > 0$ is a prerequisite for the following.

Let us now show how the notion of T^q_α regularity can be related to local l^q norms of wavelet coefficients. This will be done with the help of the following quantity.

Definition 18 *Let $f \in L^q_{loc}(\mathbb{R}^d)$. The q-leaders of f are defined by*

$$d^q_\lambda = \left(\sum_{\lambda' \subset 3\lambda_j(x_0)} |c_{\lambda'}|^q 2^{-d(j'-j)} \right)^{1/q}. \tag{51}$$

Under mild hypotheses on f, (see [29, 30, 34]) the pointwise q-exponent can be expressed by a regression on a log-log plot of the q-leaders:

$$h^q_f(x_0) = \liminf_{j \to +\infty} \frac{\log\left(d^q_{\lambda_j(x_0)}\right)}{\log(2^{-j})}. \tag{52}$$

The definition of the q-scaling function follows the one of the leader scaling function, except that wavelet leaders now are replaced by q-leaders:

$$\forall p \in \mathbb{R}, \qquad \zeta_f(p,q) = \liminf_{j \to +\infty} \frac{\log\left(2^{-dj} \sum_{\lambda \in \Lambda_j} |d^q_\lambda|^p\right)}{\log(2^{-j})}. \tag{53}$$

Note that, as above, a multifractal spectrum can be attached to the q-exponent, and a multifractal formalism can be worked out using the usual procedure; this spectrum is

obtained as a Legendre transform of the q-scaling function: We denote by $d_f^q(H)$ the multifractal spectrum associated with the q-exponent (i.e. the Hausdorff dimension of the set of points where the q-exponent takes the value H). If $\eta_f(q) > 0$, then (see [29, 30, 34]),

$$d_f^q(H) \leq \inf_{p \in \mathbb{R}} (d + Hp - \zeta_f(p, q)). \tag{54}$$

Figures 6 and 7 plot numerically computed multifractal spectra associated with q-exponents for random functions with negative q-exponents together with the theoretical spectra and the leader spectra. In each case, the condition that $f \in L^q$ has been verified; as mentioned earlier, it is implied by the following condition on the wavelet scaling function: $\eta_f(q) > 0$. It can be equivalently checked on the Legendre spectrum, by verifying that it lies below the straight line in red (joining the points $(-1/q, 0)$ and $(0, 1)$).

6.3 Variants of Random Wavelet Series

In this last section, we consider new simple examples of random multifractal functions, which are a generalization of models already considered in [5, 25]. The generalization we propose covers two aspects: First, the wavelet coefficients are not necessarily bounded, so that the random fields generated do not necessarily satisfy $H_f^{min} > 0$. It will follow that the standard multifractal analysis as developed in Sections 3 and 4 cannot be worked out, and we rather have to use the extension using q-exponent. The second generalization lies in the fact that the wavelet coefficients are not drawn at random uniformly (i.e. using the Lebesgue measure) but using a measure supported by a fractal set. This second point will allow to take into account clustering of large wavelet coefficients, a feature which is often observed in practice, and this model will generate different multifractal spectra than those obtained in [25]. In order to state the results concerning these random wavelet series, we need to start by recalling the techniques concerning points drawn at random in a fractal set; these techniques are referred to as *ubiquity methods*.

We start by describing the kind of measures which will be used in order to draw points at random.

Definition 19 *A probability measure ν supported by a compact set $B \subset \mathbb{R}^d$ is a regular δ-dimensional measure if it satisfies the following conditions:*

1. *$\exists C_1 > 0$ such that*

$$\forall x \in B,\ \forall r > 0, \qquad \nu(B(x, r)) \leq C_1 r^{\delta},$$

2. *$\exists C_2 > 0$ such that $\forall j \geq 0$, B can be covered by $C_2 \cdot 2^{\delta j}$ dyadic cubes of side 2^{-j}.*

Remark: We will also consider logarithmic corrections, i.e. (δ, γ)-dimensional measure which satisfy

1. $\exists C_1 > 0$ such that

$$\forall x \in B,\ \forall r > 0, \qquad \nu(B(x, r)) \leq C_1 r^{\delta} (\log r)^{-\gamma},$$

2. $\exists C_2 > 0$ such that $\forall j \geq 0$, B can be covered by $C_2 \cdot 2^{\delta j} j^{\gamma}$ dyadic cubes of side 2^{-j}.

Typical examples of such measures are supplied by the measures m satisfying (23) which we introduced in order to derive the Hausdorff dimensions of Cantor sets (in which case, B is precisely the corresponding Cantor set).

Let x_n be a sequence of points drawn independently using the probability measure ν, and let $\mathcal{L} = (l_n)_{n\in\mathbb{N}}$ be a decreasing sequence of positive real numbers that tends to 0. Ubiquity theory deals with the study of the sets

$$A_{\mathcal{L}} = \limsup B(x_n, l_n)$$

(i.e. $A_{\mathcal{L}}$ is the set of points that belong to an infinite number of the balls $B(x_n, l_n)$).

Note that the covering of $A_{\mathcal{L}}$ by the balls $B(x_n, l_n)$ for $n \geq N$ always yields an upper bound for the Hausdorff dimension of $A_{\mathcal{L}}$. The purpose of ubiquity methods is to obtain lower bounds under mild hypotheses on the sequence of points x_n, which express the fact that such points are "well spread", and which would coincide with the direct upper bounds obtained with the help of the sequence $B(x_n, l_n)$. Typically, drawing them independently and at random will allow such hypotheses to be fulfilled (but many situations where the points are not drawn at random are also efficiently covered by ubiquity techniques, see [8, 9]).

In the situation supplied by random points that we considered, the following result is an easy application of the Borel-Cantelli lemma.

Proposition 6 *If the sequence $\mathcal{L}$ satisfies*

$$\sum_n l_n^{\delta} = +\infty, \tag{55}$$

then, a.s. ν-almost every point of B belongs to $A_{\mathcal{L}}$.

If the stronger condition

$$\sum_n \frac{l_n^{\delta}}{(\log l_n)^3} = +\infty; \tag{56}$$

holds, then a.s. $B = A_{\mathcal{L}}$ (sufficient condition for almost sure covering).

The idea behind ubiquity techniques is that, if for a particular sequence $\mathcal{L}$ a.e. point of B is covered by $A_{\mathcal{L}}$, then, for smaller sequences, the set $A_{\mathcal{L}}$ cannot be too small, and, one can obtain lower bounds on the Hausdorff dimension of $A_{\mathcal{L}}$.

Definition 20 *Let ν be a regular δ-dimensional measure on $B \subset \mathbb{R}^d$. An ubiquity system associated with ν is a sequence $(\lambda_n, \varepsilon_n)$ such that*

- *$\lambda_n \in B$,*
- *ε_n a decreasing sequence of positive real numbers that tends to 0.*

- $\exists a > 0$ *such that* ν*-almost every point of* B *belongs to*

$$E_a = \limsup B(\lambda_n, \varepsilon_n^a).$$

The first part of Proposition 6 gives examples of ubiquity systems in a random setting. In order to estimate the size of the sets $\mathcal{D}_\alpha$, we need a refinement of the notion of Hausdorff δ-dimensional measure obtained through logarithmic corrections.

Definition 21 *Let* A *be a Borel subset of* $\mathbb{R}^d$*. Let* $\varepsilon > 0$*,* $\delta \in [0, d]$*, and* $\gamma \in \mathbb{R}$*. Let*

$$M_\varepsilon^{\delta,\gamma}(A) = \inf_R \left(\sum_i |A_i|^\delta |\log(|A_i|)|^\gamma \right),$$

where R *is an* ε*-covering of* A*. (The infimum is therefore taken on all* ε*-coverings.)*

For any $\delta \in [0, d]$*, the* (δ, γ)*-dimensional outer Hausdorff measure of* A *is*

$$mes_{\delta,\gamma}(A) = \lim_{\varepsilon \to 0} M_{\varepsilon,}^{\delta,\gamma}(A). \tag{57}$$

The following result is usually referred to as the *ubiquity theorem*. It is a slight variant of similar results proved in [10, 25].

Theorem 7 *Let* $(\lambda_n, \varepsilon_n)$ *be an ubiquity system associated with* ν*. If* ν*-almost every point of* B *belongs to* E_a*, then*

$$\forall b > a \quad mes_{\delta a/b,2}(E_b) > 0. \tag{58}$$

We now describe the model of Random Wavelet Series that we will consider. We assume that the wavelet basis used belongs to the Schwarz class.

Definition 22 *Let* $\eta \in]0, \delta[$ *et* $\alpha \in \mathbb{R}$*. Let* (x_n) *be a sequence of points drawn independently using the probability measure* ν*. This sequence is split into subsets* E_j *of length* $[2^{\eta j}]$*. For each* $x_n \in E_j$ *let* $\lambda_{n,j}$ *be the dyadic cube which contains* x_n*. The random field* $X_{\alpha,\eta}$ *is*

$$X_{\alpha,\eta}(x) = \sum_j \sum_{\lambda_{n,j}} 2^{-\alpha j} \psi_{\lambda_{n,j}}(x), \tag{59}$$

where the sum on the $\lambda_{n,j}$ *is such that the corresponding points* x_n *belong to* E_j*.*

It follows form this definition that $H_f^{\min} = \alpha$, and that the sample paths of $X_{\alpha,\eta}$ belong to $C^\alpha(\mathbb{R}^d)$.

Let us now determine under which condition the sample paths of $X_{\alpha,\eta}$ belong to L^q. For each dyadic bloc

$$\Delta_j f = \sum_{\lambda_{n,j} \in \Lambda_j} 2^{-\alpha j} \psi_{\lambda_{n,j}},$$

using the localization of the wavelets, $\| \Delta_j f \|_q \sim 2^{-(\alpha+(d-\eta)/q)j}$. So that the sample paths of $X_{\alpha,\eta}$ belong to $L^q(\mathbb{R}^d)$ if and only if

$$\alpha + \frac{d-\eta}{q} > 0. \tag{60}$$

Note that, if $\alpha < -d$, this condition will never be fulfilled. In such situations, a multifractal analysis of $X_{\alpha,\eta}$ can still be performed, either using the notion of *weak scaling exponent* (which allows to associate a pointwise regularity index in the general setting of tempered distributions, see [30, 31, 46]) or by first performing a fractional integration on the data (see [2, 30, 31] for a discussion on the practical implementation of the method as well as numerical results).

By the same argument as in the L^q case, one obtains that the sample paths of $X_{\alpha,\eta}$ belong to $L^{p,s}(\mathbb{R}^d)$ if and only if

$$\alpha - s + \frac{d-\eta}{p} > 0. \tag{61}$$

Let us now estimate the size of the sets of points where the series defining $X_{\alpha,\eta}$ will converge or diverge at a given rate. Let A be a dyadic subcube of the support of ψ of width 2^{-a} where

$$\forall x \in A, \qquad |\psi(x)| \geq C > 0.$$

The image of A under the mapping $x \to 2^j x - k$ yields a sub-cube $\lambda'(\lambda)$ of λ of width 2^{-j-a} where $|\psi_\lambda(x)| \geq C$. Let

$$D = \limsup_{\lambda_{n,j} \in E_j} \lambda'(\lambda_{n,j}).$$

If $\alpha < 0$, the series (59) diverges at rate $-\alpha$ on D; and if $\alpha > 0$, the series (59) converges at rate at most α on D. It follows from the ubiquity theorem that

$$\dim(D) = \eta.$$

Let us now check that indeed these series yield the optimality of Proposition 5 (and therefore, taking $s = 0$, also of J.-M. Aubry's result concerning the divergence of wavelet series in L^p). We pick an $\varepsilon > 0$ arbitrarily small such that $\alpha - s + \frac{d-\eta}{p} = \varepsilon$. It follows from (61) that the sample paths of $X_{\alpha,\eta}$ belong to L^p. We therefore obtain a set of dimension

$$\eta = d - (s-\alpha)q - \varepsilon q$$

where (59) diverges at rate $-\alpha$ if $\alpha < 0$, or converges at rate at most α if $\alpha < 0$. Since ε can be picked arbitrarily small, the optimality of both parts of Proposition 5 follows.

Note that one could construct random wavelet series whose set of divergence has dimension exactly given by Proposition 5 by using logarithmic corrections in the measure ν, as mentioned in the remark following Definition 19

We now study the pointwise regularity of the sample paths of $X_{\alpha,\eta}$. Clearly, the sample paths of $X_{\alpha,\eta}$ are C^∞ outside of set B. The pointwise regularity at points of B follows from the application of Theorem 7, together with (52), and one obtains the following description of the multifractal q-spectrum.

Theorem 8 *Let*

$$H_X^{min} = \alpha \qquad \text{and} \qquad H_X^{max} = \frac{\alpha\delta}{\eta} + \frac{d}{q}\left(\frac{\delta}{\eta} - 1\right).$$

If (60) holds, then the q-spectrum of $X_{\alpha,\eta}$ is supported by the interval $[H_X^{min}, H_X^{max}]$, and, on this interval, the multifractal spectrum of $X_{\alpha,\eta}$ is the segment joining the two points (H_X^{min}, η) and (H_X^{max}, δ).

Let us sketch how the pointwise q-exponent of $X_{\alpha,\eta}$ can be obtained. A first remark is that, if $\lambda'(\lambda)$ denotes the (random) subcube of 3λ of smallest scale j' such that $C_{\lambda'} \neq 0$, then, for j large enough, the p-leaders satisfy the following property:

$$\text{a.s.} \quad \forall\lambda \qquad d_\lambda^p \leq C j^3 c_{\lambda'};$$

this statement follows from the sparsity of the series, and an application of the Borel-Cantelli lemma. In turn, it implies that we can disregard the factor j^3 in the estimation of pointwise exponents (because it only brings logarithmic corrections in the pointwise regularity); it follows that the sets of points with the same p-exponents are of the same type as those that appear in "classical random" wavelet series (where α is positive, and the exponent considered in the Hölder exponent), and these sets of points can be expressed in terms of the sets which show up in the ubiquity setting. More precisely, they are of the form

$$\bigcap_{\gamma<\Gamma} E_\gamma - \bigcup_{\gamma>\Gamma} E_\gamma. \tag{62}$$

Upper bounds for the dimension of such sets follow from upper bounds for the dimension of each of the E_γ for a given $\gamma < \Gamma$, and lower bounds follow from the construction of a subset of positive Hausdorff measure on E_Γ (which is supplied by the ubiquity theorem), together with the fact that the upper bounds for dimensions yield that this Haudorff measure vanishes for the set $\bigcup_{\gamma>\Gamma} E_\gamma$ (using the remark that the union can be rewritten as a countable union).

Acknowledgement: The authors acknowledge the supports of the GDR "Analyse Multifractale", the Bézout Labex and the ANR project AMATIS.

References

[1] P. Abry, H. Wendt, S. Jaffard, H. Helgason, P. Goncalves, E. Pereira, C. Gharib, P. Gaucherand, and M. Doret. Methodology for multifractal analysis of heart rate variability: From lf/hf ratio to wavelet leaders. In *Nonlinear Dynamic Analysis of Biomedical Signals EMBC conference (IEEE Engineering in Medicine and Biology Conferences) Buenos Aires,*, 2010. 36

[2] S. Abry, P. Jaffard and H. Wendt. Irregularities and scaling in signal and image processing: Multifractal analysis,. *Benoit Mandelbrot: A Life in Many Dimensions, M. Frame, Ed., World scientific publishing*, to appear. 26, 28, 36, 41

[3] F. Ai-Hua, L. Liao, J. Schmeling, and M. Wu. Multifractal analysis of some multiple ergodic averages. *preprint*, 2012. 15

[4] J.-M. Aubry. On the rate of pointwise divergence of fourier and wavelet series in lp. *Journal of Approx. Th.*, 538:97–111, 2006. 15, 34, 35

[5] J.M. Aubry and S. Jaffard. Random wavelet series. *Communications In Mathematical Physics*, 227(3):483–514, 2002. 24, 38

[6] J. Barral, J. Berestycki, J. Bertoin, A. Fan, B. Haas, S. Jaffard, G. Miermont, and J. Peyrière. *Quelques interactions entre analyse, probabilités et fractals.* S.M.F., Panoramas et synthèses 32, 2010. 6, 24

[7] J. Barral, A. Durand, S. Jaffard, and S. Seuret. Local multifractal analysis. *Contempory Mathematics, D. Carfi, M. L. Lapidus, E. J. Pearse, and M. van Frankenhuijsen eds.*, to appear. 26

[8] J. Barral, N. Fournier, S. Jaffard, and S. Seuret. A pure jump Markov process with a random singularity spectrum. *Annals of Probability*, 38(5):1924–1946, 2010. 26, 39

[9] J. Barral and S. Seuret. A localized Jarnik-Besicovich theorem. *Adv. Math.*, 226(4):3191–3215, 2011. 39

[10] V. Beresnevich and S. Velani. A mass transference principle and the duffin-schaeffer conjecture for hausdorff measures. *Ann. of Math*, 164:971–992, 2006. 40

[11] A.P. Calderon and A. Zygmund. Local properties of solutions of elliptic partial differential equations. *preprint Studia Math.,*, 20:171–223, 1961. 36

[12] L. Calvet, A. Fisher, and B. Mandelbrot. The multifractal model of asset returns. *Cowles Foundation Discussion Papers: 1164*, 1997. 6

[13] I. Daubechies. Orthonormal bases of compactly supported wavelets. *Comm. Pure and App. Math.*, 41:909–996, 1988. 17

[14] A. Durand and S. Jaffard. Multifractal analysis of Lévy fields. *Proba. Theo. Related Fields (to appear)*, 2011. 28

[15] K. Falconer. *Fractal Geometry.* Wiley, 1990. 16

[16] K. Falconer. *Techniques in Fractal Geometry.* Wiley, 1997. 16

[17] Y. Gousseau and J.-M. Morel. Are natural images of bounded variation? *SIAM J. Math. Anal.*, 33:634–648, 2001. 22

[18] T.C. Halsey, M.H. Jensen, L.P. Kadanoff, I. Procaccia, and B.I. Shraiman. Fractal measures and their singularities - the characterization of strange sets. *Phys. Rev. A*, 33(2):1141–1151, 1986. 15

[19] Y. Heurteaux and F Bayart. Multifractal analysis of the divergence of fourier series. *Ann. Scient. ENS.*, 45:927–946, 2012. 15, 34

[20] S. Jaffard. Wavelets and nonlinear analysis. *preprint Wavelets: Mathematics and applications J. Benedetto and M. Frazier Eds., Studies in Advanced Mathematics, CRC*, pages 467–503, 1994. 34

[21] S. Jaffard. The spectrum of singularities of riemann's function,. *preprint Rev. Mat. Iberoamericana.*, 12:441–460, 1996. 27

[22] S. Jaffard. Multifractal formalism for functions. *SIAM J. of Math. Anal.*, 28(4):944–998, 1997. 17, 21

[23] S. Jaffard. Oscillation spaces: Properties and applications to fractal and multifractal functions. *Journal of Mathematical Physics*, 39(8):4129–4141, 1998. 25, 26

[24] S. Jaffard. The multifractal nature of Levy processes. *Probability Theory and Related Fields*, 114(2):207–227, 1999. 6, 15, 28

[25] S. Jaffard. On lacunary wavelet series. *Annals of Applied Probability*, 10(1):313–329, 2000. 38, 40

[26] S. Jaffard. On davenport expansions. *Fractal Geometry and Applications: A Jubilee of Benoit Mandelbrot - Analysis, Number Theory, and Dynamical Systems, Pt 1*, 72:273–303, 2004. 15, 17, 23

[27] S. Jaffard. Wavelet techniques in multifractal analysis. In M. Lapidus and M. van Frankenhuijsen, editors, *Fractal Geometry and Applications: A Jubilee of Benoît Mandelbrot, Proc. Symp. Pure Math.*, volume 72(2), pages 91–152. AMS, 2004. 21

[28] S. Jaffard. Beyond Besov spaces, part 2: Oscillation spaces. *Constructive Approximation*, 21(1):29–61, 2005. 25, 26

[29] S. Jaffard. Wavelet techniques for pointwise regularity. *Ann. Fac. Sci. Toul.*, 15(1):3–33, 2006. 37, 38

[30] S. Jaffard, P. Abry, and S.G. Roux. Function spaces vs. scaling functions: tools for image classification. *Mathematical Image processing (Springer Proceedings in Mathematics) M. Bergounioux ed.*, 5:1–39, 2011. 21, 25, 26, 36, 37, 38, 41

[31] S. Jaffard, P. Abry, S.G. Roux, B. Vedel, and H. Wendt. *The contribution of wavelets in multifractal analysis*, pages 51–98. Series in contemporary applied mathematics. World scientific publishing, 2010. 23, 24, 36, 41

[32] 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. 17, 24

[33] S. Jaffard, B. Lashermes, and P. Abry. Wavelet leaders in multifractal analysis. In *Wavelet Analysis and Applications, T. Qian, M.I. Vai, X. Yuesheng, Eds.*, pages 219–264, Basel, Switzerland, 2006. Birkhäuser Verlag. 24

[34] S. Jaffard and C. Melot. Wavelet analysis of fractal boundaries. *Communications In Mathematical Physics*, 258(3):513–565, 2005. 37, 38

[35] A.N. Kolmogorov. The Wiener spiral and some other interesting curves in Hilbert space (russian),. *Dokl. Akad. Nauk SSSR*, 26:(2):115118, 1940. 4

[36] 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. 16

[37] A.N. Kolmogorov. The local structure of turbulence in incompressible viscous fluid for very large Reynolds numbers. *Comptes Rendus De L'Academie Des Sciences De L'Urss*, 30:301–305, 1941. 16

[38] A.N. Kolmogorov. A refinement of previous hypotheses concerning the local structure of turbulence in a viscous incompressible fluid at high Reynolds number. *J. Fluid Mech.*, 13:82–85, 1962. 16

[39] S. Osher L. Rudin and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D*, 60:259–268, 1992. 21

[40] B. Lashermes, S. Jaffard, and P. Abry. Wavelet leader based multifractal analysis. *2005 Ieee International Conference On Acoustics, Speech, and Signal Processing, Vols 1-5*, pages 161–164, 2005. 17

[41] L. Liao and Seuret S. Diophantine approximation by orbits of expanding markov maps. *Ergod. Th. Dyn. Syst.*, 33:585–608, 2013. 15

[42] S. Mallat. *A Wavelet Tour of Signal Processing.* Academic Press, San Diego, CA, 1998. 17

[43] B. Mandelbrot. The variation of certain speculative price. *The Journal of Business*, 36(4):394–419, 1963. 6

[44] B. Mandelbrot and J.W. van Ness. Fractional Brownian motion, fractional noises and applications. *SIAM Reviews*, 10:422–437, 1968. 4

[45] Y. Meyer. *Ondelettes et Opérateurs.* Hermann, Paris, 1990. English translation, *Wavelets and operators*, Cambridge University Press, 1992. 5, 17, 20

[46] Y. Meyer. *Wavelets, vibrations and scalings.* CRM Ser. AMS Vol. 9,, Presses de l'Université de Montréal, Paris, 1998. 41

[47] D. Monrad and L.D. Pitt. Local nondeterminism and Hausdorff dimension. *Progress in Probability and Statistics, Seminar on Stochastic Processes 1986, (E. Cinlar, K.L. Chung, R.K. Getoor, Eds.), Birkhäuser, Boston*, page 163189, 1986. 28

[48] A.M. Obukhov. Some specific features of atmospheric turbulence. *J. Fluid Mech.*, 13:77–81, 1962. 16

[49] G. Parisi and U. Frisch. Fully developed turbulence and intermittency. In M. Ghil, R. Benzi, and G. Parisi, editors, *Turbulence and Predictability in geophysical Fluid Dynamics and Climate Dynamics*, Proc. of Int. School, page 84, Amsterdam, 1985. North-Holland. 15, 17

[50] H. Touchette and C. Beck. Nonconcave entropies in multifractals and the thermodynamic formalis,. *preprint J. Stat. Phys.*, 125:459–475, 2006. 34

[51] H. Wendt, P. Abry, and S. Jaffard. Bootstrap for empirical multifractal analysis. *IEEE Signal Processing Mag.*, 24(4):38–48, 2007. 17

[52] H. Wendt, P. Abry, S.G. Roux, and S. Jaffard. Analyse multifractale d'image : l'apport des coefficients dominants. In *21e colloque sur le Traitement du Signal et des Images. GRETSI 2007*, 11–14 Sep. 2007. 36

[53] H. Wendt, S.G. Roux, P. Abry, and S. Jaffard. Wavelet leaders and bootstrap for multifractal analysis of images. *Signal Processing*, 89(6):1100–1114, 2009. 36

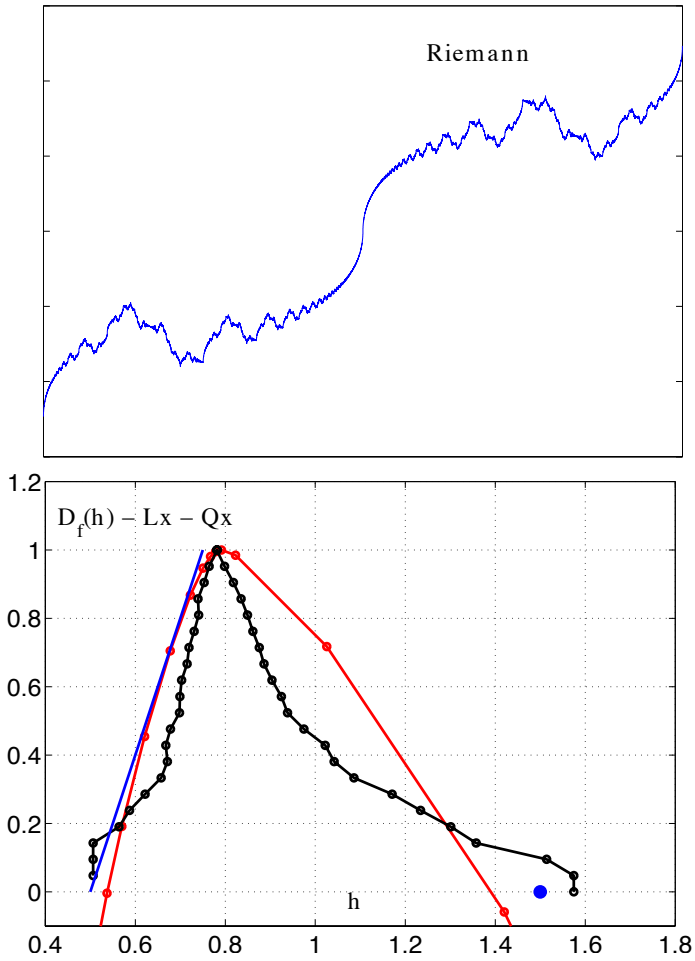

Figure 1: Top: Riemann's non-differentiable function. Bottom: theoretical spectrum (blue), leaders Legendre spectrum (red), leaders quantile spectrum (black).

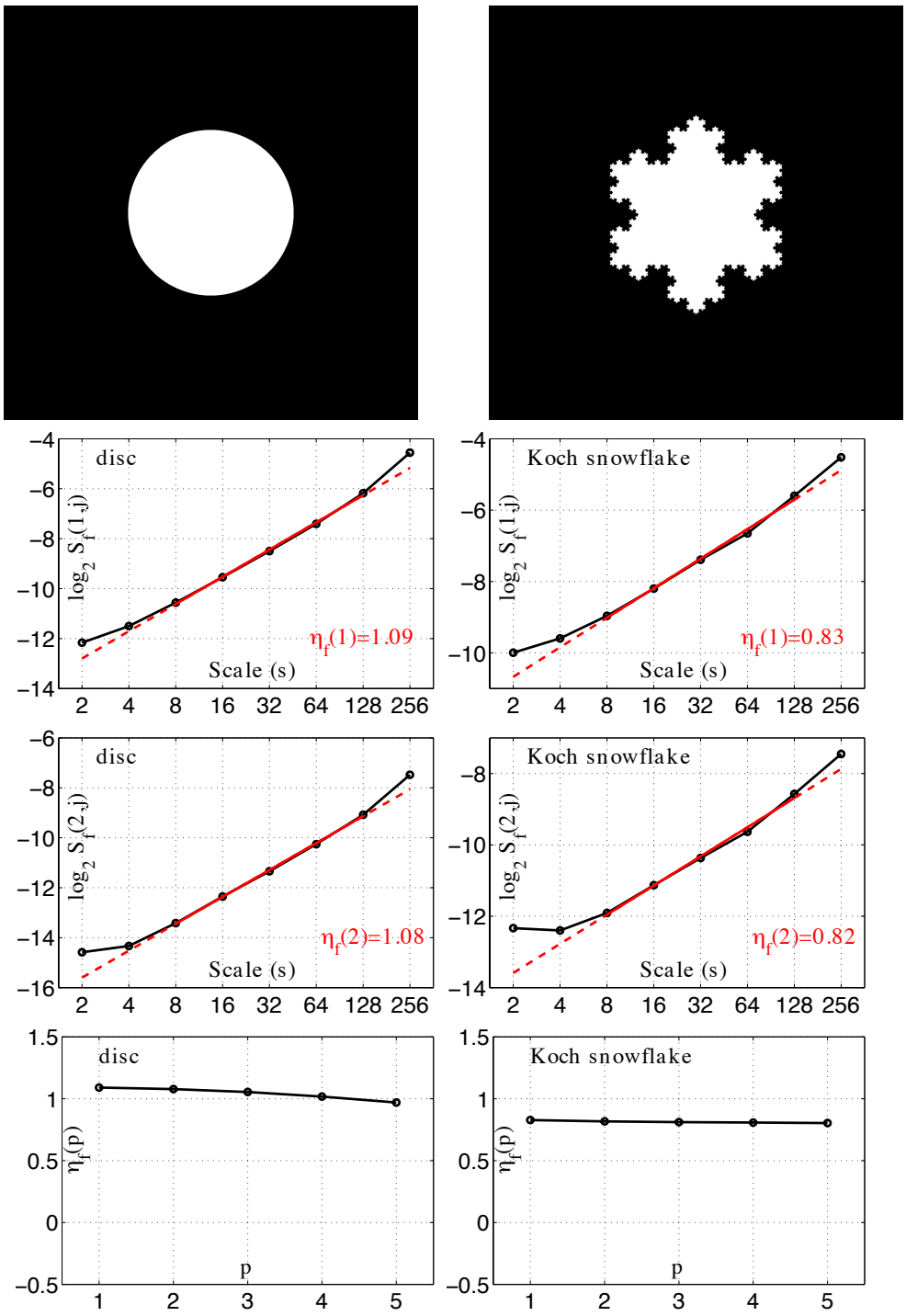

Figure 2: Indicator functions (top), structure functions for moments $p = 1$ and $p = 2$ (second and third rows), and scaling exponents for $p = \{1, 2, 3, 4, 5\}$ (bottom row): disc (left column) and Von Koch snowflake (right column). The scaling functions of indicator functions are constant and measure the fractal dimension of the sets as $D = 2 - \eta_f(p)$ ($D = 1$ for the disc and $D = 2 - \ln(4)/\ln(3) \approx 0.74$ for the Koch snowflake).

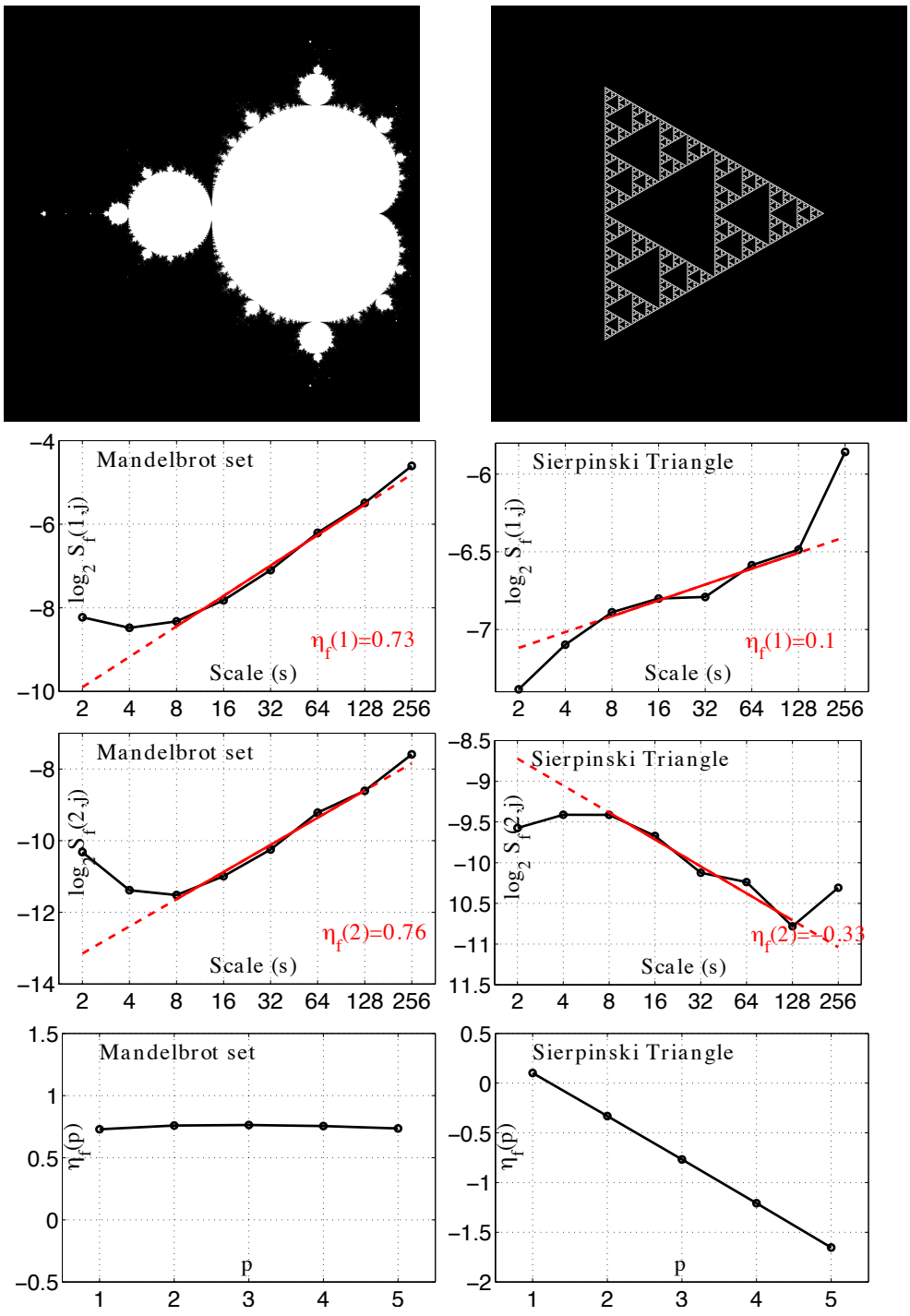

Figure 3: Indicator functions (top), structure functions for moments $p = 1$ and $p = 2$ (second and third rows), and scaling exponents for $p = \{1, 2, 3, 4, 5\}$ (bottom row): Mandelbrot set (left column) and Sierpinski triangle (right column). The scaling functions of indicator functions are constant and measure the fractal dimension of the sets as $D = 2 - \eta_f(p)$ ($D = 1$ for the disc and $D = 2 - \ln(4)/\ln(3) \approx 0.74$ for the Koch snowflake).

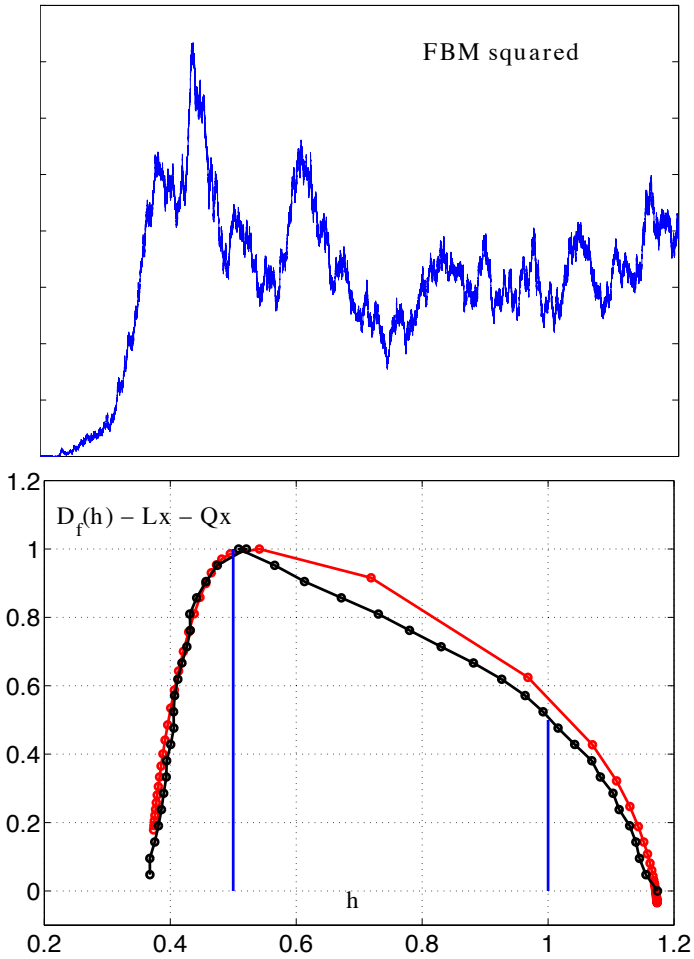

Figure 4: Square of fractional Brownian motion ($H = 0.5$). Top: single realization. Bottom: theoretical spectrum (blue), leaders Legendre spectrum (red), leaders quantile spectrum (black). Results obtained as means over 50 realizations of length $N = 2^{19}$.

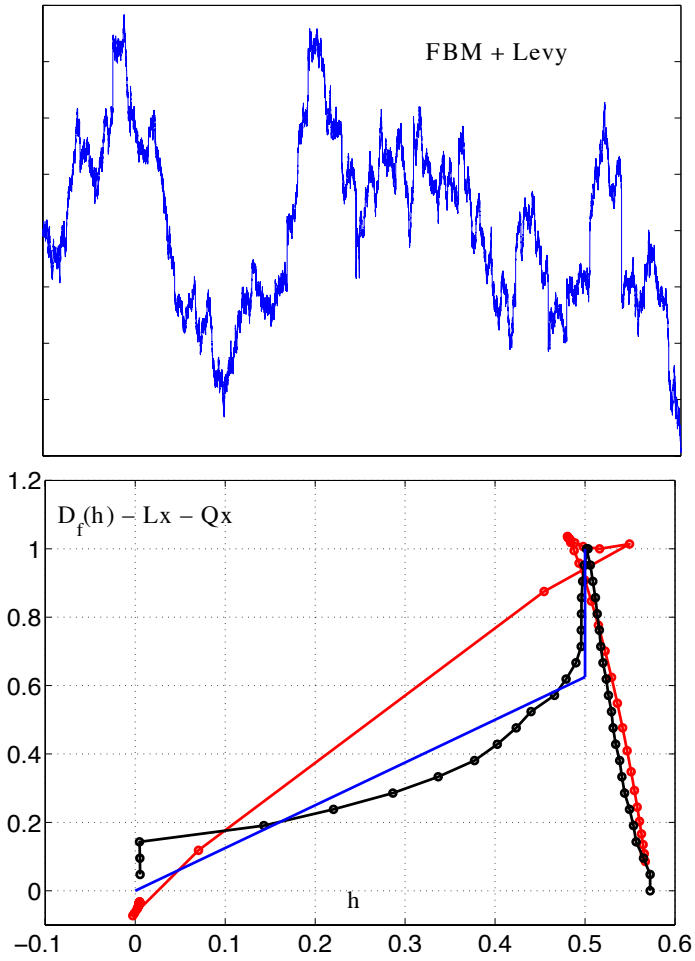

Figure 5: Lévy process which is the sum of a Brownian motion ($H = 0.5$) and pure jump Levy process ($\alpha = 1.25$). Top: single realization. Bottom: theoretical spectrum (blue), leaders Legendre spectrum (red), leaders quantile spectrum (black). Results obtained as means over 50 realizations of length $N = 2^{19}$.

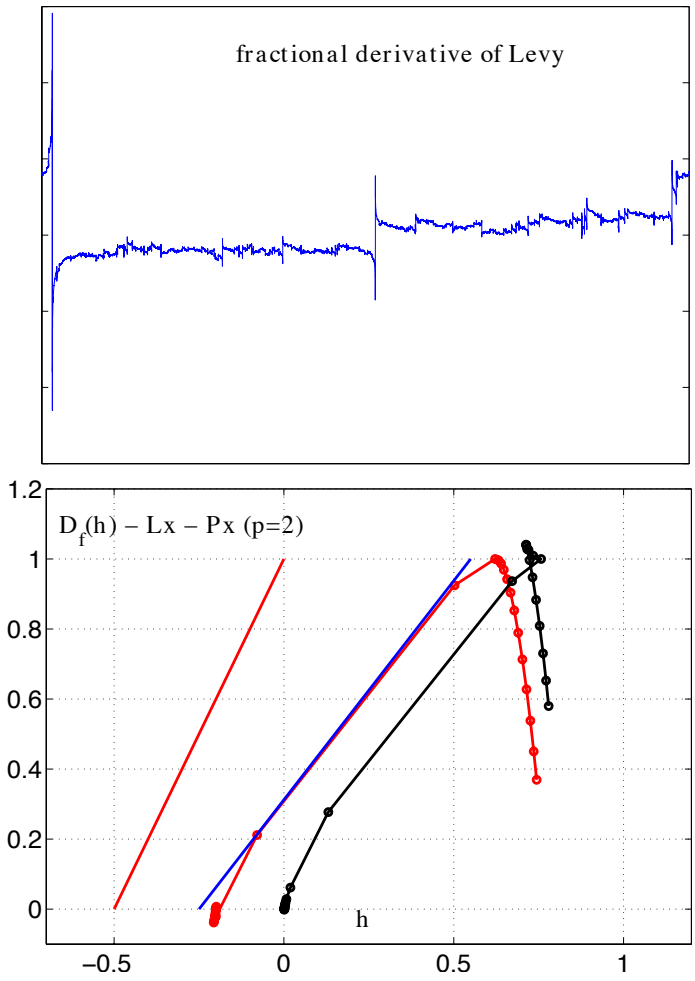

Figure 6: Fractional derivative of order $1/\alpha - H$ of Levy-stable motion ($\alpha = 1.25$, $H = 0.55$). Top: single realization. Bottom: theoretical spectrum (blue), leaders Legendre spectrum (black solid with circles), p-Leaders Legendre spectrum (red solid with circles). The red line indicates the limit of L^p spaces ($q = 2$, results obtained as means over 50 realizations of length $N = 2^{16}$).

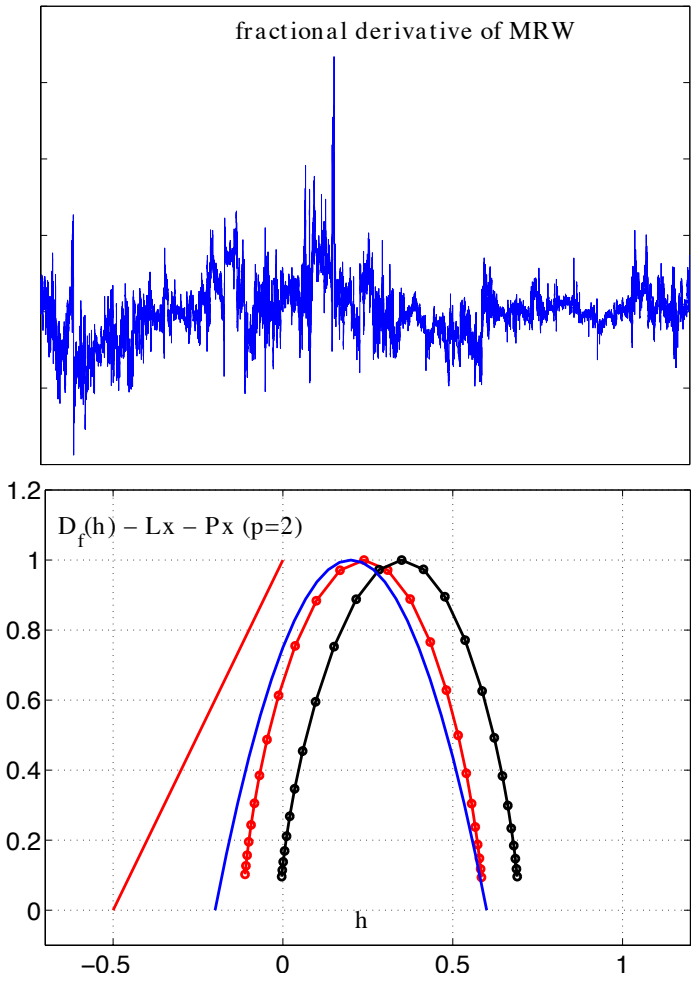

Figure 7: Fractional derivative of order 0.6 of MRW ($H = 0.72$, $\lambda = \sqrt{0.08}$). Top: single realization. Bottom: theoretical spectrum (blue), leaders Legendre spectrum (black solid with circles), p-Leaders Legendre spectrum (red solid with circles). The red line indicates the limit of L^q spaces ($q = 2$, results obtained as means over 50 realizations of length $N = 2^{16}$).