Minireview

Multifractal methodology

Hadrien Salat [a,*], Roberto Murcio [b], Elsa Arcaute [a]

[a] Centre for Advanced Spatial Analysis, University College London, London, UK
[b] Consumer Research Data Centre, University College London, London, UK

HIGHLIGHTS

- An in depth comparison between the 4 main multifractal methods is proposed.
- A didactic link is established between multifractal theory and methodology.
- The codes to implement different methods to compute multifractality are provided.

ARTICLE INFO

ABSTRACT

Various methods have been developed independently to study the multifractality of measures in many different contexts. Although they all convey the same intuitive idea of giving a "dimension" to sets where a quantity scales similarly within a space, they are not necessarily equivalent on a more rigorous level. This review article aims at unifying the multifractal methodology by presenting the multifractal theoretical framework and principal practical methods, namely the moment method, the histogram method, multifractal detrended fluctuation analysis (MDFA) and wavelet transform modulus maxima (WTMM), with a comparative and interpretative eye.

Contents

* Corresponding author.
  E-mail addresses: hadrien.salat.14@ucl.ac.uk (H. Salat), r.murcio@ucl.ac.uk (R. Murcio), e.arcaute@ucl.ac.uk (E. Arcaute).

1. Introduction

Since its introduction in the mid eighties to study turbulence signals [1–4], multifractal theory has found numerous applications such as the study of the hierarchical resistor network model [5,6], diffusion limited aggregation (DLA) models [7,8], DNA sequences [9,10], satellite and microscopic images [11–13], financial time series [14–19], urban growth and hierarchies [20–22], quantum dynamical theory [23], land use and prices [24–26], street networks [27,28], and even music [29].

From a mathematical point of view, the idea of applying fractal theory to measures was hinted by Mandelbrot as soon as 1982 [30], and was theorized more in depth later, notably by Evertsz and Mandelbrot [31], Brown et al. [32], Olsen [33], Riedi [34], and Pesin [35], and by Falconer [36]. Based on this theoretical framework, four principal multifractal methodologies have been established to solve practical problems.

The moment method was the first method to be introduced in the mid eighties [1,37,38]. It can still be considered the reference method in the field because of the simplicity of its implementation, adaptability to many types of data, as well as the existence of many variants to enhance its accuracy or computational efficiency. The histogram method [4,31,39] on the other hand improves greatly the run time over the moment method and is less reliant on error generating techniques. However, it only works for data offering a wide variety of scaling ranges. Multifractal detrended fluctuation analysis (MDFA) [40,41] is a generalization of detrended fluctuation analysis (DFA), which was originally created to detect long-range monofractal correlations in DNA nucleotide sequences [9,10]. It is used to remove artifacts created by nonstationarities in one-dimensional time series and uses the core idea of the moment method as its mechanic. The simplicity of its implementation allows to extend it to higher dimensions [13]. Wavelet transform modulus maxima (WTMM) is another method originally invented for time series [42,43]. Better suited for a generalization to higher dimensions than MDFA and able to extend, in a sense, the multifractal formalism to nonconservative and continuous phenomena, it is unfortunately more challenging to implement as it relies on a continuous framework while the three other methods are discretized.

This diversity of practical methods as well as the variety of domains and data types they can be applied to enlighten the depth of the multifractal formalism. It is also one of its drawbacks, since most methodologies have been developed independently so that multifractality lacks the unity present in some older fields. This article aims at reducing this drawback by presenting the main methodologies in a common intuitive and comparative framework. Its intent is to help making an informed choice of a multifractal methodology for someone willing to study real datasets.

The paper is organized as follows. Section 2 proposes a unified intuitive approach of the core concepts behind monofractals and multifractals, embedded in the field of measure theory. The general multifractal framework can be grasped without prior knowledge of the Hausdorff measure and the box-counting dimension, although to fully understand the details these definitions are essential. In Section 3, the main mathematical methodologies as well as the four practical multifractal methodologies mentioned above are detailed, compared and applied to binomial cascades. Each methodology is explained from the ground up and can be understood on its own, keeping in mind that the concepts explained in Section 2 help to understand how they relate to one another. The main elements of interpretation as well as the limits of multifractal analysis are discussed in Section 4. Finally, an appendix explains some of the mathematical subtleties allowing the methodology to be applied to objects that are not *explicitly* measures.

2. From monofractals to multifractals

The purpose of this section is to give a brief heuristic approach of monofractals, simply referred to as fractals, and of multifractals. All subsequent practical definitions and methodologies, as unrelated as they may appear at first glance, are only different interpretations of the core concepts explained here.

2.1. Monofractals: Characterizing space

The use of the word *fractal* in various overlapping yet different contexts makes it quite confusing for someone new to fractality. The root of all its meanings was planted by Benoit Mandelbrot [30] who coined it from the Latin word "fractus" which means "broken", as in "too irregular to fit into classical geometry". Over the years, some have restricted its use to sets which present self-similarity or to subsets whose dimension is intuitively a fraction of the integer dimension of the set they are embedded in. The reason why most of the focus has turned towards the former type of sets is that self-similarity makes most subtly different definitions equivalent and provides effective computational tricks for practical uses.

Informally speaking, given an irregular subset *D* of a space *A* whose properties are well known, the fractal analyst is usually interested in either quantifying how much of the set *A* is filled by the subset, or measuring the complexity of *D* through the scale invariance of its details. Both goals are achieved simultaneously by choosing a well adapted definition of *fractal dimension* and a method to compute it. In most practical situations, *A* will be in fact $\mathbb{R}^n$ and the chosen dimension will be the box-counting dimension. Meanwhile, the mathematician may be more interested in the Hausdorff dimension, the

Fig. 1. Fractal middle third Cantor set. From an initial segment of length 1, two sub-segments of length one third are created, and so on for each new segment, generating a self-similar fractal of dimension log(2)/log(3).

canonical measure of local size. Other, more rarely seen, definitions include the correlation dimension for sets of random points [44] and the packing dimension, a dual to the Hausdorff dimension [45].

For most rigorously self-similar subsets encountered, Hausdorff and box-counting dimensions are in fact the same thing. Finding how much of the set A is filled by the fractal subset D is the same as finding by how much one needs to grow a sub-element of the figure to find the whole figure again. This is done through the relation

$$\dim_{\mathcal{H}} = -\frac{\log(\text{number of copies})}{\log(\text{scaling factor})}, \tag{1}$$

where $\dim_{\mathcal{H}}$ is the Hausdorff dimension. Because of this, self-similarity is often treated as a synonym of fractality in the literature. By extension, the word "fractal" is also used for phenomena described by *self-similar* functions, i.e. functions $f : D \subset \mathbb{R}^n \to \mathbb{R}$ for which there exists an α such that

$$\forall \lambda \in \mathbb{R}, x \in D, f(\lambda x) = \lambda^{\alpha} f(x), \tag{2}$$

and in particular power-laws [46], which are, in a sense, representations of scale invariance within the space D.

On the other hand, any dense subset made of a countable number of points is of dimension 0 for the Hausdorff dimension and of dimension equal to its closure for the box-counting dimension. For example $\mathbb{Q} \cap [0, 1]$ in $\mathbb{R}$ is of dimension 0 for the Hausdorff dimension and of dimension 1 for the box-counting dimension. Those definitions are therefore far from equivalent in all generality. A detailed discussion on what elements are desirable to define a suitable fractal dimension can be found in chapter 3 from Falconer [36].

A simple example is given in Fig. 1. The middle third Cantor set is created from an initial segment of length 1, from which two sub-segments of length one third are extracted. This process is then repeated for each new segment, and so on. Since the resulting set is self-similar with two new copies of itself each scaled at a ratio of 1/3, one would get from Eq. (1) a Hausdorff dimension of log(2)/log(3). Calculating directly the Hausdorff dimension without using Eq. (1) is more involving than one would expect even for such a simple set, hence the motivation to restrict the notion of fractals to self-similar sets. The value log(2)/log(3) represents how much of the initial segment is still present in the Cantor set after an infinite number of iterations of the generating process.

2.2. Multifractals: A theory of measures

While monofractals are mostly concerned with spaces, multifractals deal with measures.[1] Even if the idea behind multifractals is also to study the complexity and reveal the scaling properties of a mathematical object, those two concepts are distinct. Indeed, a measure can be a multifractal despite its support not being a monofractal [47]. Let us consider a subset $D \subset \mathbb{R}^n$ on which are defined:

- a "fractal" measurement method $\mathcal{M}$;
- a finite measure μ which we want to study.

Here, $\mathcal{M}$ can be any method providing a way to compute a monofractal dimension, such as those quoted in the previous section, as deemed appropriate for the nature of the space D.

A *multifractal measure* μ on D is characterized by a distribution such that around any $x \in D$, the measure in a ball of radius r around x scales with r, i.e. is proportional to r^{α} for some α, provided r is small enough, and such that the sets formed by all points around which the scaling exponent is the same are monofractals for $\mathcal{M}$. The fractal dimension of the set corresponding to the local exponent α is usually denoted $f(\alpha)$.

Most methods $\mathcal{M}$ are based on defining a self-similar local measurement $\mathcal{M}_r$, such as the number of boxes of radius r necessary to cover the set for the box-counting dimension or the quantity $\mathcal{H}_r^s$ in the definition of the s-dimensional Hausdorff measure (see Falconer [36] for a definition). In that case, the multifractality of μ is equivalently characterized by a distribution such that the two following fundamental scaling relations hold for r small enough:

---

[1] Here, *measure* is to be understood as the mathematical notion. In practice, data can be mathematical measures (a probability measure, for example) or any value distribution that can be transformed into a measure (a weighted set, a fractal landscape, a time-series, for example, or even a function using WTMM), see the appendix.

1. $\mu_r(x) \sim r^{\alpha_x}$ for an α_x around any $x \in D$, where $\mu_r(x)$ is the measure in a ball of radius r around x;
2. $\mathcal{M}_r(\alpha) \sim r^{-f(\alpha)}$ for an $f(\alpha)$, where $\mathcal{M}_r(\alpha)$ is the $\mathcal{M}_r$-measurement of the set $\{x, \alpha_x = \alpha\}$.

The *multifractal spectrum* is the curve $f(\alpha)$ against α. It gives, roughly speaking, the "fractal dimension" $f(\alpha)$ of sets where the measure scales locally with the same exponent α. *Multifractal analysis* should be understood as a method to characterize and compare measures defined on D when they present enough scaling properties to alleviate the intrinsic complexity of (D, μ).

An example is given in Fig. 2. The middle third Cantor set is made multifractal by weighting every right sub-interval twice as much as every left sub-interval, the total weight being normalized to 1 at each step. The first three steps of this process are illustrated in the top figure. Denote by r_k the size of the new sub-intervals at step k, and fix $r_0 = 1$. Then, at step $k = 3$, height sub-intervals are obtained, each of size $r_3 = (1/3)^3$ and carrying a weight that can be expressed as r_3^α for some α. At this macroscopic state, a broad $\mathcal{M}$ can be defined such that $\mathcal{M}_{r_k}(\alpha)$ denotes the number of sub-intervals scaling with r_k for an exponent α. This number can be in turn expressed as $r_k^{-f(\alpha)}$. For the particular α chosen in Fig. 2, that is $\alpha = 1 - \frac{\log(2)}{3\log(3)}$, there are $3 = (1/3^3)^{-1/3}$ sub-intervals carrying this measure, hence $f(\alpha) = 1/3$. By repeating this calculation for each of the four different weights carried by the sub-intervals at step $k = 3$, the spectrum corresponding to the bottom line of the bottom figure is obtained.

Of course, at such a low level of iteration, $\mathcal{M}$ does not make much sense. But as k grows to infinity, the spectrum resulting from this $\mathcal{M}$ converges to the actual spectrum one would obtain for the Hausdorff measure and the proper totally disconnected weighted middle third Cantor set. The first 500 iterations are illustrated in Fig. 2. It can be noted that the multifractality comes from the measure created by the weights, not from the physical support itself which is only the monofractal Cantor set presented in the previous section. In particular, the dimension of the support, here $\log(2)/\log(3)$, can be found at the peek of the spectrum.

3. Multifractals in the field

In this section, different ways of defining the "multifractal spectrum" are given, along with methods to compute it wherever possible. Those are all equivalent definitions in the sense that they convey the same intuitive idea of providing a fractal dimension to iso-scaling sets in data. As such, it is expected that the spectra resulting from each of them will share the same symmetries when applied to a particular dataset. They are not equivalent however on a more precise level, and the resulting spectra may differ in width or lead to an overshooting of the $f(\alpha)$ value. The first formal definition should be considered as canonical and subsequent definitions may be considered as approximations of it.

Methods are sorted in three groups depending on the type of data best suited to support the studied measure: mathematical abstract sets, singular measures and two-dimensional data such as maps and images, and one-dimensional (discrete or continuous) time series. Nonetheless, these methods can usually be extended easily to measures supported by any subset of $\mathbb{R}^n$. Methods from the two latter groups are tested against two samples obtained from binomial cascades. The first sample corresponds to a theoretical binomial cascade of parameter $p = 0.6$, that is the measure one would obtain after averaging over an infinite number of random walks through the cascade, while the second one corresponds to one particular realization of a random walk through the cascade.

3.1. Abstract sets—Formal definitions

Falconer [36] distinguishes two variants of spectra of particular interest for mathematicians: The *singularity spectrum*, which is the most canonical definition and encompasses the universality sought in mathematics, and the *coarse spectrum*, which is more adequate for practical purposes.

Consider a topological space D and a finite measure μ on D. The local scaling exponent α_x of μ at $x \in D$ is given by the Hölder dimension $\dim_{loc}$, defined by

$$\dim_{loc}\mu(x) := \lim_{r \to 0} \frac{\log \mu(B(x, r))}{\log r}, \tag{3}$$

where $B(x, r)$ is the ball of center x and radius r for the topology of D. The *singularity spectrum* is then defined by the function

$$f_{\mathcal{H}}(\alpha) := \dim_{\mathcal{H}} \{x \in D, \; \dim_{loc}\mu(x) = \alpha\}, \tag{4}$$

where $\dim_{\mathcal{H}}$ is the Hausdorff dimension.

Note that the Hausdorff dimension is chosen for $\mathcal{M}$ instead of box-counting since $\{x \in D, \; \dim_{loc}\mu(x) = \alpha\}$ is often dense in the support of μ, in which case box-counting would give a constant spectrum equal to the dimension of the support of μ.

Let us now consider an r-mesh grid covering D and count the number of cells for which μ is roughly r^α. Define,

$$N_r(\alpha) := \# \{r\text{-mesh cubes } C, \; \mu(C) \geq r^\alpha\}, \tag{5}$$

where # stands for "the number of". Provided the limits do exist, the *coarse spectrum* is defined by the function

$$f_C(\alpha) := \lim_{\varepsilon \to 0} \lim_{r \to 0} \frac{\log^+(N_r(\alpha + \varepsilon) - N_r(\alpha - \varepsilon))}{-\log r}, \tag{6}$$
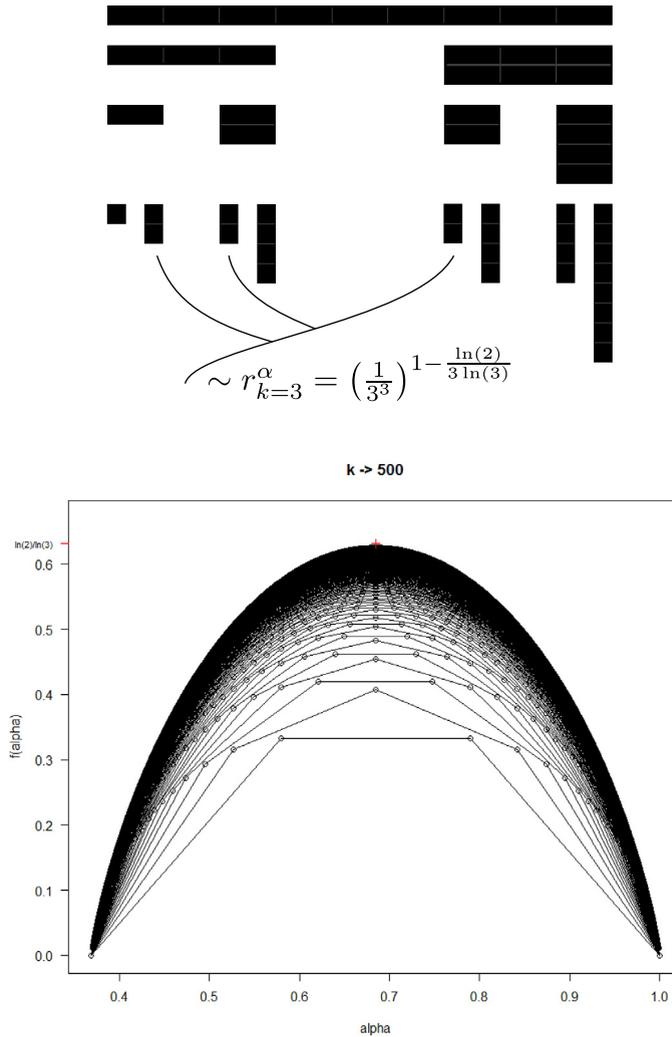
where $\log^+(\cdot)$ stands for $\max(\log(\cdot), 0)$.

$$\sim r^{\alpha}_{k=3} = \left(\frac{1}{3^3}\right)^{1-\frac{\ln(2)}{3\ln(3)}}$$

k -> 500

Fig. 2. Multifractal middle third Cantor set. On the top, the first three iterations of the generation of the weighted multifractal Cantor set are represented, while on the bottom the first 500 spectra corresponding to each successive iteration for $\mathcal{M}$ are plotted.

When f_C does exist, then for all α,

$$f_{\mathcal{H}}(\alpha) \leq f_C(\alpha),$$
(7)

and the equality holds true for self-similar measures (Proposition 17.9 from Falconer [36]). When f_C does not exist, one can define the *lower* and *upper* spectra by

$$\underline{f}_C(\alpha) := \lim_{\varepsilon \to 0} \liminf_{r \to 0} \frac{\log^+ (N_r(\alpha + \varepsilon) - N_r(\alpha - \varepsilon))}{-\log r},$$
(8)

and

$$\bar{f}_C(\alpha) := \lim_{\varepsilon \to 0} \limsup_{r \to 0} \frac{\log^+ (N_r(\alpha + \varepsilon) - N_r(\alpha - \varepsilon))}{-\log r}.$$
(9)

In that case, according to lemma 17.3 of Falconer [36],

$$f_{\mathcal{H}}(\alpha) \leq \underline{f}_C(\alpha) \leq \bar{f}_C(\alpha).$$
(10)

An example is given in Fig. 3 for a binomial cascade of parameter $p = 0.6$. The binomial cascade is a simpler version of the multifractal middle-third Cantor set introduced in Section 2.2. Here, an original interval of size 1 is divided into two sub-intervals of length $1/2$ carrying a probability 0.6 for the left one and 0.4 for the right one. This process is then iterated
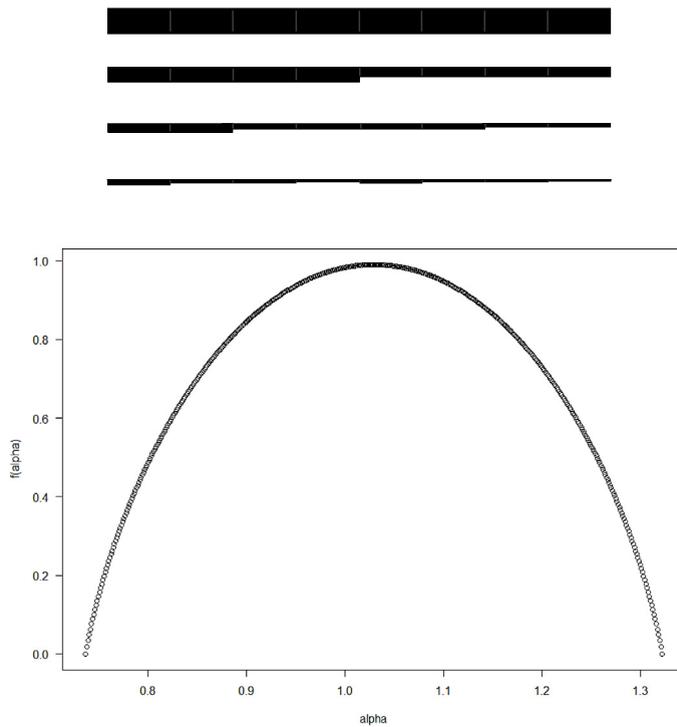
Fig. 3. Multifractal binomial cascade. On the top, an original interval of size 1 is divided into two sub-intervals of length 1/2 carrying a probability 0.6 for the left one and 0.4 for the right one. On the bottom, the corresponding multifractal spectrum.

on each resulting sub-intervals, and so on. The resulting singularity spectrum $f_{\mathcal{H}}$ in the bottom figure is computed using the same trick as in Section 2.2 and is identical to the coarse spectrum $f_C = f_{\underline{C}}(\alpha) = \bar{f}_C(\alpha)$, since the measure is self-similar. Here, the fractal dimension of the support is 1 since the iterative process does not create "holes" in the initial segment.

3.2. Spatial data—Moment and histogram methods

Moment and histogram methods are the elementary components of practical multifractal spatial data and image analysis. Both methods rely on counting the measure at different levels of aggregation and include a multitude of variants depending on the aggregation method chosen. The most basic way to aggregate consists in applying square grids of increasing resolutions to the data (box-counting). Instead of grids, one can use ball neighborhoods of increasing radius as they can be easily calculated for two-dimensional geographic data by GIS software, or gliding boxes to increase the number of data points. Grids can also be based on any regular unit shape, such as diamonds or equilateral triangles, instead of squares, to enhance computational complexity or suitability with the data. Other variants exist to counter the reliance on error generating techniques such as linear fits or Legendre transforms. The core mechanics of both methods will be detailed and reference to the literature will be given for their variants.

Box-counting techniques are meant to give correct results for "static" singular measures but may not be adapted to continuous data and to nonconservative phenomena like turbulent velocity profiles (see Kestener and Arneodo [51]). For those, wavelet based techniques (see Section 3.3 below) are better suited.

Let us consider a mesh grid of unit r covering a domain D, and a phenomenon occurring N times in D. Denote

$$p_i := N_i/N = \int_{i^{th} box} d\mu(x) \tag{11}$$

the probability that an instance of the phenomenon occurs in the ith box, where μ is the corresponding probability measure. To interpret the two scaling rules from Section 2.2, one simply needs

1. $p_i \sim r^{\alpha_i}$;
2. $N(\alpha_i) \sim \rho(\alpha_i)d\alpha_i r^{-f(\alpha_i)}$,

where $N(\alpha_i)$ is the number of times α falls in each interval $[\alpha_i, \alpha_i + d\alpha_i]$, and ρ is a density function used to take into account the dimension of D.

To effectively compute f, the trick generally used is the *moment method* [1,37,38]. By raising p_i to its moment p_i^q for different q, one can force only one value of alpha for each q to make a significant contribution to the total value of the measure. Consider

$$Z(q) := \sum_i p_i^q \sim \sum_i r^{\alpha_i q}$$

$$\sim \int_\alpha N(\alpha) r^{\alpha q}$$

$$\sim \int \rho(\alpha) r^{\alpha q - f(\alpha)} d\alpha, \tag{12}$$

then, for r small enough, the value of $Z(q)$ is almost entirely given by the α such that

$$\tau(q) := \alpha q - f(\alpha) \tag{13}$$

is minimal. Let us call $\alpha(q)$ this value of α. It is easy to show by a Legendre transform that the minimality condition yields

$$\alpha(q) = \frac{d\tau(q)}{dq}; \tag{14}$$

and

$$f(\alpha(q)) = \alpha(q)q - \tau(q), \tag{15}$$

so that computing $\tau(q)$ from $Z(q)$ for each q between $-\infty$ and $+\infty$ is enough to obtain the full spectrum.

It is not possible in practice to use an infinite range of values for q, nor is it desirable since the method becomes less and less accurate for extreme values of q. To select an appropriate range of q, one should set a threshold for the error generated by linear fitting and dismiss all the values of q for which the threshold is exceeded. It is also necessary to select q in order to ensure that $f(\alpha) > 0$ and that the *generalized dimension D_q*, defined by $D_q := \tau(q)/(q-1)$, remains lower than the dimension of the physical support of the phenomenon. The reason for that last constraint will become clear when the physical meaning of D_q will be explained in Section 4.

In practice, $\tau(q)$ is found directly as the slope in a log–log plot of $\sum_i \mu_i^q(r)$ versus r obtained for different grid sizes r, where $\mu_i(r)$ is the total measure of cell i of size r. Since this slope is independent of the normalization of the measure μ, μ does not need to be weighted as a probability measure. Explicitly, $\tau(q)$ is found as the limit

$$\tau(q) = \lim_{r \to 0} \frac{\log(\sum_i \mu_i^q(r))}{\log(r)}. \tag{16}$$

In accordance with the idea that practical methods tend to create overshooting, one finds that

$$f_{\mathcal{H}}(\alpha) \le \underline{f}_{-C}(\alpha) \le \bar{f}_C(\alpha) \le f_M(\alpha), \tag{17}$$

where f_M is the spectrum resulting from the moment method (corollary from Proposition 17.2 from Falconer [36]).

Both finding $\tau(q)$ through linear fitting and applying numerical Legendre transforms have a cost on the accuracy of the results. The possibility of averaging over several samples can be extremely beneficial. There are two ways of doing this: averaging over a range of $\{f_j(\alpha_i)\}_j$ computed independently for different samples j, or averaging first over a range of $\{N_j(\alpha_i)\}_j$ and then deducing the corresponding $f(\alpha_i)$ from the relation $N(\alpha) \sim \rho(\alpha) d\alpha r^{-f(\alpha)}$ on the averaged values. The first solution guarantees to obtain a "classic" positive spectrum, but it can be unreliable if the fluctuations between the $f_j(\alpha_i)$ are too important. The second solution is more reliable, but may create an artificial negative part in the spectrum if $N(\alpha_i)$ falls below 1 for some α_i as a result of the averaging process.

Chhabra and Sreenivasan [48] argue that this artificial negative part can still be of relevance when a strong underlying probabilistic process is suspected either as a cause of the phenomenon or as a result of the experimental methodology since it could describe the rarely occurring events. Unfortunately, since $\alpha \mapsto N(\alpha)$ decreases exponentially compared to $\alpha \mapsto f(\alpha)$ in the negative part, one would need an exponentially increasing number of samples as the resolution gets smaller to maintain accuracy while supersampling. Paradoxically, for a constant number of samples, a better resolution would mean a less accurate result.

To tackle this problem, Chhabra and Sreenivasan [48] propose a multiplier method. The self-similarity of the measure implies the existence of an underlying scale-invariant multiplier distribution such that the α_i at resolution r_k are only the result of k composition of said multipliers. If there is no correlation in the underlying probabilistic process from resolution r_{k-1} to resolution r_k for some k, then one can deduce the multipliers and hence α. In particular, if all levels of resolution are uncorrelated, one can choose $k = 1$, otherwise, one should choose the smallest k for which a level of resolution is uncorrelated to the previous one.

Denote r_0 the minimal resolution and r_k the resolution chosen as described above. Then, define
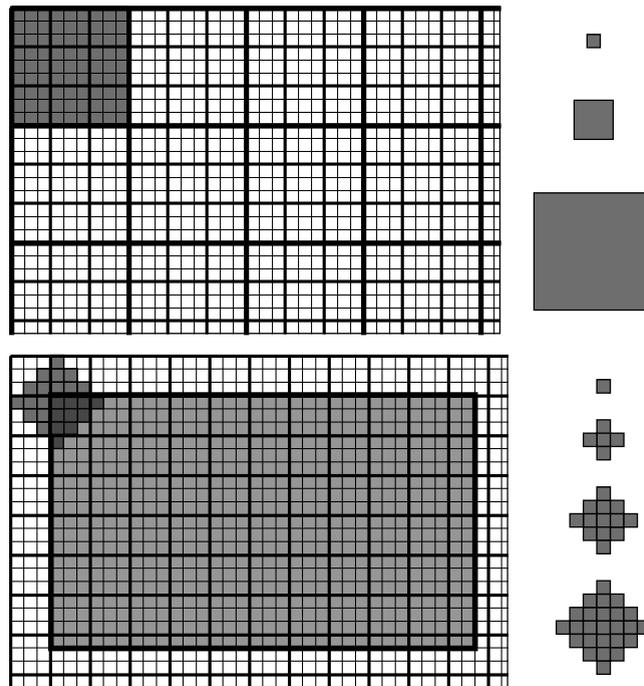
$$r := r_k/r_0, \tag{18}$$

Fig. 4. Comparison between grid and gliding box upscalings. In the top image, the third level of aggregation, which is only 3^2 times the smallest resolution, only allows to fully place 8 boxes on the figure. In contrast, gliding boxes applied on the lowest image maintain 608 sample points at the cost of forcing to remove a wide border from the analysis.

and for each sample j and box i,

$$M_{ij} := \mu_{ij}(r_0)/\mu_{ij}(r_k). \tag{19}$$

Then, according to Chhabra and Sreenivasan [48], $\tau(q)$ and $\alpha(q)$ are given by

$$\frac{1}{N(r)}\tau(q) + d \approx -\frac{\log\left(1/N(r)\sum_{i,j} M_{ij}^q\right)}{\log(r)}; \tag{20}$$

$$\frac{1}{N(r)}\alpha(q) \approx -\frac{\sum_{i,j} M_{ij}^q \log(M_{ij})}{\sum_{i,j} M_{ij}^q \log(r)}. \tag{21}$$

where $N(r)$ is the number of non zero values of M_{ij} and d is the dimension of the physical support D.

Chhabra and Sreenivasan have shown that for a binary cascade and the dissipation field of fully developed turbulence in the atmospheric surface layer, using the multiplier method allowed to expand the negative part of the spectrum and make it converge to the theoretical result more rapidly than the supersampling method, with also a gain in computational complexity.

Another way to expand the set of sample points consists in using one grid and aggregate with a gliding box for different radii of said gliding box instead of using different grid sizes [11,49]. In that case, $\tau(q)$ is given by taking the limit for $r \to 0$ in

$$\frac{1}{N(r)}\tau(q) + d \approx \frac{\log\left(1/N(r)\sum_{i=1}^{N(r)} \mu_i^q(r)\right)}{\log(r)}, \tag{22}$$

where $N(r)$ is the number of gliding boxes of size r with non zero measure, $\mu_i^q(r)$ is the measure inside the ith gliding box, and d is the dimension of the physical support D.

Since gliding boxes need not be mutually exclusive, contrary to squares from a mesh grid, the number of values contributing to the analysis remains that of the smallest resolution at all scales. The trade-off is that only boxes which are completely bounded in D should be included, so that only the "inner portion" of the data can be analyzed, or the object of study needs to be surrounded by a large neighborhood of known values (see Fig. 4). Using gliding boxes allows a higher raw number of sample points at the cost of restricting the range of study. It is of course possible to join gliding boxes and the multiplier method by adapting the definition of μ_{ij} and $N(r)$ in Eqs. (19) and (20).

Chhabra et al. [3] propose a recipe to avoid the Legendre transform of $\tau(q)$ when the measure arises from multiplicative processes. Once the p_i have been established, compute

$$\mu_i^q(r) = \frac{p_i^q}{\sum_j p_j^q}. \tag{23}$$

Then, the Legendre transform can be directly integrated in the calculation of f and α through the formulas

$$f(q) = \lim_{r \to 0} \frac{\sum \mu_i^q(r) \log \mu_i^q(r)}{\log(r)}; \tag{24}$$

$$\alpha(q) = \lim_{r \to 0} \frac{\sum \mu_i^q(r) \log p_i(r)}{\log(r)}. \tag{25}$$

Note that here $\alpha(q)$ is the average value of α at resolution q. Unfortunately, this recipe does not remove the need for linear fitting when calculating the limits, which is usually the main cause of error. It was applied to two-scaled cantor measures by Chhabra and Jensen [50] with good results when the boxes size progression matched the sub-intervals size progression, and "satisfactory" results otherwise despite the errors created by linear fitting. It was also found by Chhabra et al. [3] that the result of this direct computation were in good agreement with those obtained from Legendre transforming $\tau(q)$ for fully developed turbulence.

Kestener and Arneodo [51] argue that the moment method when applied to multifractal measures implicitly supposes that the underlying cascading process is conservative across all scales (which translates into $\tau(q) = (q - 1)D_q$, and, in particular, $\tau(1) = 0$). They indicate that one can still detect and quantify the multifractal properties of measures that have a cancellation exponent $\tau(1) < 0$ by using the WTMM method (defined in Section 3.3). Note, however, that a nonconservative measure is not a properly defined mathematical measure. The analysis of such a measure could be seen instead as the analysis of a sequence of measures approximating one "underlying" measure defined at scale $r \to 0$.

To conclude with the moment method, note that the function Z from Eq. (12) is similar to the partition function, usually denoted $Z(\beta)$, in thermodynamics. Likewise, the quantities τ, α, and $f(\alpha)$ can be interpreted respectively as the free energy, the energy, and the entropy. For more details on this analogy, see Stanley and Meakin [47], Tel [52] and Arneodo et al. [53].

A different direct approach is the *histogram method*, see for example Meneveau and Sreenivasan [4] and Arneodo et al. [39]. The idea consists in finding the cells with extremal values of total measure for different grid resolutions, and dividing the distance between those values into regular intervals to exploit the fact that exactly one value of α and $f(\alpha)$ will correspond to an extremity of one of the new sub-intervals.

Let us call μ_i^k the total measure of cell i of a grid of unit r_k, and $N(X)$ the number of boxes presenting feature X. Step by step, the method breaks down as follows.

1. Compute $X_i^k := \log(\mu_i^k)$ for each cell i of different grids of unit r_k;
2. Divide $[X_{min}^k, X_{max}^k]$ regularly in n smaller intervals for each k, where $X_{min}^k := \min\{X_i^k\}$ and $X_{max}^k := \max\{X_i^k\}$;
3. Deduce one value of α and $f(\alpha)$ from the slopes of X^k and $N(X^k)$ versus $\log(r_k)$ for each sub-interval;
4. Repeat for different grid positions to get a better estimate.

In step 3, for $1 \leq j \leq n$, the value α_j is given by the slope of $X_{j,k}$ versus $\log(r_k)$ where $X_{j,k}$ is one of the extremities of the jth sub-interval for grid resolution r^k. According to Meneveau and Sreenivasan [4], the correct normalization of the total measure leads to an expression of $f(\alpha)$ as the slope of $\log(N(X_{j,k}) \Delta X^{1/2})$ versus $\log(r_k)$, where $N(X_{j,k})$ is the number of boxes of size r_k containing an X falling in the interval of size ΔX around $X_{j,k}$.

It is indeed a problem to find the correct normalizations of α and $f(\alpha)$ because the relations $p_i \sim r^{\alpha_i}$ and $N(\alpha_i) \sim \rho(\alpha_i)d\alpha_i r^{-f(\alpha_i)}$ depend on prefactors that are unknown *a priori*. Such a problem is absent in previous methods since those factors are canceled while taking the limit for $r \to 0$, which is not done here.

Meneveau and Sreenivasan [4] have applied the histogram method in for a binomial measure, a period doubling attractor for a specific logistic map and the dissipation field of turbulent kinetic energy in turbulence flows. They found good agreement with the results obtained from the moment method for the first two cases but it was evidenced that errors are generated by the histogram method for measures with small scaling ranges such as the third case. In fact it was evaluated that the exponent found by this method is only accurate up to order $\log(L/r)^{-2}$, where L represents a characteristic value intrinsic to the problem (a translation of the unknown prefactor). It was recommended to use this method only for measures such that the largest obtainable scale is at least 10^3 times bigger than the smallest measurable scale.

The spectra resulting from moment and histogram methods applied to binomial cascades are given in Figs. 5–7. In Fig. 5, the plots on the top are obtained for the theoretical binomial cascade of probability $p = 0.6$ introduced in Section 3.1, that is the result of averaging an infinite number of random walks through the iterative process defining the cascade. The plots on the bottom are obtained for one particular realization of a random walk through the cascade. The iterative process is stopped after the 20th step for which over a million sub-intervals have already been created. This is done to ensure a reasonable run time and use of memory, and because it is comparable to the size of many encountered data types such as pixels of an HD image or data collected from a city-size human settlement. Unfortunately, it also means that the studied cascades are not equivalent to the one used in Fig. 3, for which the iterative process is repeated an infinite number of times. As such the
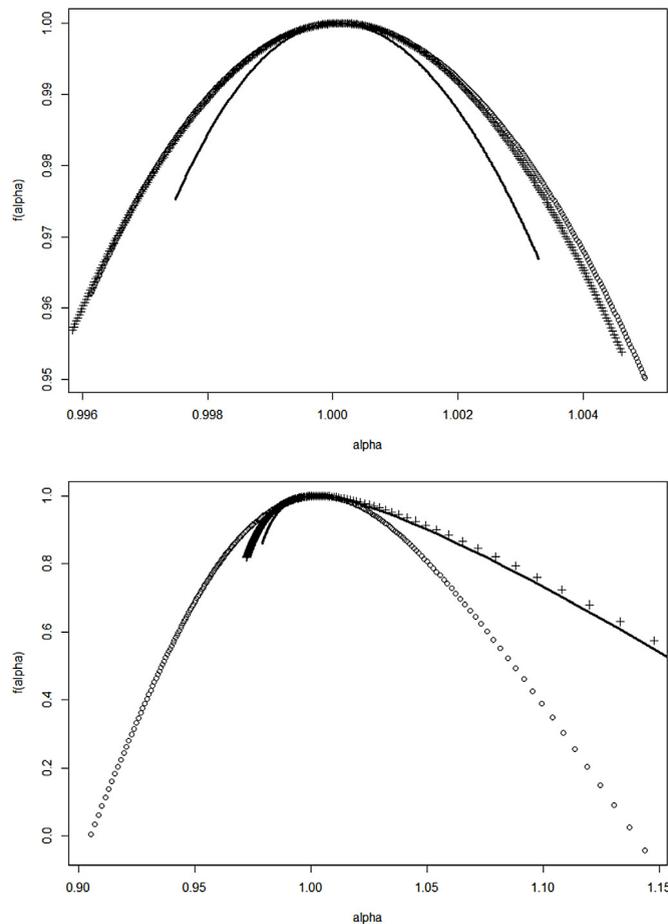
Fig. 5. Moment method and variants applied to multifractal binomial cascades. The standard moment method (plain line), the gliding box method (circles), and the multiplier method (crosses) are applied to the theoretical cascade on the top and to a particular random walk through the cascade on the bottom.

resulting spectra are expected to be similar, but not necessarily identical to the theoretical curve of Fig. 3 depending on the sensitivity of the chosen multifractal method. The chosen range of q for the moment method goes from -20 to 20.

The results of the standard moment method and its variants, the gliding box and the multiplier methods, are illustrated in Fig. 5. In the theoretical case on the top, the range of α is a lot narrower than what is expected based on the reference curve (Fig. 3). The variants help improve the situation, but not by much. This problem is due to the fact that the iterative process was stopped too soon. The top figure of Fig. 6 evidences that stopping the iterative process at step 25 results in a wider spectrum (circles) than stopping it at step 10 (triangles). Aside from this problem, the resulting spectrum is similar to the reference one and a simple rescaling of the range of α is enough to make both spectra harmonious.

For the particular realization of the cascade on the bottom of Fig. 5, the accuracy of the standard and gliding box moment methods deteriorates rapidly for negative q. The multiplier method gives the best results overall. To keep the curve above zero, the range of q had to be restricted to $q \geq -4$ for the first two variants and to $17.5 \geq q \geq -10$ for the multiplier variant.

Results of the histogram method applied to the binomial cascades can be found in Fig. 7. Unfortunately, for such a small range of scaling, the histogram method is not well adapted and the resulting spectra are not smooth. The error generated by the method makes results difficult to interpret in this case. It has however the advantage of being faster than the moment method and gives a range of α closer to the reference in this particular case.

3.3. Time series - MDFA & WTMM

Some methods have been independently developed for the specific purpose of studying time series, an important object in physics. They are therefore particularly well suited for one-dimensional data, but can be extended to any dimension at the expense of computational complexity. For our purpose, time series will be defined as one dimensional arrays of discrete values representing observations taken at regular intervals for the MDFA method and more generally as one-dimensional continuous functions for the WTMM method.
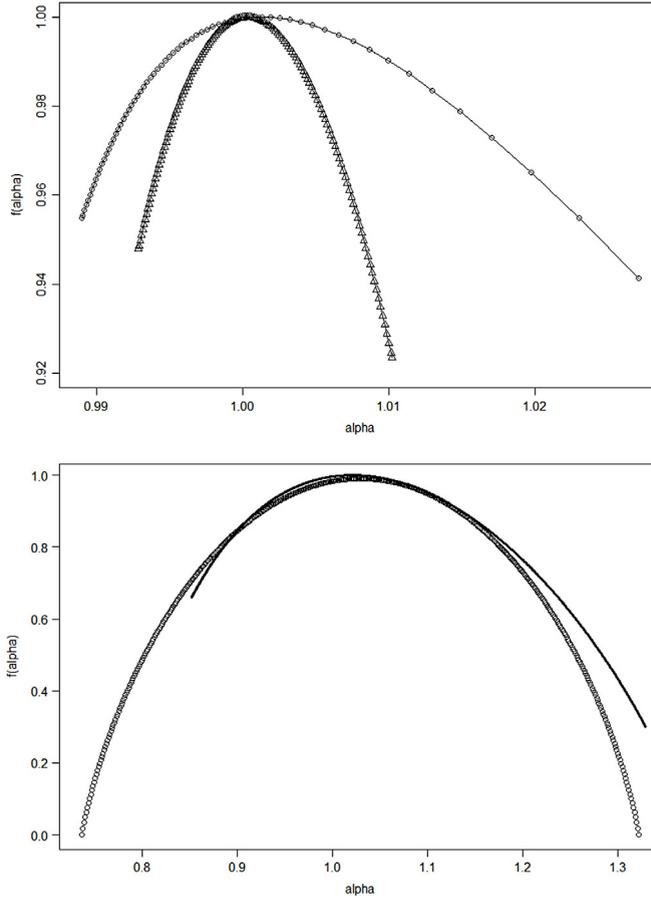
Fig. 6. Influence of the number of iterations on the spectrum. On the top, the spectrum made of triangles is obtained for 10 repetitions of the iterative process generating the cascade, while the spectrum made of circles is obtained for 20 repetitions. Increasing the number of iterations makes the spectrum larger and therefore closer to the theoretical curve. On the bottom, a rescaling of the range of alpha on the curve resulting from the moment method (plain line) is enough to make it harmonious with the reference curve (circles).

Multifractal Detrended Fluctuation Analysis (*MDFA*) is thoroughly described by Kantelhardt et al. [40] and Kantelhardt [41]. In the basic approach, time series are first sub-divided into smaller segments on which is subtracted a least-squares best-fit polynomial of a chosen order to remove the artifacts created by nonstationarities in the time series. A method similar to the moment method is then applied to the resulting detrended series. In details, MDFA consists of the following steps.

1. Replace a time series $f(\cdot)$ with its cumulative sum $F(\cdot)$;
2. Divide F into N_s segments containing s elements each for an array of s;
3. Detrend by removing a least-squares fitted polynomial of order n to F on each segment;
4. Denoting $\bar{F}$ the result of step 2, compute

$$F_q(s) := \left(\frac{1}{N_s} \sum_{v=1}^{N_s} \bar{F}(v,s)^q \right)^{1/q} ;$$

5. Find the scaling relation $F_q(s) \sim s^{h(q)}$.

Here, $h(q)$ is the *hurst exponent*, which relates to the classical $\tau(q)$ through the relation $\tau(q) = qh(q) - D_f$, where D_f is the fractal dimension of the physical support of f.

Step one ensures that the transformed series F is σ-additive, which is a necessary property to define a measure that the original series f may not possess. It also allows the use of simple polynomials of the form $a_n i^n + \cdots + a_0$ with $i \in \mathbb{N}$ to detrend in step three. It should be noted that the expression of F_q given above is not well defined for $q = 0$. It is indeed necessary to set

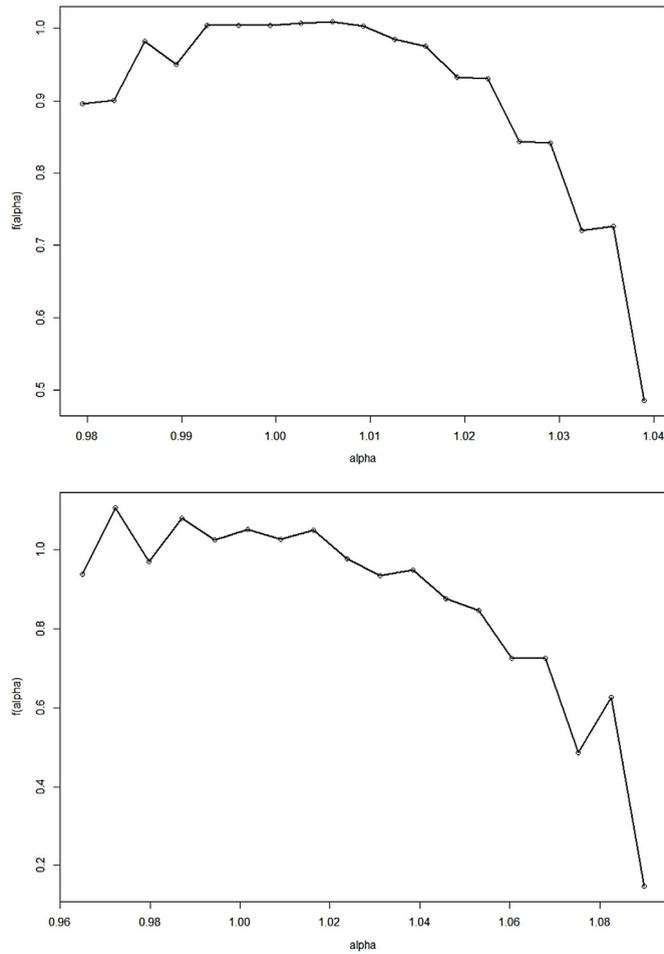$$F_0(s) = \exp\left(\frac{1}{N_s} \sum_{v=1}^{N_s} \log\left(\bar{F}(v,s) \right) \right). \tag{26}$$

Fig. 7. Histogram method applied to binomial cascades. On the top, a theoretical binomial cascade of parameter $p = 0.6$ and on the bottom a particular realization of it.

According to Kantelhardt et al. [40], MDFA works only for positive h and becomes inaccurate for h close to 0. A solution consists in integrating by considering the sum $\sum F(\cdot, s)$ instead of F. Following the same steps, one would obtain $h(q) + 1$ instead of $h(q)$.

The use of $\tau(q)$ as an intermediary step is given to link the method to previous techniques (see Eqs. (14) and (15)), but one can compute directly α and $f(\alpha)$ using the expressions

$$\alpha(q) = h(q) + q\frac{dh(q)}{dq};$$ (27)

$$f(\alpha(q)) = q(\alpha - h(q)) + D_f.$$ (28)

An original application of MDFA, presented by Kantelhardt et al. [40], is to distinguish the underlying cause of multifractality between long-range correlations and a broad probability density function. Indeed, if one shuffles the time series, all correlations are destroyed. Hence, when applying MDFA to the shuffled time series, if the resulting spectrum shifts towards monofractality, that is if the shuffled h is constant, then multifractality is probably due to long range correlations. Obviously, multifractality can have several causes, but an important influence of correlations should be noticeable in the alteration of h.

MDFA can be extended to 2 or more dimensions using multivariate polynomials, as proposed by Gu and Zhou [13]. The 3 dimensions extension is particularly useful to study the evolution of a two-dimensional spatial pattern simultaneously in space and time. Unfortunately, the necessity to choose a common array of s for all directions at the same time, and therefore constraining the precision and accuracy of the method to the direction along which the data is the most scarce or irregular, as well as the rapidly growing computational complexity are significant limiting factors.

Wavelet Transform Modulus Maxima (WTMM) replaces square intervals (and square grids for higher dimensions) by highly customizable wavelets. If those are chosen orthogonal to low order polynomials, a natural detrending happens. The "modulus

maxima" part of the name refers to an observation that analyzing the data along maxima lines is enough to bring out the underlying multifractal structure. With a judicious choice of wavelets, one can therefore integrate the advantages of MDFA and extend it efficiently to higher dimensions while maintaining adequate computational complexity. The wavelet transform idea is introduced by Muzy et al. [42] and a more extensive study of WTMM can be found from Arneodo et al. [43].

Consider a function $f : \mathbb{R} \to \mathbb{R}$ representing either a continuous signal or the interpolation of a discrete time series and a wavelet ψ orthogonal to low-order polynomials. Note that the function f is not a mathematical measure, but represents instead the density of a measure. It is implicitly transformed into a measure when integrated in the first step below. The wavelet is a real valued function, preferably of zero mean to ensure the method is invertible. WTMM is divided in the following steps.

1. Operate the wavelet transform by defining for any x_0:

$$T_\psi[f](x_0, r) := \frac{1}{r} \int_{-\infty}^{+\infty} f(x)\psi\left(\frac{x - x_0}{r}\right) dx;$$

2. Sum along the local maxima lines $\mathcal{L}(r)$ at scale r:

$$Z_q(r) = \sum_{l \in \mathcal{L}(r)} \left(\sup_{(x, \tilde{r}) \in l} \left| T_\psi[f](x, \tilde{r}) \right| \right)^q;$$

3. Find the scaling relation $Z_q(r) \sim r^{\tau(q)}$.

The set of maxima lines $\mathcal{L}(r)$ is defined as follows. Consider the set of extrema $L(r)$ defined by

$$L(r) := \left\{ x, \ \frac{\partial}{\partial x} \left(x \mapsto \left| T_\psi[f](x, r) \right| \right) = 0 \right\}. \tag{29}$$

Then, the set $\{(x, r), \ x \in L(r)\}$ is formed of connected curves called *maxima lines*. The set $\mathcal{L}(r)$ is then obtained as the set of all maxima lines defined for all $r' \leq r$. Explicitly,

$$\mathcal{L}(r) := \left\{ \left(x(r'), r'\right), \ \forall 0 \leq r' \leq r, x(r') \in L(r') \right\}. \tag{30}$$

Analyzing wavelets can be obtained from several ways. A classical one is to use the successive derivatives of the Gaussian function $\exp(-x^2/2)$. Indeed, the derivative of order n is orthogonal to polynomials of order up to n and of zero mean if n is greater than 1. See Fig. 8 for a representation of these wavelets for order 0–5.

Another possible way is to process convolutions of the unit box over Dirac type distributions. On Fig. 9, three successive convolutions of three variants of Dirac distributions are represented. The plot *Dij* is obtained from the Dirac distribution denoted *Dirac i* by applying *j* number of convolution. Note that only the last two Dirac distributions produce zero mean wavelets and that the unit box has been centered on 0 for aesthetic preferences.

WTMM can be easily extended to n dimensions by considering the wavelets formed by the partial derivatives $\psi = (\psi_1, \ldots, \psi_n)$ of a function ϕ such as the Gaussian function $\exp(-|X|^2/2)$, where $X = (x_1, \ldots, x_n)$. The wavelet transform is then replaced by the higher dimensional version

$$T_\psi[f](X_0, r) := \nabla T_\phi[f](X_0, r). \tag{31}$$

More details on the two-dimensional case and examples are provided by Arneodo et al. [43].

Despite its apparent complexity, the WTMM method offers a unique interface allowing to directly use functions representing continuous phenomena as data instead of fully defined mathematical measures. Heuristically, the wavelet transform does transform functions into measures (see the appendix). However, the ability to work directly with functions is more suitable for nonconservative phenomena, or if one wants to integrate the multifractal analysis in the equations the studied functions are a solution of. The simplicity of box-counting techniques and the adaptability offered by their many variants should nonetheless make those preferable for "static" measures.

In Fig. 10, MDFA is applied to the theoretical binomial cascade and to the random realization of it already used in Section 3.2. When the values of s are chosen as powers of 2 in the theoretical case, the N_s intervals are only translated copies of themselves, resulting in a completely flat spectrum (on the top). MDFA is particularly well suited for data such as the random realization (on the bottom) and gives the closest results to what is expected from the mathematical study (see Fig. 3), with only a slight offset to the left of the range of α which is explained by the fact that the iteration process generating the cascade was stopped at a relatively low level of iterations.

Since WTMM is meant to act on a different framework compared to the three previous methods, no application to binomial cascades are given. For illustration of the WTMM, the reader is referred to the work of Arneodo et al. [43] for numerous examples, and the work of Kantelhardt [41] and Audit et al. [54] for comparison with the MDFA method.
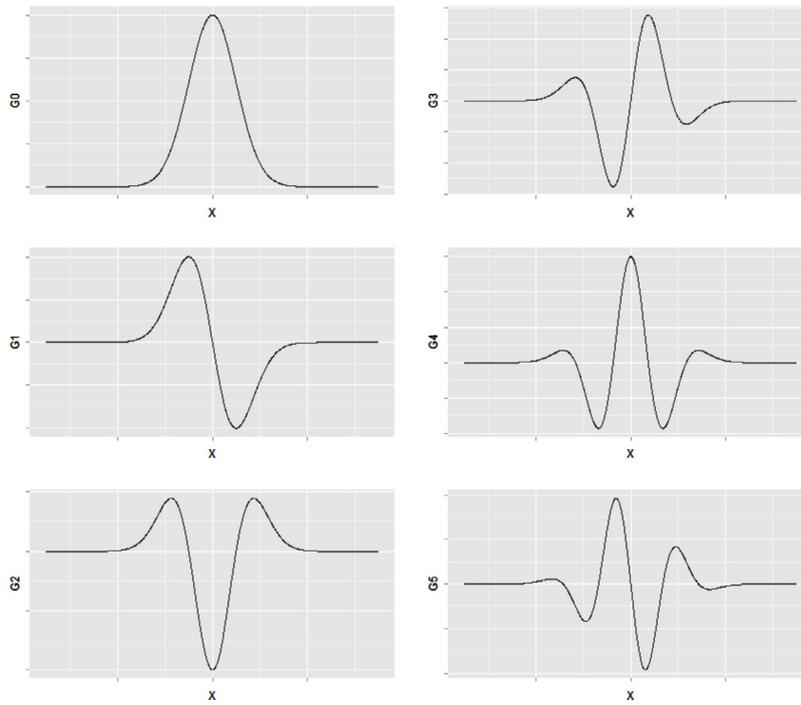
Fig. 8. Analyzing wavelets obtained from derivatives of the Gaussian function $\exp(-x^2/2)$.
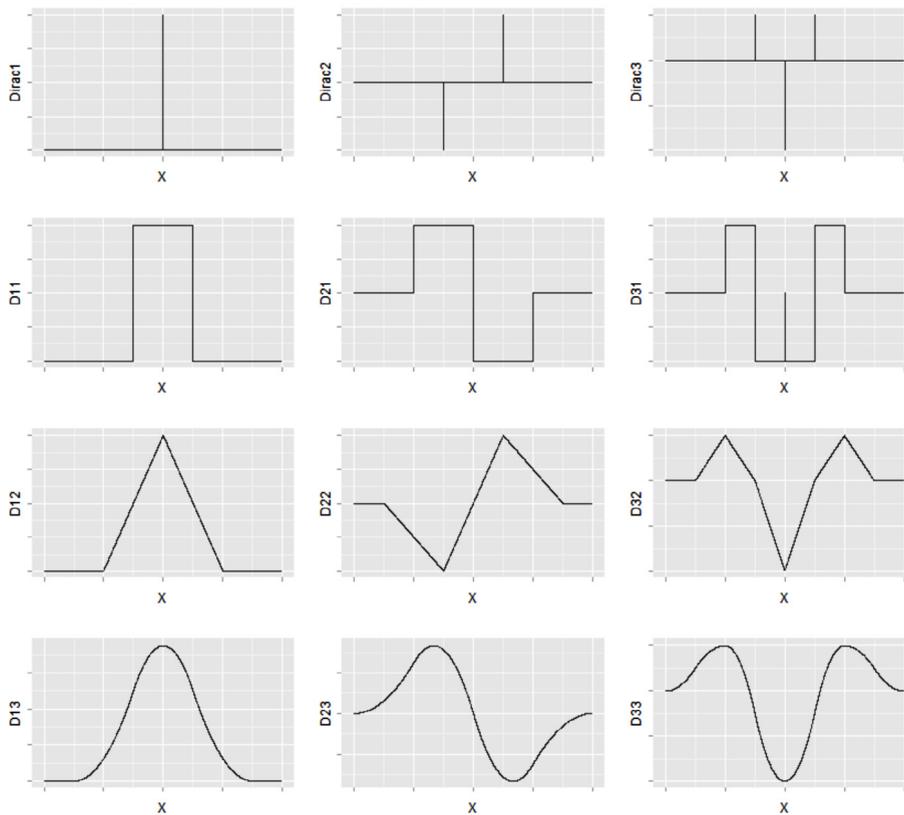

Fig. 9. Analyzing wavelets obtained from convolutions of the unit box over Dirac type distributions.
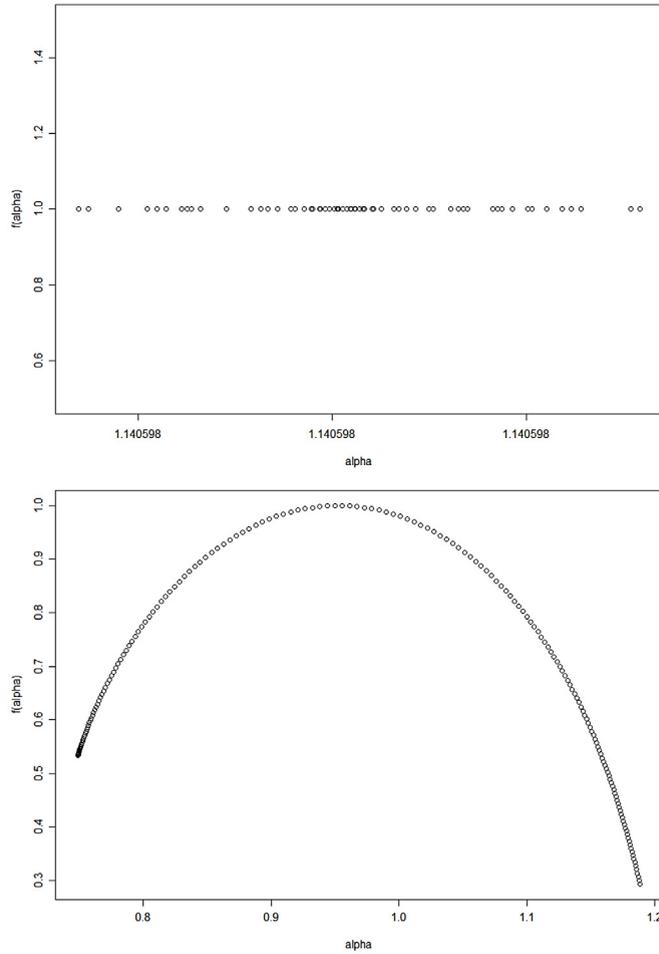
Fig. 10. MDFA applied to binomial cascades. On the top, a theoretical binomial cascade of parameter $p = 0.6$ and on the bottom a particular realization of it.

4. Interpretation and limits

The main elements of interpretation are presented in 4.1 and links are established with some classical measures of heterogeneity such as Shannon's entropy. The type of information one can expect to gain from multifractal analysis is illustrated with the binomial cascade and results from a study on the multifractality of London's street network [28]. In 4.2, limits of multifractal analysis are discussed, in particular the lack of consistency between methods.

4.1. Link to usual measures and interpretation

Recall Eq. (12) and introduce the approximation that only the exponent $\tau(q) = \alpha(q)q - f(\alpha(q))$ is making a significant contribution to the value of $Z(q)$.

$$Z(q) \approx r^{\tau(q)} \int_\alpha \rho(\alpha)d\alpha \tag{32}$$

$$= r^{\tau(q)}. \tag{33}$$

Then, one defines the *generalized dimension* as the family $\{D_q\}_q$, where

$$\forall q \neq 1, \; D_q := \lim_{r \to 0} \left[\frac{1}{q-1} \frac{\log Z(q)}{\log r} \right] = \frac{\tau(q)}{q-1}, \tag{34}$$

$$D_1 := \lim_{r \to 0} \frac{-\sum_i p_i \log p_i}{-\log(r)}. \tag{35}$$

Three values of the generalized dimension are of particular interest. D_0 is the usual box-counting dimension of D and therefore gives information on how much the data fills its physical support. D_1, referred to as the *information dimension*,
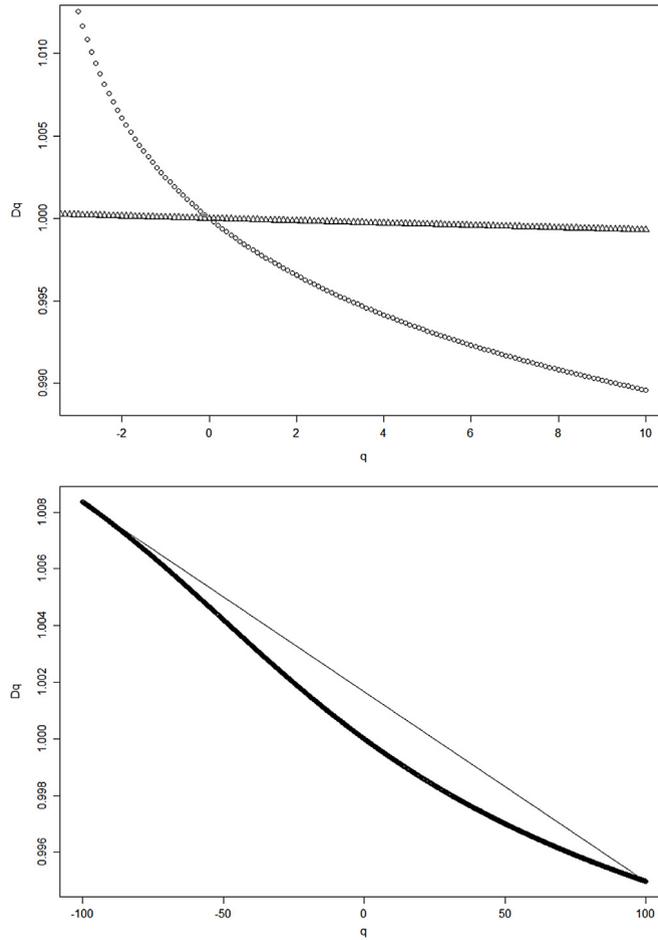
Fig. 11. Comparison of two sets of D_q. On the top, the standard moment method is applied to the theoretical random cascade of parameter $p = 0.6$ (triangles) and to a particular realization of it (circles), showing weak multifractality in the theoretical case and stronger multifractality in the particular case. On the bottom, the D_q curve of the theoretical case is zoomed in over a wider range of q and compared to a straight line to show its slight curvature.

relates to Shannon's entropy and captures how even the data density is, with higher values of D_1 meaning a more uniform density. D_2 is the probability of pairs of independent events occurring in the same box and measures how scattered the data is, with increasing compactness for increasing values of D_2. It is similar to the correlation dimension quoted in 2.1, (see Grassberger and Procaccia [44]).

The full plot of D_q versus q is representative of the strength of the multifractality of μ. The more constant the plot is, the weaker the multifractality is. See Fig. 11. It is also a good way to select the range of q to study. Indeed, Eq. (34) may lead to obtaining D_q values greater than d the dimension of D, in particular for negative values of q. As such a D_q would loose its physical meaning, one may want to restrict the range of q to ensure $D_q \leq d$.

Another expression to obtain D_q for $q \neq 1$ directly from $f(\alpha)$ and α is

$$D_q = \frac{f(\alpha) - q\alpha}{1 - q}. \tag{36}$$

It is mainly useful for the histogram method since $\tau(q)$ is never calculated when applying it.

The limit of the generalized dimension is that it only gives global measurements of the whole data. In contrast, the multifractal spectrum gives one dimension for each set where the data scales similarly. In a sense, the variable q selects different resolutions, with higher values of q selecting a local scaling $\alpha(q)$ of lower order. The variable $f(\alpha(q))$ then gives the local fractal dimension at resolution q.

It is easy to see that the spectrum's peak is achieved for $q = 0$, where $f(\alpha_0) = D_0$ is the fractal dimension of the physical support. Of particular relevance are therefore the asymmetries between the left and the right part of the spectrum. The spread of α indicates the variety of scaling present in the sample while the value of $f(\alpha)$ indicates the strength of the contribution of each α.

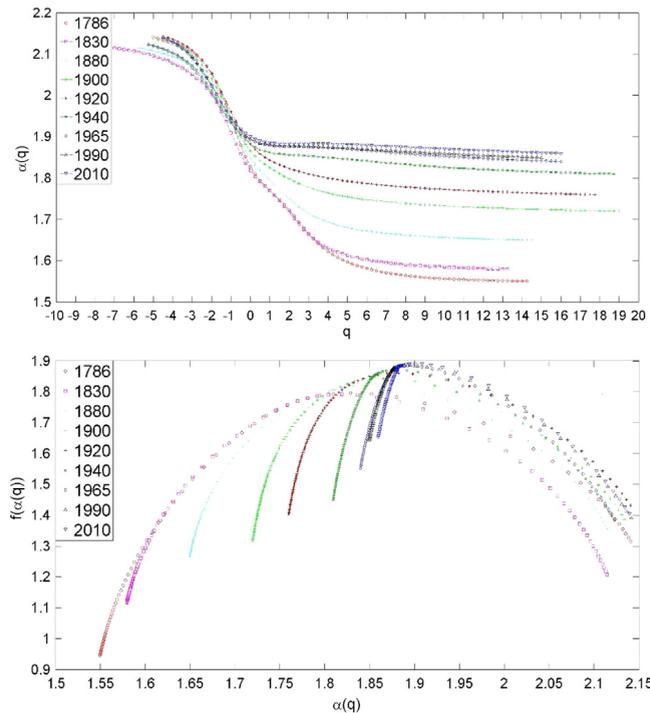**Fig. 12.** London's street network multifractality. On the top, the curve α versus q shows the variety of zones with low probabilities of intersection not evolving much over the years (left part of the curve), while disappearing in zones corresponding to higher probabilities (right part of the curve). On the bottom, the spectrum indicates an evolution that favors more and more areas with an increasingly lower density of intersections while presenting less and less variety.
Source: Murcio et al. [28].

For the binomial cascade, the range of $\alpha(q)$ is symmetric relatively to $q = 0$ and the spectrum is quite "round", denoting a well balanced repartition of each scaling and its contribution. A more interesting example is given by the evolution of London's street network.

Fig. 12 presents a clear picture of the structural differences that the London's street network has experimented in the last 200 years. The strong differences in value and shape between the first and last years is an evidence of how this street network evolved to lose its multifractal nature, and has become more homogeneous in terms of intersection density across the whole city.

Firstly, the singularity exponent $\alpha(q)$ accounts for the balance between areas with more/less street intersections (Fig. 12, top). In 1786, the $\alpha(q)$ values for positive q are relatively low compared with the $\alpha(q)$ values for negative q, a situation that confirms that the number of areas with major intersection densities are not that common. As we move forward in time, these differences are less and less evident, until 2010, where basically the same density can be found across the whole network.

The multifractal spectrum curves (Fig. 12, bottom) represent the distribution of intersection densities between the different regions. The left part of each curve (related with positive q, i.e., with the denser areas) became less and less wide as we move forward in time, while the right section remains stable for all nine networks. This is a clear indicator of how most of the network is evolving to become more similar through its different areas.

One last property of the curve $f(\alpha)$ versus α can be used to test the correctness of the result: from chapter 4 (pp. 46–53) of Nakayama and Yakubo [55], it can be deduced that $f(\alpha(1)) = \alpha(1)$, i.e. that $\tau(1) = 0$. Moreover $df(\alpha)/d\alpha = q = 1$, implying that the spectrum lies below the diagonal and touches it exactly in the point corresponding to $q = 1$. See Fig. 13 for the middle third Cantor set.

4.2. Limits

Using a multifractal study written by someone else can be a perilous initiative. The definition of what is called here $\mathcal{M}$ is often left implicit and some authors restrict the analysis to either the multifractal spectra or the generalized dimension, or even to the three values D_0, D_1 and D_2 only. Furthermore, most methods are equivalent for rigorously abstract self-similar sets and measures, but practical estimations often create an overshooting that is hard to measure. It is unfortunately a necessity to tailor the experimental protocol to the type and amount of data studied. One should therefore be extremely careful when comparing analyses made under different circumstances.
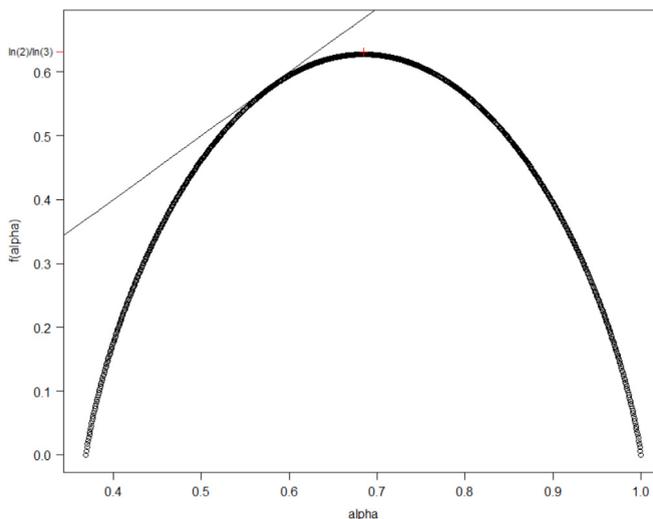
Fig. 13. Tangency with the diagonal. On the spectrum obtained for the middle third Cantor set, the diagonal $f(\alpha(q)) = \alpha(q)$ is tangent to the spectrum on the point where $q = 1$.

Results of isolated multifractal studies should be seen as giving "trends" or describing an evolution. In the work of Murcio et al. [28] for example, the loss of multifractality in London's street pattern through time is striking. The study provides rather precise information on which part of the spectrum is collapsing and from what point in time it becomes noticeable, only because it is consistently applied over sufficiently similar datasets of London's street network.

Another main issue when computing the spectra is controlling the error. To the theoretical error inherent to the chosen method one has to add the practical concessions for computational efficiency, and the measurement error that inevitably arises from the large amount of data required. The most common sources of error are the edge effects and the linear fit that is generally used to deduce $\tau(q)$ from the slopes of $Z(q)$ against r in log–log plots. The value $\tau(q)$ is quite sensitive and its computation often has to be automated and operated over only few values of r.

Choosing an adequate range of q is critical for moment based methods (including MDFA). In theory, q should go from $-\infty$ to $+\infty$, which is not possible in practice. For negative values of q, one may find values of D_q greater than the dimension of the support of the measure, indicating that those values of q should be discarded. Furthermore, values for which the linear fitting of $Z(q)$ versus r is not obtained with a sufficiently good level of confidence should also be discarded. For example, Lee and Stanley [56] found that for diffusion-limited aggregation, a phase transition occurs. Below some critical value of q, the function $Z(q)$ does not scale as a power law and forcing a linear fit for these values would create important discrepancies.

As evidenced through the example of the binomial cascades, the smallest resolution allowed by the studied dataset may play an important role in the results' accuracy. Multifractality aims at studying measures at an infinitely small scale and a scale as small as 10^{-6} might not be small enough to be considered "infinitely" small. In particular, when comparing two similar objects for which the measure cannot be evaluated with the same level of precision, one should use the largest minimal resolution of the two as the starting one for both.

It should be noted that predicting beforehand the shape of the spectrum is often difficult. There exist exactly self-similar nonrandom measures for which half of the spectrum is not defined [57], proving that unusual shapes may be normal. Another remark is that if μ_1 and μ_2 are finite measures on $\mathbb{R}^n$ with disjoint supports, then it can be proven that

$$f_{\mathcal{H}}^{\mu_1+\mu_2}(\alpha) = \max\left\{f_{\mathcal{H}}^{\mu_1}(\alpha), f_{\mathcal{H}}^{\mu_2}(\alpha)\right\}, \qquad (37)$$

where $f_{\mathcal{H}}^{\mu}$ is the fine spectrum of measure μ as defined in 3.1. In particular, $f_{\mathcal{H}}(\cdot)$ does not need to be concave.

5. Conclusions

Four methods were presented in the multifractal context. They are the moment method, the histogram method, MDFA and WTMM. Each is particularly well suited for a particular type of data: spatial data for the moment method and also for the histogram method provided the range of scaling is large enough, one-dimensional time series for MDFA and WTMM, with a convenient possibility to extend to higher dimensions for the last one.

Through the study of binomial cascades, it was found that the MDFA method gave the best results in addition to the possibility to detect the source of multifractality by applying the shuffling procedure. The moment method proved both simple and reliable with many variants giving it great adaptability to different contexts.

It was finally evidenced that multifractality can provide useful information about both the local and the global complexity and inhomogeneity of a phenomenon. In addition, it can be an effective comparative tool for measures and spaces that are

too complex for classical geometry provided the methodology used is consistent. It was emphasized that one should focus on trends and evolutive aspects in a phenomenon rather than expect to obtain a foolproof numerical result.

Acknowledgment

Appendix A. From diverse datasets to measures

Let X be a set and Σ a σ-algebra over X. A measure is any function $\mu : \Sigma \to \bar{\mathbb{R}}$ such that

1. $\forall A \in \Sigma, \mu(A) \geq 0$;
2. $\mu(\emptyset) = 0$;
3. for all countable or finite collection $\{A_i\}$ of pairwise disjoint sets in $\Sigma, \mu\left(\bigcup A_i\right) = \sum_i \mu(A_i)$.

Consider the Borel set $\mathcal{B}^n$ over $\mathbb{R}^n$ and a nonnegative integrable function $f : \mathbb{R}^n \to \mathbb{R}$. Define $\mu : \mathcal{B}^n \to \bar{\mathbb{R}}$, such that $\forall a_1, \ldots, a_n, b_1, \ldots, b_n \in \mathbb{R}$,

$$\mu(]a_1, b_1] \times \cdots \times]a_n, b_n]) := \int_{a_1}^{b_1} \cdots \int_{a_n}^{b_n} f(x)dx.$$

Then μ is a measure.

More generally, any increasing right continuous function F (uniquely) defines a measure by considering $\mu(]a, b]) = F(b) - F(a)$, it is not limited to the primitives of a nonnegative function.

Applying the wavelet transform to a function, explicitly, defining T_ψ such that

$$T_\psi[f](x_0, a) = \frac{1}{a} \int_{-\infty}^{+\infty} f(x)\psi\left(\frac{x - x_0}{a}\right) dx,$$

is heuristically equivalent to evaluating $\mu(W)$ for the above measure, where W is an element of the (heuristic) Σ-algebra created by the wavelets. In particular, if one considers the unit box as the wavelet and normalizes correctly, then T_ψ is formally equivalent to the above measure. Note that the integrability of f in the definition of μ is also required to apply the wavelet transform, so no information is lost by this hypothesis.

MDFA accomplishes the same transformation into a measure by dividing the original segment into smaller segments and summing on each of them, which is the discrete equivalent of μ defined above.

Finally, the methodology can be adapted to many other datasets even when those do not respect directly the σ-additivity hypothesis (third one in the definition of measure above). For example, average values should not be taken directly, as those are not additive, but can be considered when normalized by a fixed unit (unit of land, fixed number of particles or agents for example). Fractal landscape in particular become measures when the volume under the surface is considered. The only issue is to interpret accurately what aspects of the phenomenon are measured after the transformation and what aspects are not.

Appendix B. Supplementary data: R codes

Basic codes implemented in R for the first three methods described in this paper can be found online at http://dx.doi.org/10.1016/j.physa.2017.01.041.

References

[1] U. Frisch, G. Parisi, Turbulence and predictability of geophysical flows and climate dynamics, in: Proc. Varenna Summer School LXXXVIII.
[2] R. Benzi, G. Paladin, G. Parisi, A. Vulpiani, On the multifractal nature of fully developed turbulence and chaotic systems, J. Phys. A 17 (1984) 3521–3531. http://dx.doi.org/10.1088/0305-4470/36/10/312.
[3] A.B. Chhabra, C. Meneveau, R.V. Jensen, K. Sreenivasan, Direct determination of the $f(\alpha)$ singularity spectrum and its application to fully developed turbulence, Phys. Rev. A 40 (1989) 5284–5294. http://dx.doi.org/10.1103/PhysRevA.40.5284.
[4] C. Meneveau, K. Sreenivasan, Measurement of $f(\alpha)$ from scaling of histograms, and application to dynamical systems and fully developed turbulence, Phys. Lett. A 137 (1989) 103–112. http://dx.doi.org/10.1016/0375-9601(89)90093-5.
[5] L. de Arcangelis, S. Redner, A. Coniglio, Anomalous voltage distribution of random resistor networks and a new model for the backbone at the percolation threshold, Phys. Rev. B 31 (1985) 4725–4727. http://dx.doi.org/10.1103/PhysRevB.31.4725.
[6] L. de Arcangelis, S. Redner, A. Coniglio, Multiscaling approach in random resistor and random superconducting networks, Phys. Rev. B 34 (1986) 4656–4674. http://dx.doi.org/10.1103/PhysRevB.34.4656.
[7] C. Amitrano, A. Conoglio, F. di Liberto, Growth probability distribution in kinetic aggregation processes, Phys. Rev. Lett. 57 (1986) 1016–1019. http://dx.doi.org/10.1103/PhysRevLett.57.1016.
[8] P. Meakin, A. Coniglio, H. Stanley, T. Witten, Scaling properties for the surfaces of fractal and nonfractal objects: An infinite hierarchy of critical exponents, Phys. Rev. B 34 (1986) 3325–3340. http://dx.doi.org/10.1103/PhysRevA.34.3325.

[9] C. Peng, S. Buldyrev, M. Simons, H. Stanley, A. Goldberger, Mosaic organization of DNA nucleotides, Phys. Rev. E 49 (1994) 1685–1689. http://dx.doi.org/10.1103/PhysRevE.49.1685.

[10] S. Ossadnik, S. Buldyrev, A. Goldberger, S. Havlin, R. Mantegna, C. Peng, M. Simons, H. Stanley, Correlation approach to indentify coding regions in DNA sequences, Biophys. J. 67 (1994) 64–70. http://dx.doi.org/10.1016/S0006-3495(94)80455-2.

[11] Q. Cheng, The gliding box method for multifractal modeling, Comput. Geosci. 25 (1999) 1073–1079. http://dx.doi.org/10.1016/S0098-3004(99)00068-0.

[12] Q. Cheng, Multifractality and spatial statistics, Comput. Geosci. 25 (1999) 949–961. http://dx.doi.org/10.1016/S0098-3004(99)00060-6.

[13] G.-F. Gu, W.-X. Zhou, Detrended fluctuation analysis for fractals and multifractals in higher dimensions, Phys. Rev. E 74 (2006) 061104. http://dx.doi.org/10.1103/PhysRevE.74.061104.

[14] J. Muzy, J. Delour, E. Bacry, Modelling fluctuations of financial time series: from cascade process to stochastic volatility model, Eur. Phys. J. B 17 (2000) 537–548. http://dx.doi.org/10.1007/s100510070131.

[15] E. Bacry, J. Delour, J. Muzy, Modelling financial time series using multifractal random walks, Physica A 299 (2001) 84–92. http://dx.doi.org/10.1016/S0378-4371(01)00284-9.

[16] R. Morales, T.D. Matteo, R. Gramatica, T. Aste, Dynamical generalized Hurst exponent as a tool to monitor unstable periods in financial time series, Physica A 391 (2012) 3180–3189. http://dx.doi.org/10.1016/j.physa.2012.01.004.

[17] J. Barunik, T. Aste, T. D. Matteo, R. Liu, Understanding the source of multifractality in financial markets, Physica A 391 (2012) 4234–4251. http://dx.doi.org/10.1016/j.physa.2012.03.037.

[18] R. Morales, T.D. Matteo, T. Aste, Non-stationary multifractality in stock returns, Physica A 392 (2013) 6470–6483. http://dx.doi.org/10.1016/j.physa.2013.08.037.

[19] R. Buonocore, T. Aste, T.D. Matteo, Measuring multiscaling in financial time-series, Chaos Solitons Fractals 000 (2015) 1–10. http://dx.doi.org/10.1016/j.chaos.2015.11.022.

[20] Y. Chen, Y. Zhou, Multi-fractal measures of city-size distributions based on the three-parameter Zipf model, Chaos Solitons Fractals 22 (2004) 793–805. http://dx.doi.org/10.1016/j.chaos.2004.02.059.

[21] Y. Chen, J. Wang, Multifractal characterization of urban form and growth: the case of Beijing, Environ. Plann. B 40 (2013) 884–904. http://dx.doi.org/10.1068/b36155.

[22] Y. Chen, Multifractals of central place systems: models, dimension spectrums, and empirical analysis, Physica A 402 (2014) 266–282. http://dx.doi.org/10.1016/j.physa.2014.01.061.

[23] A. Facchini, S. Wimberger, A. Tomadin, Multifractal fluctuations in the survival probability of an open quantum system, Physica A 376 (2007) 266–274. http://dx.doi.org/10.1016/j.physa.2006.10.012.

[24] D. Wang, B. Fu, K. Lu, L. Xiao, Y. Zhang, X. Feng, Multifractal analysis of land use pattern in space and time: A case study in the Loess Plateau of China, Ecol. Complex. 7 (2010) 487–493. http://dx.doi.org/10.1016/j.ecocom.2009.12.004.

[25] S. Hu, Q. Cheng, L. Wang, S. Xie, Multifractal characterization of urban residential land price in space and time, Appl. Geogr. 34 (2012) 161–170. http://dx.doi.org/10.1016/j.apgeog.2011.10.016.

[26] S. Hu, Q. Cheng, L. Wang, D. Xu, Modeling land price distribution using multifractal IDW interpolation and fractal filtering method, Landsc. Urban Plan. 110 (2013) 25–35. http://dx.doi.org/10.1016/j.landurbplan.2012.09.008.

[27] A.B. Ariza-Villaverde, F. J. Jimenez-Hornero, E.G.D. Rave, Multifractal analysis of axial maps applied to the study of urban morphology, Comput. Environ. Urban 38 (2013) 1–10. http://dx.doi.org/10.1016/j.compenvurbsys.2012.11.001.

[28] R. Murcio, A.P. Masucci, E. Arcaute, M. Batty, Multifractal to monofractal evolution of the London's street network, Phys. Rev. E 92 (2015) 062130 . http://dx.doi.org/10.1103/PhysRevE.92.062130.

[29] A.K. Maity, R. Pratihar, A. Mitra, S. Dey, V. Agrawal, S. Sanyal, A. Banerjee, R. Sengupta, D. Ghosh, Multifractal detrended fluctuation analysis of alpha and theta EEG rhythms with musical stimuli, Chaos Solitons Fractals 81 (2015) 52–67. http://dx.doi.org/10.1016/j.chaos.2015.08.016.

[30] B.B. Mandelbrot, The Fractal Geometry of Nature, Freeman, San Francisco, 1982.

[31] C.J. Evertsz, B.B. Mandelbrot, Multifractal measures, in: H. Peitgen, H. Jürgens, D. Saupe (Eds.), Chaos and Fractals, Springer, New York, 1992, pp. 849–881.

[32] G. Brown, G. Michon, J. Peyriere, On the multifractal analysis of measures, J. Stat. Phys. 66 (1992) 775–790. http://dx.doi.org/10.1007/BF01055700.

[33] L. Olsen, A multifractal formalism, Adv. Math. 116 (1995) 82–196. http://dx.doi.org/10.1006/aima.1995.1066.

[34] R. Riedi, An improved multifractal formalism and self-similar measures, J. Math. Anal. Appl. 189 (1995) 462–490. http://dx.doi.org/10.1006/jmaa.1995.1030.

[35] Y. Pesin, Dimension Theory in Dynamical Systems, University of Chicago Press, Chicago, 1997.

[36] K. Falconer, Fractal Geometry, second ed., Wiley, Chichester, West Sussex, England, 2003.

[37] T.C. Halsey, M.H. Jensen, L.P. Kadanoff, I. Procaccia, B.I. Shraiman, Fractal measures and their singularities: The characterization of strange sets, Phys. Rev. A 33 (1986) 1141–1151. http://dx.doi.org/10.1103/PhysRevA.33.1141.

[38] H. Atmanspacher, H. Scheingraber, G. Wiedenmann, Determination of $f(\alpha)$ for a limited random point set, Phys. Rev. A 40 (1989) 3954–3963. http://dx.doi.org/10.1103/PhysRevA.40.3954.

[39] A. Arneodo, G. Grasseau, E. Kostelich, Fractal dimensions and $f(\alpha)$ spectrum of the Henon attractor, Phys. Lett. A 124 (1987) 426–432. http://dx.doi.org/10.1016/0375-9601(87)90546-9.

[40] J.W. Kantelhardt, S.A. Zschiegner, E. Koscielny-Bunde, A. Bunde, S. Havlin, H.E. Stanley, Multifractal detrended fluctuation analysis of nonstationary time series, Physica A 316 (2002) 87–114. http://dx.doi.org/10.1016/S0378-4371(02)01383-3.

[41] J.W. Kantelhardt, Fractal and multifractal time series, in: R. Meyers (Ed.), Encyclopedia of Complexity and System Science, Springer, 2002, pp. 3754–3779.

[42] J. Muzy, E. Bacry, A. Arneodo, Wavelets and multifractal formalism for singular signals: Application to turbulence data, Phys. Rev. Lett. 67 (1991) 3515–3518. http://dx.doi.org/10.1103/PhysRevLett.67.3515.

[43] A. Arneodo, B. Audit, N. Decoster, J.-F. Muzy, C. Vaillant, Wavelet based multifractal formalism: applications to dna sequences, satellite images of the cloud structure and stock market data, in: A. Bunde, J. Kropp, H.J. Schellnhuber (Eds.), The Science of Disasters: Climate Disruptions, Heart Attacks, and Market Crashes, Springer Verlag, Berlin, 2002, pp. 26–102.

[44] P. Grassberger, I. Procaccia, Measuring the strangeness of strange attractors, Physica D 9 (1983) 189–208. http://dx.doi.org/10.1016/0167-2789(83)90298-1.

[45] C. Tricot, Two definitions of fractional dimension, Math. Proc. Cambridge 91 (1982) 57–74. http://dx.doi.org/10.1017/S0305004100059119.

[46] E.A. Ihlen, B. Vereijken, Multifractal formalisms of human behavior, Hum. Mov. Sci. 32 (2013) 633–651. http://dx.doi.org/10.1016/j.humov.2013.01.008.

[47] H. Stanley, P. Meakin, Multifractal phenomena in physics and chemistry, Nature 335 (1988) 405–409. http://dx.doi.org/10.1038/335405a0.

[48] A.B. Chhabra, K. Sreenivasan, Negative dimensions: Theory, computation, and experiment, Phys. Rev. A 43 (1991) 1114–1117. http://dx.doi.org/10.1103/PhysRevA.43.1114.

[49] Q. Cheng, F. Agterberg, Comparison between two types of multifractal modeling, Comput. Geosci. 28 (1996) 1001–1015. http://dx.doi.org/10.1007/BF02068586.

[50] A.B. Chhabra, R.V. Jensen, Direct determination of the $f(\alpha)$ singularity spectrum, Phys. Rev. Lett. 62 (1989) 1327–1330. http://dx.doi.org/10.1103/PhysRevLett.62.1327.
[51] P. Kestener, A. Arneodo, Three-dimensional wavelet-based multifractal method: the need for revisiting the multifractal description of turbulence dissipation data, Phys. Rev. Lett. 91 (2003) 194501 . http://dx.doi.org/10.1103/PhysRevLett.91.194501.
[52] T. Tel, Fractals, multifractals, and thermodynamics, Z. Naturforsch 43a (1988) 1154–1174. http://dx.doi.org/10.1515/zna-1988-1221.
[53] A. Arneodo, E. Bacry, J. Muzy, The thermodynamics of fractals revisited with wavelets, Physica A 213 (1995) 232–275. http://dx.doi.org/10.1016/0378-4371(94)00163-N.
[54] B. Audit, E. Bacry, J.-F. Muzy, A. Arneodo, Wavelet-based estimators of scaling behavior, IEEE Trans. Inf. Theory 48 (2002) 2938–2954. http://dx.doi.org/10.1109/TIT.2002.802631.
[55] T. Nakayama, K. Yakubo, Fractal Concepts in Condensed Matter Physics, Springer-Verlag, Berlin, Heidelberg, New-York, 2003.
[56] J. Lee, H.E. Stanley, Phase transition in the multifractal spectrum od diffusion-limited aggregation, Phys. Rev. Lett. 61 (1988) 2945–2948. http://dx.doi.org/10.1103/PhysRevLett.61.2945.
[57] B.B. Mandelbrot, C.J. Evertsz, Y. Hayakawa, Exactly self-similar left-sided multifractal measures, Phys. Rev. A 42 (1990) 4528–4536. http://dx.doi.org/10.1103/PhysRevA.42.4528.