Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Computational Physics 216 (2006) 362–390
JOURNAL OF COMPUTATIONAL PHYSICS
www.elsevier.com/locate/jcp

Numerical methods for the estimation of multifractal singularity spectra on sampled data: A comparative study

Antonio Turiel [a,*], Conrad J. Pérez-Vicente [b], Jacopo Grazzini [c]

[a] *Physical Oceanography Group, Institut de Ciencies del Mar – CMIMA (CSIC), Passeig Maritim de la Barceloneta, 37-49, 08003 Barcelona, Spain*
[b] *Complex System Group, Departament de Fisica Fonamental, Universitat de Barcelona, Diagonal, 647, 08028 Barcelona, Spain*
[c] *Regional Analysis Division, Institute of Applied Computational Mathematics (FORTH), Vassilika Vouton, 71110 Heraklion, Greece*

Received 14 March 2005; received in revised form 7 December 2005; accepted 8 December 2005
Available online 23 January 2006

Abstract

Physical variables in scale invariant systems often show chaotic, turbulent-like behavior, commonly associated to the existence of an underlying fractal or multifractal structure. However, the assessment of multifractality over experimental, discretized data requires of appropriate methods and to establish criteria to measure the confidence degree on the estimates. In this paper we have evaluated the quality of different techniques used for multifractal analysis. We have tested four different techniques: the moment (M) method, the wavelet transform modulus maxima (WTMM) method, the gradient modulus wavelet projection (GMWP) method and the gradient histogram (GH) method, which are used to estimate the singularity spectra of multifractal signals. The test consists in analyzing synthetic multifractal 1D signals with given multifractal spectrum. We have compared the results, studying the sensibility of each method to the length of the series, size of the ensemble and type of spectrum. Our results show that GMWP method is the one attaining the best performance, providing reliable estimates which can be improved when the statistics is increased. All the other methods are affected by problems such as the linearization of the right tail of the spectrum, and some of them are very demanding in data.

Keywords: Fractals; Multifractals; Wavelets; Singularity analysis; Numerical methods; WTMM

1. Introduction

The multifractal formalism provides a scale invariant mechanism for the analysis and generation of complex signals which fits well with the observed experimental properties of fully developed turbulence (FDT) [1] as well as other physical systems which range from natural images [2] to heartbeat dynamics [3] or econometric signals [4] to cite some examples. The multifractal formalism also allows to implement new coding strategies [5,6] and to highlight relevant dynamical features of the systems under study [7]. In addition, theoretical

* Corresponding author. Tel.: +34 932 309 602; fax: +34 932 309 555.
E-mail addresses: turiel@icm.csic.es (A. Turiel), conrad@ffn.ub.es (C.J. Pérez-Vicente), grazzini@iacm.forth.gr (J. Grazzini).

0021-9991/$ - see front matter
doi:10.1016/j.jcp.2005.12.004

models can be devised to fit the observed multifractal properties [8,4]. These facts explain why it is triggering an increasing interest in the scientific community.

The process of characterization of multifractality from experimental data must be performed with care. In essence, multifractality is a property verified in the infinitesimal limit only, while empirical data have an inherent discrete nature. Therefore, any tool designed to validate the multifractal character of a given signal faces several difficulties elicited by the finite size and the discretization of the data record. In addition, any technique used to validate the multifractal behavior of a signal necessarily involves some interpolation scheme which can make it prone to some bias. For that reason, it is very convenient to know beforehand the range of validity, limitations and biases as well as the theoretical foundations of any method (''validation method'', in the rest of the paper) devoted to carry out this task. This knowledge allows to determine the degree of reliability of the estimates and to define a framework where alternative methodologies can be compared.

In this paper, we will study the theoretical foundations, performance and reliability of four different validation techniques for the analysis of multifractality from experimental data. We will discuss the performance of each method according to their capability to retrieve the correct singularity spectrum from a given data set. As a test benchmark we will use synthetic data generated following a simple model originally devised for the description of turbulent data. With this setup we have a prior knowledge of the singularity spectra of the data and, as a consequence, the possibility to calibrate the quality of the answers. We have worked with 1D series, but the generating model, methods and results can be generalized to higher dimensions. In this paper's webpage [9] the reader can find C source codes which generate synthetic 1D and 2D multifractals and which also implement the validation techniques.

The paper is structured as follows: Next section is devoted to introduce general aspects of multifractal theory, required for a proper understanding of the methods. In Section 3, we will present the four validation techniques. Section 4 is devoted to the definition and presentation of the three benchmarks of synthetic series as well as the different ensembles of signals that will be processed. In Section 5 we will use the first benchmark to illustrate the different particularities of each method, discussing their potential advantages and known drawbacks. Section 6 is devoted to provide quantitative measurements on the performance of the different validation techniques, using the two other benchmark of synthetic signals to study the dependence on the amount and type of statistics of the studied data. A general discussion, with the overall evaluation and comparison of the methods is given in Section 7. Finally, in Section 8 we present the conclusions of our work.

2. Multifractal theory

The origins of multifractal theory can be tracked back to the seminal works by Kolmogorov [10]. Under conditions of intense turbulence (fully developed turbulence), variables as the velocity or the local dissipation of energy vary sharply from one location to another and can be regarded not as deterministic quantities but as random ones. Let, for instance, $\epsilon_r(\vec{x})$ be the local dissipation of energy around the point $\vec{x}$ and over a neighborhood of radius r, whose expression is given by the following formula:

$$\epsilon_r(\vec{x}) = \frac{1}{|B_r(\vec{x})|}\int_{B_r(\vec{x})} \mathrm{d}\vec{x}' \sum_{ij} [\partial_i v_j(\vec{x}') + \partial_j v_i(\vec{x}')]^2, \tag{1}$$

where v_i are the components of the velocity vector and $B_r(\vec{x})$ stands for the ball of radius r centered around $\vec{x}$. Kolmogorov's intuition was that the energy is transmitted from the larger scales (L) to the smaller ones (r) by means of an injection process defined by a variable η_{rL} which in fact only depends on the ratio r/L, as

$$\epsilon_r = \eta_{rL}\epsilon_L. \tag{2}$$

In Kolmogorov's original work the energy injection variable η_{rL} has a fixed value, $\eta_{rL} = [\frac{r}{L}]^{-\alpha}$, from which it can immediately deduced that the order p moments of ϵ_r can be related with those of ϵ_L in a very simple way, namely

$$\langle \epsilon_r^p \rangle = \left[\frac{r}{L}\right]^{-\alpha p} \langle \epsilon_L^p \rangle \propto r^{-\alpha p},$$

where $\langle\cdot\rangle$ denotes an average over an ensemble of $\vec{v}$ realizations. All the dependence in r of the order-p moment of ϵ_r is concentrated in the power-law $r^{-\alpha p}$, which is similar to what experimental measures show, namely:

$$\langle \epsilon_r^p \rangle \propto r^{\tau_p}, \tag{3}$$

a property which is known as self-similarity (SS). Unfortunately, the exponents τ_p obtained in the experiments do not have a linear dependence on p. In general, τ_p as a function of p is a convex curve, a phenomenon which is known as "anomalous scaling" [11], in opposition to the "normal" (linear) scaling. To describe the "anomalous scaling" Kolmogorov's decomposition can still be applied[1] but now η_{rL} in Eq. (2) has to be interpreted as a random variable, independent of ϵ_L. Moreover, in order to produce a consistent scheme, the variables η_{rL} have to be infinitely divisible [12], to ensure that downwards process from scale L to scale r is verified directly or in several stages (forming the famous *cascade* process). The experimental property of self-similarity led researchers to propose a model for its generation based on the existence of local scale-invariant laws. First, it is assumed that at any point $\vec{x}$ the following equation holds:

$$\epsilon_r(\vec{x}) \propto r^{h(\vec{x})} \quad \text{as } r \searrow 0, \tag{4}$$

that is, the dependency on a vanishing positive scale parameter r is conveyed by the power-law factor $r^{h(\vec{x})}$. The exponent $h(\vec{x})$, which is a function of the point $\vec{x}$ under study, is called the *singularity exponent* of the point. Then, the singularity exponents can be arranged in special sets called *singularity components* F_h defined as

$$F_h = \{\vec{x} : h(\vec{x}) = h\}. \tag{5}$$

In order to close the model, it is required that the singularity components are of fractal character (and for that reason they are also known as *fractal manifolds*). The *singularity spectrum* (also known as the Hausdorff spectrum) associated to the multifractal hierarchy of fractal components is the function $D(h)$ defined by the fractal (Hausdorff) dimension of each component F_h, namely

$$D(h) = \dim_H F_h. \tag{6}$$

Following the famous derivation by Parisi and Frisch [13], under some assumptions on the homogeneity and isotropy of the statistics of local singularities, it is possible to derive a relation between self-similarity exponents τ_p and the singularity spectrum $D(h)$. They proved that the SS exponents τ_p can be computed from the Legendre transform of the singularity spectrum $D(h)$,

$$\tau_p = \inf_h\{hp + d - D(h)\}, \tag{7}$$

where d stands for the dimension of the embedding space. By means of Eq. (7) it is possible to relate statistics and geometry. It is then evident that the singularity spectrum contains all the information about SS, hence about the multiplicative process, Eq. (2), what is the same as saying that it describes the statistics of changes in scale. How much information is actually conveyed by the singularity spectrum? In fact, there are features of a signal which are not described by the singularity spectrum. There is no known way to split a signal in a multifractal part, responsible of the singularity spectrum, and a mono-fractal or something similar what would explain the remaining features, although some attempts have been made in that direction [14,15]. Anyway, a precise knowledge of the singularity spectrum will provide valuable information on the inner structure of a system.

The Hausdorff spectrum, i.e., the one given by Eq. (6), is in fact what we need, but it is difficult to estimate. The main inconveniences with such a definition is that is very difficult to explicitly assign a singularity exponent to each point (see discussion in Section 3.2), to decompose the signal into its fractal manifolds and, in addition, even when the multifractal decomposition of the signal is available there are numerical problems to provide a good estimate of the fractal dimensions. Box-counting and related measures [16] are usually ill-behaved, specially when the fractal components are (ideally) dense sets. And a direct application of the definition of Hausdorff dimension (which implies to search optimal coverings with sets of diameter *at most* ϵ, then let $\epsilon \to 0$) is impossible for values of ϵ below the discretization size.

[1] However notice that Eq. (2) applies in a statistical sense only, that is, for a given basis point $\vec{x}$ you cannot conclude the following: $\epsilon_r(\vec{x}) = \eta_{rL}(\vec{x})\epsilon_L(\vec{x})$. This last relation can only hold in an appropriate representation basis [6].

Thus, the singularity spectrum is often estimated using indirect methods which are numerically more stable. We will concentrate on the so-called *Legendre spectrum*. One of the advantages of Parisi–Frisch's formula, Eq. (7), is that it can be inverted. By definition, the Legendre spectrum corresponds to the Legendre transform of the SS exponents τ_p

$$D_l(h) = \inf_p\{hp + d - \tau_p\},$$

where, as before, d stands for the dimension of the environment space. It is simple and accessible for numerical calculations. The main disadvantage is that, by construction, the Legendre spectrum is convex, so if the Hausdorff spectrum is not convex the Legendre spectrum will equal its convex hull.

In mathematical literature, it is common to consider that a system is multifractal only when the Hausdorff spectrum, the Legendre spectrum and another spectrum not introduced here, the *large deviation spectrum*, coincide. This point of view is a bit restrictive; in fact, from the physical point of view the only relevant spectrum is the one which links statistics and geometry: the Hausdorff spectrum. It should be noticed that the first two methods are based on the estimation of the Legendre spectrum while the last two methods analyzed give direct access to the Hausdorff spectrum, which makes them more reliable. Anyway, Hausdorff and Legendre spectra are coincident for the signals used in our benchmark due to the model used to generate them.

3. Presentation of the different methods

We have analyzed the performance of four methods on the estimation of the singularity spectrum of experimental data. In the experiences shown in this paper, and to simplify the analysis, we will only consider scalar signals defined in a 1D space, although all the derivations could be extended to scalar signals in higher dimensional spaces. For that reason, in order to keep the presentation as general as possible all the quantities will be described in terms of a vector spatial variable (position) $\vec{x}$. Thus, a generic multifractal signal will be denoted by $s(\vec{x})$. Let us first introduce the theoretical foundations of each method.

3.1. M method

Moment (M) method is primarily a statistical method. It is based on getting the Legendre spectrum from SS exponents associated [1,8] either to an appropriate scale-dependent variable like ϵ_r (in the simplest formulation) or to a partition function (in a more sophisticated version). In the study of turbulence, the variable ϵ_r is usually identified with the energy dissipation over a ball of radius r [12] or the wavelet projection at scale r of a velocity component [17]. The advantages and limitations of using wavelet projections are discussed in the following section. In this paper, we will use the first, simpler version. Then, our implementation is done according the following steps:

1. *Definition of the variable* $\epsilon_r(\vec{x})$*:* Following [2,5,7,18] and inspired by the local dissipation of energy in turbulent flows, Eq. (1), we define ϵ_r at each point $\vec{x}$ as

$$\epsilon_r(\vec{x}) = \frac{1}{|B_r|}\int_{B_r(\vec{x})} \mathrm{d}\vec{x}'|\nabla s|(\vec{x}'). \tag{8}$$

   The integral above is estimated by simple step rectangles. Such a variable has been shown to exhibit SS in many contexts [2,5].
2. *Calculus of the moments:* We average $\epsilon_r^p(\vec{x})$ over all the points $\vec{x}$ in the ensemble (all points are considered, as we assume that spatial correlations do not significantly alter the value of moments). For each value of r, we compute the moments for a particular, finite sequence of p's: $p_1, p_2, \ldots, p_n$. For each order p_i we evaluate the value of the associated τ_{p_i} as the slope of the linear regression of $\log\langle\epsilon_r^{p_i}\rangle$ vs. $\log r$.
3. *Computing the spectrum:* Interpolating τ_p by a second-order spline, we evaluate the slope of the curve τ_p vs. p at each p_i. The value, an estimate of the derivative of τ_p at p_i, corresponds to a value h_i in the spectrum: $h_i = \frac{\mathrm{d}\tau_p}{\mathrm{d}p}(p_i)$. The corresponding value of the Legendre spectrum is given by $D(h_i) = h_i p_i + d - \tau_{p_i}$.

3.2. Wavelet transform modulus maxima method

Inspired in early works by Mallat [19] on the use of wavelet projections to provide a multiresolution representation of signals, the concept of wavelet transform modulus maxima (WTMM) was originally introduced by Mallat and Zhong [20] in the context of signal processing and later on applied by Muzy et al. [17] in the analysis of turbulent flows, in order to solve some processing problems with experimental data. The main idea of WTMM method [17,21–23] is to obtain the multiresolution skeleton of the signal, a characteristic set of scale-space lines which, following Mallat's conjecture [24], would be enough to provide a complete description of the signal and which would be concentrated on the separation of structures according to their scale of influence. This multiresolution skeleton are the lines of WTMM, that is, the set of lines where the wavelet transform reaches a local maximum (with respect to the position coordinate). It was claimed in [17] that the branching structure of the skeleton defined by these lines in the scale-space plane highlights the hierarchical organization of the singularities.

To understand the basic elements of WTMM we need to remind what wavelet projections are. In the context of this paper, a wavelet Ψ is a function which can be used to analyze the local properties of any signal. Given a signal $s(\vec{x})$, the wavelet projection (or wavelet transform, WT) of this signal over the wavelet Ψ at the point $\vec{x}$ and the scale of observation r is denoted by $T_\Psi s(\vec{x}, r)$, and it is defined by

$$T_\Psi s(\vec{x}, r) \equiv \int \mathrm{d}\vec{x}' s(\vec{x}') \frac{1}{r^d} \Psi\left(\frac{\vec{x} - \vec{x}'}{r}\right) = s \otimes \Psi_r(\vec{x}), \tag{9}$$

where $\otimes$ stands for the convolution product and the function $\Psi_r(\vec{x})$ is given by $\Psi_r(\vec{x}) \equiv \frac{1}{r^d} \Psi(\frac{\vec{x}}{r})$. Therefore, the wavelet projection consists in convolving the signal with an appropriate "focus kernel" Ψ_r which can be tuned with the scale parameter r in order to zoom in and out the details surrounding each point $\vec{x}$ under analysis. A signal s would be considered multifractal if for any Ψ such that Ψ annihilates a certain number m of polynomial moments (in order to filter away long-range correlations which could mask the local multifractal structure [17,22]) it is verified that:

$$|T_\Psi s(\vec{x}, r)| \sim r^{h(\vec{x})} \quad \text{as } r \searrow 0, \tag{10}$$

where the obtained local singularity exponent $h(\vec{x})$ should be otherwise independent of the wavelet used for all $h(\vec{x}) < m$ (where m is the wavelet order) [22,25]. The WTMM method takes advantage of the scale-space partitioning given by the skeleton of maxima lines to define a partition function $Z(r,p)$ of order p and scale r as the sum over the set $\{\vec{x}_\alpha(r)\}_\alpha$ of coordinates supporting maxima of the modulus of WT; namely,

$$Z(r, p) \equiv \sum_\alpha |T_\Psi s(\vec{x}_\alpha(r), r)|^p. \tag{11}$$

Summing only over those well-chosen points (the maxima lines) rather than making an average over the whole sampling interval makes sense since most of the information is carried by the wavelet maxima lines [24]. Namely, the locations and values of the singularity exponents can be recovered from the scaling of WT along the maxima lines. If the signal is multifractal and the wavelet appropriate [17], $Z(r,p) \sim r^{\tau_p}$ as $r \searrow 0$ and so we can proceed to obtain the Legendre spectrum, similarly to M method. However, some technical difficulties must be solved. First, in order to get rid of some maxima with very low WT values (which may induce spurious divergences), a refinement of Eq. (11) summands must be introduced, leading to a scale adaptive version of the WTMM method

$$Z(r, p) \equiv \sum_\alpha \left(\sup_{\vec{x}, r' \leqslant r} |T_\Psi s(\vec{x}_\alpha(r'), r')| \right)^p, \tag{12}$$

where $\sup_{\vec{x}, r' \leqslant r}$ means that the supremum is taken for $(\vec{x}, r')$ on each line of maxima at scales $r' \leqslant r$. Secondly, to avoid other numerical difficulties associated with the approximation of the Legendre transform (and mainly due to the divergence of negative moments or some end points of the sampled singularities), it is often preferred to use the so-called canonical approach to compute the $D(h)$ [26]. This method consists in computing the average over a range of scales (r_0, r_1) of some appropriate quantities derived from the partition function

$Z(r,p)$. According to theory, these quantities must scale as $r_1^{h(p)}$ and as $r_1^{D(h(p))}$ and hence, just taking logarithms and dividing by the logarithm of r_1 the couple $(h(p), D(h(p)))$ is obtained. The leading terms in the averages are the ones defined by the smallest scale r_1 and they also scale as $r_1^{h(p)}$ and as $r_1^{D(h(p))}$, but to average over the full scaling line is convenient, as it allows to diminish the influence of errors and oscillations. For further information, see [27].

It is important to remark that, as we are processing the signal itself, and not its gradient (as in M method), the exponents τ_p are shifted by $+p$ with respect to those of M method. Conversely, you can say that the singularities h are shifted by $+1$ and thus the singularity spectrum appears horizontally displaced by this factor [5,28]. We have taken that shift into account when representing the results from WTMM, shifting WTMM spectra by -1 to obtain results directly comparable to those of other methods.

Although our code includes a home-made implementation of the WTMM method, we have implemented WTMM method using the algorithms provided by some built-in software very often referenced in the literature: a GPL stand-alone software called LastWave [29], and two add-on toolboxes for the computing environment Matlab: WaveLab [30] and FracLab [31]. In this way, we intended to obtain closer outcomes to those of standard literature. The scripts which perform all the calculations with the different softwares can also be found in [9]. For simplicity, in this paper we will show only the results provided by LastWave. The basic implementation of WTMM therein consists in the following basic steps:

1. *Compute the WT:* After choosing an appropriate analyzing wavelet Ψ, we compute the WT $T_\Psi s(\vec{x}, r)$ according to Eq. (9) for several translations and dilations. We will discuss about the different parameters of the projection (wavelet, minimal scale, etc.) below. The wavelet projections are obtained via FFT with aliasing corrections.
2. *Find the extremal representation associated to the WT:* We locate the local maxima of the absolute value of the WT as a function of position at each scale. The modulus maxima are points $(\vec{x_0}, r_0)$ such that $|T_\Psi s(\vec{x}, r_0)| < |T_\Psi s(\vec{x_0}, r_0)|$ when $\vec{x}$ belongs to either the right or left part of the neighborhood of $\vec{x_0}$.
3. *Connect (chain) the maxima lines:* We have to check whether a local maximum $(\vec{x_0}, r_0)$ at a given scale r_0 is located close to a maximum at a smaller scale $r < r_0$. A simple search on the nearest neighbors is performed, in this case. The maxima lines α are obtained as the set of connected curves in the scale space $(\vec{x}, r)$ along which all points are modulus maxima.
4. *Track the maxima across the scales:* Then, we track maxima lines for increasing scale r by choosing at each scale the supremum $\sup_{\vec{x}, r' \leqslant r}$ between all previous values at smaller scales $r' \leqslant r$. This way, we replace the WT value of an extremum by the maximum value along its line.
5. *Compute the partition function:* We sum the maximum values of the WT along the modified maxima lines to get the $Z(r,p)$ according to Eq. (12).
6. *Compute the scaling exponents:* SS exponents τ_p are obtained as the slopes of the log–log plots $\log Z(r,p) \approx \tau_p \log r + C(p)$ by linear regression over r.
7. *Estimate the spectrum:* Finally, the multifractal spectrum $(h, D(h))$ are computed from the τ_p through the canonical approach.

To end let us remark that when working with ensembles consisting in more than a single signal, the partition function of each signal is evaluated, then all the partition functions are summed up, because WTMM partition functions are additive [22] (contrarily to the standard concept of partition function in Statistical Mechanics).

3.3. Gradient modulus wavelet projection method

The method based on the analysis of gradient modulus wavelet projections (GMWP) is primarily a geometric method. GMWP's give explicit estimates of the values of $h(\vec{x})$ at every point $\vec{x}$. In that sense, GMWP method can be considered as a direct check of multifractality because it explicitly performs the multifractal decomposition of a signal (something that can also be obtained with some refinements of WTMM techniques, see [23]).

GMWP method recovers the idea underlying the definition of local energy dissipation which was already used in the moment method, and combines it with wavelet projections in order to get the assessment of the local singularity exponents. Namely, we will work on the wavelet projections of the "function" $|\nabla s|$ (may it be used in a distribution sense). A more precise formalism includes defining a multifractal measure [32]

$$\mathrm{d}\mu(\vec{x}) = \mathrm{d}\vec{x}|\nabla s|(\vec{x}). \tag{13}$$

The existence of a multifractal hierarchy of singularity exponents, Eq. (4), is translated to the measure μ by stating

$$\mu(B_r(\vec{x})) \propto r^{d+h(\vec{x})} \quad \text{as } r \searrow 0. \tag{14}$$

What is done in practice is to study the wavelet projections of the measure, $T_\Psi\mu(\vec{x},r) \equiv T_\Psi|\nabla s|(\vec{x},r)$, that is, to study GMWP's. The existence of a multifractal hierarchy then reads as

$$T_\Psi|\nabla s|(\vec{x},r) \propto r^{h(\vec{x})}. \tag{15}$$

Eq. (15) has been taken as the starting point to check multifractality of signals in systems as diverse as natural images [5], meteorological satellite images [18] or econometric series [7]. Obtaining explicit values of singularity exponents for each point has the appeal of revealing the geometrical arrangement of multifractality. In addition, it can be checked point by point to which extent multifractality holds, just looking at the coefficients of the log–log regression by which Eq. (15) is verified. The use of GMWP's has another advantage with respect to WTMM from the numerical point of view, which will be further discussed in Section 7: to compute GMWP's we do not need to use a wavelet annihilating a certain number of moments and even a positive function can be used. Positive functions are not *admissible wavelets*, that is, they cannot be used to represent signals [33]; but they can be used as *analyzing wavelets*, that is, they are able to describe local behaviors. Positive analyzing wavelets provide finer resolution, because the minimal resolution that a wavelet projection can attain is related to the number of zero-crossings of the wavelet [34].

Having access to the singularity exponents $h(\vec{x})$ at every point $\vec{x}$ we can directly estimate the singularity spectrum using a histogram method [16]. If the exponents were obtained at a minimum resolution r_0 then the distribution of singularities $\rho(h)$ verifies

$$\rho(h) \propto r_0^{d-D(h)}. \tag{16}$$

It should be remarked that Eq. (16) is not only the starting point for Parisi–Frisch's derivation which was used to define Legendre spectrum; it also allows to access to the Hausdorff spectrum, $D(h)$, although in a probabilistic way. Something interesting about this formula is that it allows to interpret negative dimensions as events with small probability [35].

Applying Eq. (15) we will obtain the singularity exponents at only one resolution r_0, that is, the minimum resolution in the regression. So that, we need some additional information in order to invert Eq. (16). One possibility is to repeat the analysis at other resolutions, then to perform a linear regression of the logarithm of $\rho(h)$, Eq. (16), vs. $\log r$. Although effective, this procedure can be numerically highly time consuming, as it implies analyzing several times the whole series. But there is a simpler alternative when we can assume that the multifractal has total support, which is valid except for monofractals (monofractals should then be treated by the general scale regression method). In that case, there exists a fractal component F_{h_1} of maximal fractal dimension, $D(h_1) = d$; according to Eq. (16) this fractal component is also that of maximal probability. Assuming that there exists such a fractal component we can easily obtain the singularity spectrum

$$D(h) = d - \frac{\log\left(\rho(h)/\rho(h_1)\right)}{\log r_0}. \tag{17}$$

The main advantages of GMWP are its ability to provide good spatial localization of singularities, a direct access to this value for each point and its relatively easy implementation and simpler connection with theory, with more direct measures. According to what has been discussed, our implementation of the GMWP method runs as follows:

1. *Calculus of the singularity exponents* $h(\vec{x})$: We estimate the gradient by finite differences, then we take its modulus and the resulting function is wavelet-projected for every point at different resolutions $r_0, r_1, \ldots, r_n$. Applying Eq. (15), at each fixed point $\vec{x}$ we estimate the exponent $h(\vec{x})$ by a linear log–log regression (taking r as the variable).
2. *Obtention of the empirical histogram* $\rho(h)$.
3. *Evaluation of singularity spectrum from histogram*: We apply directly Eq. (17). In addition, we could also propagate errors to obtain error bars, as indicated in Section 7.

3.4. Gradient histogram method

The gradient histogram (GH) method is a crude but fast version of GMWP method, which has the additional advantage of providing an explanation to a frequently observed feature of derivative histograms, namely the existence of a small mode and heavy tails. The starting point for GH method is the multifractal measure μ, Eq. (14). Taking into account that $\mu(B_r(\vec{x})) = \int_{B_r(\vec{x})} \mathrm{d}\vec{x}'|\nabla s|(\vec{x}')$, for a discretized signal, the gradient at one point $\vec{x}$ can be approximated by the multifractal measure at that point and radius r_0 (where r_0 is the minimum scale), in the way

$$|B_1(0)|r_0^d|\nabla s|(\vec{x}) \approx \mu(B_{r_0}(\vec{x})) = \alpha_\mu(\vec{x}) r_0^{d+h(\vec{x})}, \tag{18}$$

where for the last equality we have made use of Eq. (14) and $\alpha_\mu(\vec{x})$ denotes the non-scaling constant in that relation. By statistical translational invariance,[2] $\langle \mu(B_{r_0}) \rangle = \alpha_0 r_0^d$ for an appropriate non-scaling factor $\alpha_0 = |B_1(0)|\langle|\nabla s|\rangle$ so we obtain

$$\log \frac{|\nabla s|(\vec{x})}{\langle|\nabla s|\rangle} = \log a(\vec{x}) + h(\vec{x}) \log r_0 \tag{19}$$

(where $a(\vec{x}) = \alpha_\mu(\vec{x})/\alpha_0$). Defining now the log-normalized gradient $\tilde{h}$,

$$\tilde{h}(\vec{x}) \equiv \frac{\log\left(|\nabla s|(\vec{x})/\langle|\nabla s|\rangle\right)}{\log r_0}, \tag{20}$$

it follows that

$$\tilde{h}(\vec{x}) = h(\vec{x}) + \frac{\log a(\vec{x})}{\log r_0}. \tag{21}$$

According to Eq. (21), the variable $\tilde{h}$ is a good estimate of the singularity exponent h provided a r_0 small enough to make the variations of the non-scaling term $\log a(\vec{x})$ negligible. As r_0 goes to zero the probability distributions of $\tilde{h}$ and h get closer and closer, so we obtain

$$\rho(\tilde{h}) \approx \rho(h) = K r_0^{d-D(h)} \approx K r_0^{d-D(\tilde{h})}, \tag{22}$$

where we have made use of Eq. (16). We can now proceed as in the previous method, applying the histogram formula, Eq. (17), and derive from it the singularity spectrum. The main advantages of GH method concern its great simplicity, both theoretical and numerical (which diminishes artifacts due to numerical effects). In addition, it directly connects multifractality with intermittency: intermittent variables usually have small values (and hence a small mode) but depart significantly from the average with a certain probability (giving rise to heavy tails). GH is a histogram-based method, so the spectrum is directly the Hausdorff spectrum and less distortions are expected (as a contrast with methods based on the estimation of p-moments, which are strongly affected by numerical instabilities associated to negative order moments). Our implementation of the GH method has been performed in the following way:

[2] Monofractals are not translational invariant and so require a different treatment.

1. *Evaluation of the histogram for the log-normalized gradient variable:* We compute finite differences at each point, taking their moduli. We estimate the average of the modulus, $\langle|\nabla s|\rangle$, over the sampled signal. From that, we evaluate the log-normalized gradient at each point in direct application of Eq. (20) and construct the histogram.
2. *Direct application of the histogram formula:* As for the GMWP method.

4. Definition of the benchmark

To test the previous methods we have created as a benchmark a set of multifractal signals with known singularity spectra. The signals have been generated with the model introduced by Benzi et al. [8]. Originally it was developed to explain some properties of turbulent flows [8,36–38]. It has been also very efficient in the study of the statistical properties of natural images [39]. One of the advantages of Benzi et al.'s model is that for synthetic signals the three singularity spectra coincide, so the generative model generates simple, well-behaved multifractals. The main disadvantage of the model is that it does not introduce intra-scale correlations, although extensions are possible. In the following section we briefly explain the model.

4.1. Benzi, Biferale, Crisanti, Paladin, Vergassola and Vulpiani's model for the generation of multifractal signals

Benzi et al.'s model works around a simple idea: to represent images in a wavelet basis [33] for which the multiplicative cascade, Eq. (2), takes place point by point; that is, the equation is not only valid in a statistical sense with an abstract variable η_{rL} but some corresponding spatial locations and scales can be connected by a physical variable $\eta_{rL}(\vec{x})$ which has the appropriate infinitely divisible distribution and is independent from the wavelet projection at scale L. Let us write Benzi et al.'s model more explicitly.

We will consider a wavelet Ψ defining a dyadic wavelet basis. As it is dyadic, any change in scale is a power of 2, and the positions are discretized in terms of the scale unit. We define the notation $\Psi_{j\vec{k}}(\vec{x}) = \Psi(2^j\vec{x} - \vec{k})$ with integer j, k_1 and k_2: $\vec{k} = (k_1, k_2)$. A synthetic signal $s(\vec{x})$ will be expressed as a combination of the elements of this basis, in the way

$$s(\vec{x}) = \sum_j \sum_{\vec{k}} \alpha_{j\vec{k}} \Psi_{j\vec{k}}(\vec{x}). \tag{23}$$

The wavelet coefficients $\alpha_{j\vec{k}}$ are not chosen in an arbitrary way, but in order to guarantee that $s(\vec{x})$ has a given multifractal structure. The coefficients $\alpha_{j\vec{k}}$ realize the multiplicative cascade between consecutive scales, in the way

$$\alpha_{j\vec{k}} = \eta_{j\vec{k}} \alpha_{j-1,[\frac{\vec{k}}{2}]}, \tag{24}$$

where the operator $[\cdot]$ means the vector with components the integer part of its argument. Eq. (24) is the discretized counterpart of Eq. (2), except for an explicit spatial dependence in the η variable. By construction, all the variables $\eta_{j\vec{k}}$ are taken independently of $\alpha_{j-1,[\frac{\vec{k}}{2}]}$ and having the same distribution. In fact the distribution of η is defined by its p-moments as

$$\langle \eta^p \rangle = \left(\frac{1}{2}\right)^{\tau_p}, \tag{25}$$

where the SS exponents τ_p are given by the multifractal spectrum to be implemented, Eq. (7), and the basis $\frac{1}{2}$ is a consequence of dealing with a dyadic representation (so $L/r = 2$ constantly). To construct a multifractal of known singularity spectrum using Benzi et al.'s model, you should apply Eq. (7) to retrieve τ_p and from τ_p derive the distribution of η using Eq. (25). As a possible alternative for obtaining the distribution of η, you could make use of Eq. (16) to retrieve the distribution of h and take $\eta = 2^{-h}$. Once the distribution of η's is known, you would start generating the values $\eta_{j\vec{k}}$ at random and independently according to that distribution, then use Eq. (24) to recursively construct the wavelet coefficients $\alpha_{j\vec{k}}$ and from them synthesize the signal as in Eq. (23). It can be proved that such synthetic signals $s(\vec{x})$ have the desired multifractal spectrum [8,36,37].

4.2. Test spectra

We have focused our analysis in two basic types of multifractal spectra, namely log-Poisson [12,40,41] and log-Normal [42]. They are the best to illustrate the differences between bounded and unbounded singularity spectra. Log-Poisson has been commonly used to model turbulence while log-Normal is always a good first-order approximation to any spectrum around its mode.

Log-Poisson spectra correspond to translationally invariant infinitely divisible processes, which under infinitesimal changes in scale become binomial. By construction, log-Poisson spectra are bounded below, that is, there exists a minimum value for the singularity exponents, that we will denote by h_∞.[3] This parameter, together with the dimension of the associated fractal dimension, D_∞, can be used to parameterize the spectrum. Log-Poisson spectra in dimension d spaces have the following form:

$$D(h) = D_\infty + (d - D_\infty)\omega(h)(1 - \log\omega(h)), \tag{26}$$

where

$$\omega(h) \equiv \frac{h - h_\infty}{(d - D_\infty)(-\log\beta)} \tag{27}$$

and the dissipation parameter $0 \leqslant \beta \leqslant 1$ is given by

$$\beta = 1 + h_\infty/(d - D_\infty). \tag{28}$$

The log-Poisson spectrum defined in the equation above corresponds to the following SS exponents τ_p:

$$\tau_p = h_\infty p + (d - D_\infty)(1 - \beta^p). \tag{29}$$

Log-Normal processes correspond to unbounded spectra of the simplest type. Log-Normal spectra are parabolic,

$$D(h) = d - \frac{1}{2}\left(\frac{h - h_m}{\sigma_h}\right)^2. \tag{30}$$

As for log-Poisson multifractals, log-Normal multifractals are also characterized by two parameters: the mean singularity h_m and the singularity dispersion, σ_h. The corresponding τ_p exponents are given by the following expression:

$$\tau_p = h_m p - \frac{1}{2}\sigma_h^2 p^2. \tag{31}$$

Log-Poisson multifractals have a minimum singularity greater than -1. Let us recall that we are talking about the singularities for the gradient, which are shifted by -1 with respect to those of the signal [5,28]. This means that all log-Poisson signals have a minimum singularity greater than 0: these series are bounded on any compact set, they are ''physical signals'' as they cannot diverge to infinity at any point. In contrast, series with log-Normal statistics are unbounded, what implies that for any interval you can (and you will) find points at which the signal diverges to infinity in the limit of infinite resolution. Of course, over discretized series no point diverges to infinity, but what you get is that some nearby points can differ by orders of magnitude and as you increase the level of resolution in Benzi et al.'s model these differences increase. Unbounded series are more difficult to deal with, and they carry on specific processing problems, as for instance the lack of integrability. In spite of the popularity of some unbounded multifractals as log-Normal multifractals, those models fail to describe some essential properties present in actual physical signals, and should be always taken as an approximation to real situations (see discussion in [7]).

[3] Notice that, according to the model [12,43,5] $-1 < h_\infty < 0$.

4.3. Construction of the benchmarks

We have designed a set of benchmarks to analyze how different variables affect the performance of the validation methods. The selected variables are the length L (the total number of points) of each individual series, the number N of series that will be processed together (forming a coherent statistical ensemble) and finally the type of spectrum, which in turn depends on the choice of the appropriate parameters. For this article, we have combined them to generate several datasets. The first set defines the *Illustration Benchmark*, constructed for illustration purposes, which consists of 12 ensembles representative of some typical choices of the three selected variables. The second set defines the *Size Benchmark* in which length and number of series vary exhaustively. The third set allows to define the *Spectrum Benchmark*, which consists of ensembles formed by just one series of maximum length but with varying type of spectrum and choice of parameters. The details of each benchmark are given below.

Illustration Benchmark: For each type of spectrum we have modified the parameters to generate narrow and wide spectra. Thus, we have denoted by PN and PW the narrow and the wide log-Poisson spectrum while NN and NW denote the narrow and the wide log-Normal spectra. For each case we have generated three ensembles of varying sizes. Table 1 summarizes the characteristics and labels associated to each ensemble in this benchmark.

Size Benchmark: For the construction of the size benchmark, we take again the four spectra considered above (namely PW, PN, NW and NN) but instead of the three choices of size we have previously selected, we let N and L to take any possible value in a given range. For N we consider the values 1, 10, 100, 1000 and 10000; L can take the values 1024, 4096, 16,384 and 65,536. We have thus 20 possible ensembles (5 values of $N \times 4$ values of L) for each spectrum, so that this benchmark totalizes 80 ensembles of varying sizes. Notice that the C-codes given in [9] are prepared to analyze extensive benchmarks of this type.

Spectrum Benchmark: This benchmark is formed by ensembles of constant size: $N = 1$ and $L = 65{,}536$. Here, we change the parameters defining the singularity spectra. For log-Normal spectra we have chosen all the possible combinations of h_m and σ_h, where $h_m = -1$, -0.5, 0, 0.5, 1 and $\sigma_h = 0.33$, 0.5, 1, 2 (20 possibilities). In the case of log-Poisson spectra, as there are constraints on the parameters ($1 \geqslant d - D_\infty > -h_\infty > 0$), one cannot vary them freely over all the range of possible values. We will use all the couples (h_∞,D_∞) compatible with the constraints above, where h_∞ can take the values -0.2, -0.4, -0.6, -0.8 and D_∞ takes the values 0, 0.2, 0.4, 0.6. This leads to 10 possible spectra.

4.4. Influence of the variables

In the design of the benchmarks, we have paid special attention to produce ensembles of varying length, number of series and type of spectra, as we know that they are going to influence the performance of the validation methods. The length L of the series is important if the method requires of a sequence long enough to uncover some range of exponents which could be absent due to the discretization. The number N of series has a different effect than L: as N increases, we gain access to negative dimension manifolds, and at the same the uncertainty in the positive dimension manifolds is decreased. The total size (that is, $L \times N$) would be the relevant variable if the method is very demanding in statistics, which is usually the case for methods based on the Legendre spectrum. The third relevant variable is the multifractal spectrum of the generated series. Some spectra covering a wide range of singularities are worst analyzed by methods which rely on too strong regularity conditions or require comparing quantities with are numerically difficult to control (for instance, when moments diverge very quickly).

Table 1
Illustration benchmark

	$N = 1$, $L = 16{,}384$	$N = 10$, $L = 4096$	$N = 100$, $L = 1024$
$h_\infty = -0.5$, $D_\infty = 0$	PW1	PW2	PW3
$h_\infty = -0.25$, $D_\infty = 0.25$	PN1	PN2	PN3
$h_m = 0.5$, $\sigma_h = 1$	NW1	NW2	NW3
$h_m = 0.25$, $\sigma_h = 0.33$	NN1	NN2	NN3

The boxes contain the labels associated to each ensemble.

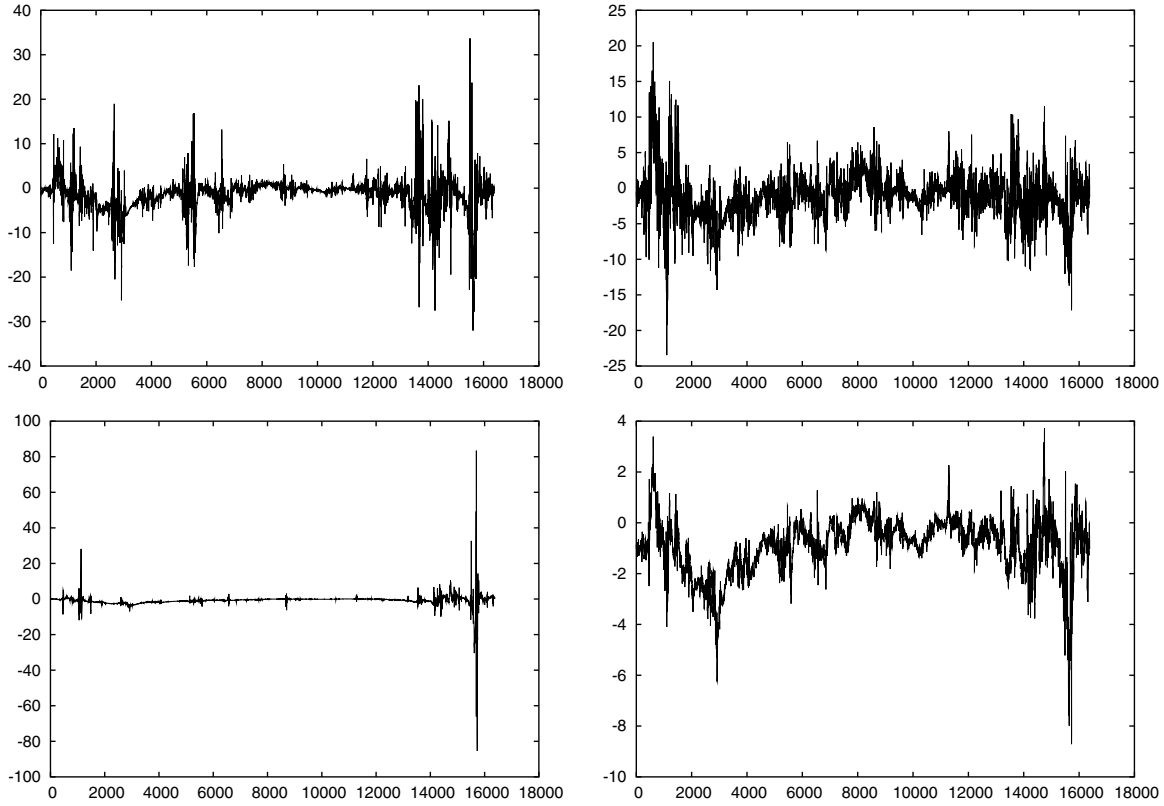

Fig. 1. Example series in the benchmark. Top: series from ensembles PW1 (left) and PN1 (right). Bottom: series from ensembles NW1 (left) and NN1 (right).

In Fig. 1, you can see the graphical representation of the long series for the four multifractal spectra in Illustration and Size Benchmarks. Series on the left side of Fig. 1 have wider spectra than those on the right. For the two series on top of the figure (log-Poisson type) the differences are not so dramatic, because log-Poisson multifractals have a singularity range bounded from below (which excludes explosions, bursts to infinity). On the contrary, the two series on bottom of the figure (log-Normal type) are unbounded and hence any value of singularity is allowed. This is dramatic for the series on the bottom left, that with wider spectrum, in which the orders of magnitude vary enormously in the same series. The log-Normal series on the bottom, right, has a narrow singularity spectrum and hence looks more physically reasonable. However, notice that a log-Normal series would eventually have bursts of any order if it is long enough.

5. Results: description of the general features of each method

5.1. M method on the Illustration Benchmark

In Fig. 2 we present the results of applying M method over the benchmark. To make the comparison easy, the theoretical spectra are plotted together with the corresponding estimates. In this implementation the range of scales considered for the regression runs from $r = 4$ points to $r = 100$ points. We have considered 65 moments non-uniformly sampled in the range $p \in [-4,8]$. The list is: $p_{\text{list}} = \{-4, -3.6, -3.2, -3, -2.8, -2.6, -2.4, -2.2, -2, -1.8, -1.6, -1.4, -1.2, -1.1, -1, -0.9, -0.8, -0.7, -0.6, -0.5, -0.4, -0.3, -0.2, -0.1, 0, 0.05, 0.1, 0.15, 0.2, 0.25, 0.3, 0.35, 0.4, 0.45, 0.5, 0.55, 0.6, 0.65, 0.7, 0.75, 0.8, 0.85, 0.9, 0.95, 1, 1.05, 1.1, 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9, 2, 2.3, 2.6, 3, 3.5, 4, 5, 6, 7, 8\}$.

This technique is very demanding in data, requiring larger datasets to obtain the parts of the spectrum corresponding to small dimensional components. This problem comes from the fact that we need to compute high

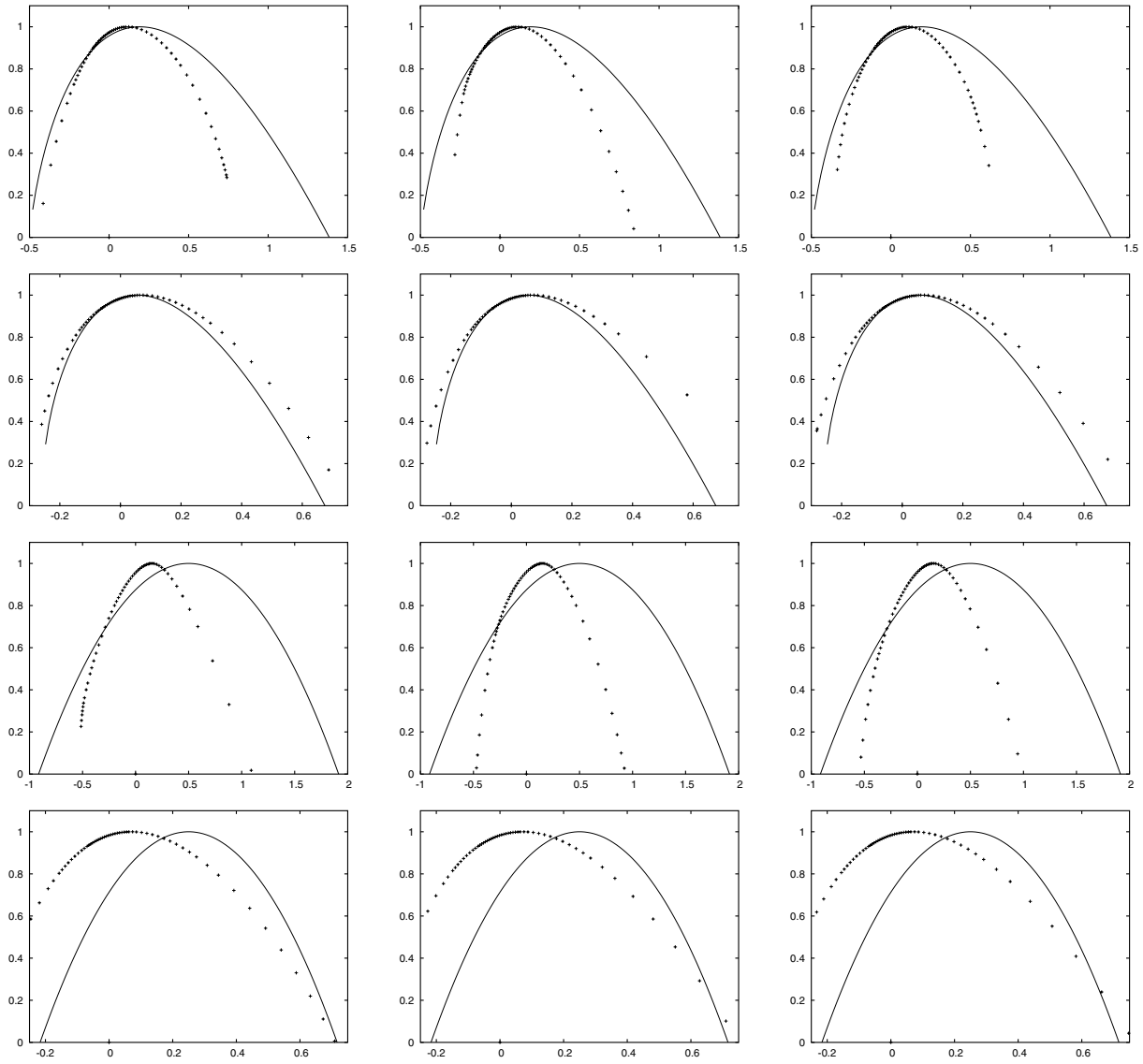

Fig. 2. Singularity spectra with M method. First row: ensembles PW1, PW2 and PW3; second row: PN1, PN2 and PN3; third row: NW1, NW2 and NW3; fourth row: NN1, NN2 and NN3. Solid lines: theoretical spectra.

order positive moments in order to sample the most singular values (left tail of the spectrum), while we need to sample high order negative moments to obtain the most regular values (right tail of the spectrum). Take into account that the distribution of ϵ_r has a large kurtosis even for the most uniform of our ensembles (see Fig. 3) and that, in addition, it has a very small mode. These two properties (high kurtosis, smallness of the mode) are characteristic to all multiplicative processes and make the evaluation of the histogram by means of the M method more complicated, as we next detail.

As the distribution of ϵ_r has high kurtosis, it has heavy tails which are difficult to estimate since the probability of observing events beyond a certain threshold is very small and many samplings must be made before getting a value on that region. Thus, this region is undersampled. Let us now suppose that the multifractal distribution is such that the variable ϵ_r has a finite maximum, denoted by $\|\epsilon_r\|_\infty$. For large values of p, the value of $\langle \epsilon_r^p \rangle$ is dominated by the maximum, $\langle \epsilon_r^p \rangle \sim \|\epsilon_r\|_\infty^p$. On the contrary, the value of the empirical estimate of the moment p, $\langle \epsilon_r^{\exp,p} \rangle$, is dominated by the empirical maximum M_r, so $\langle \epsilon_r^{\exp,p} \rangle \sim M_r^p$ for large p. Hence, the quality in the estimation of the most singular exponent h_∞, as $\|\epsilon_r\|_\infty \sim r^{h_\infty}$, depends on whether M_r provides a good estimate of

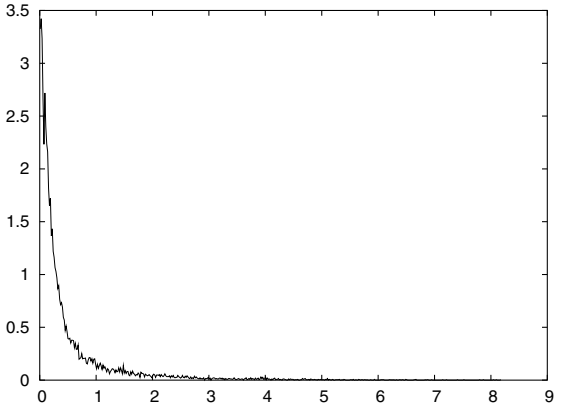

Fig. 3. Empirical histogram of ϵ_r for $r = 16$, obtained for Ensemble PW1.

$\|\epsilon_r\|_\infty$ or not. In general, M method will tend to truncate the left tail due to the difficulty of giving an accurate, steady estimate of $\|\epsilon_r\|_\infty$ in a broad range of scales. The situation is even more dramatic if the maximum is not finite, as the behavior of the estimates tend to deviate more and more as the order of the moment grows.

To understand the second problem affecting M method let us consider Fig. 3. The shape of the probability distribution of ϵ_r is a direct consequence of the multiplicative process associated to the multifractal model under consideration. All the distributions we are considering are unimodal, so there is only one mode (i.e., most probable value). As the mode is very close to zero and the variable ϵ_r is positive, the left side of the probability distribution changes very fast from zero to the highest probability (all the multifractals we have generated verify that the probability of $\epsilon_r = 0$ is strictly zero). The difficulties to resolve this narrow region between $\epsilon_r = 0$ and the mode affect the quality in the estimation of negative moments, which in turn affects to the quality of the estimates of positive singularity exponents and the right tail of the singularity spectrum (notice that the *left* tail of the histogram describes the *right* tail of the spectrum, i.e., the right part of $D(h)$ corresponds to moments $p < 0$). In fact, the issue "mode smallness" affects all the methods. For instance, notice that in many cases the right tail of the spectrum scales linearly, which suggests a linearization (due to lack of sampling) of the narrow range between 0 and the mode in the histogram of the corresponding measure.

Fig. 2 shows that M method provides a good estimation on bounded singularity spectra when they are narrow enough (ensemble PN), being much worst when the spectrum is broader. The problems are, in general, accentuated on the right tail due to the already discussed "mode smallness" issue. Another drawback characteristic to this method is the tendency to shift the curve, in the case of unbounded spectra. This problem seems an effect of the integrability issue discussed in Section 5.3. The question is that, in order to evaluate $\langle \epsilon_r^p \rangle$, we need to integrate over the whole space, that is, we assume that there is some weak type of translational invariance. Translational invariance implies $\tau_1 = 0$ for total support multifractals (i.e., the support of the multifractal is the whole real interval) [5], but in general log-Normal spectra have a value of $\tau_1 = h_m - \frac{1}{2}\sigma_h^2$ which is different from zero if $h_m \neq h_m^0$, where $h_m^0 = \frac{\sigma_h^2}{2}$. The way in which moments are evaluated seems to shift all the singularities by $h_m^0 - h_m$ in order to re-establish translational invariance; in fact, curves appear to be centered around h_m^0. In general, as one does not have prior knowledge about the actual value of h_m, there is no simple way to correct this problem for M method.

5.2. WTMM method on Illustration Benchmark

In Fig. 4 we present our results when applying the LastWave scripts to the data. The list of moments is identical to that used for the moments method, the analysis wavelet was that set as default (namely, second derivative of a Gaussian). There is a sensitive dependence of the performance of WTMM method in the choice of the minimum and maximum scales. For the Illustration Benchmark, we have carefully chosen the minimum and maximum scale, taking into account the size of the ensemble, its sampling rate and the analyzing wavelet, in order to get the best possible results. This procedure returned close but anyway different values for those scales depending on the

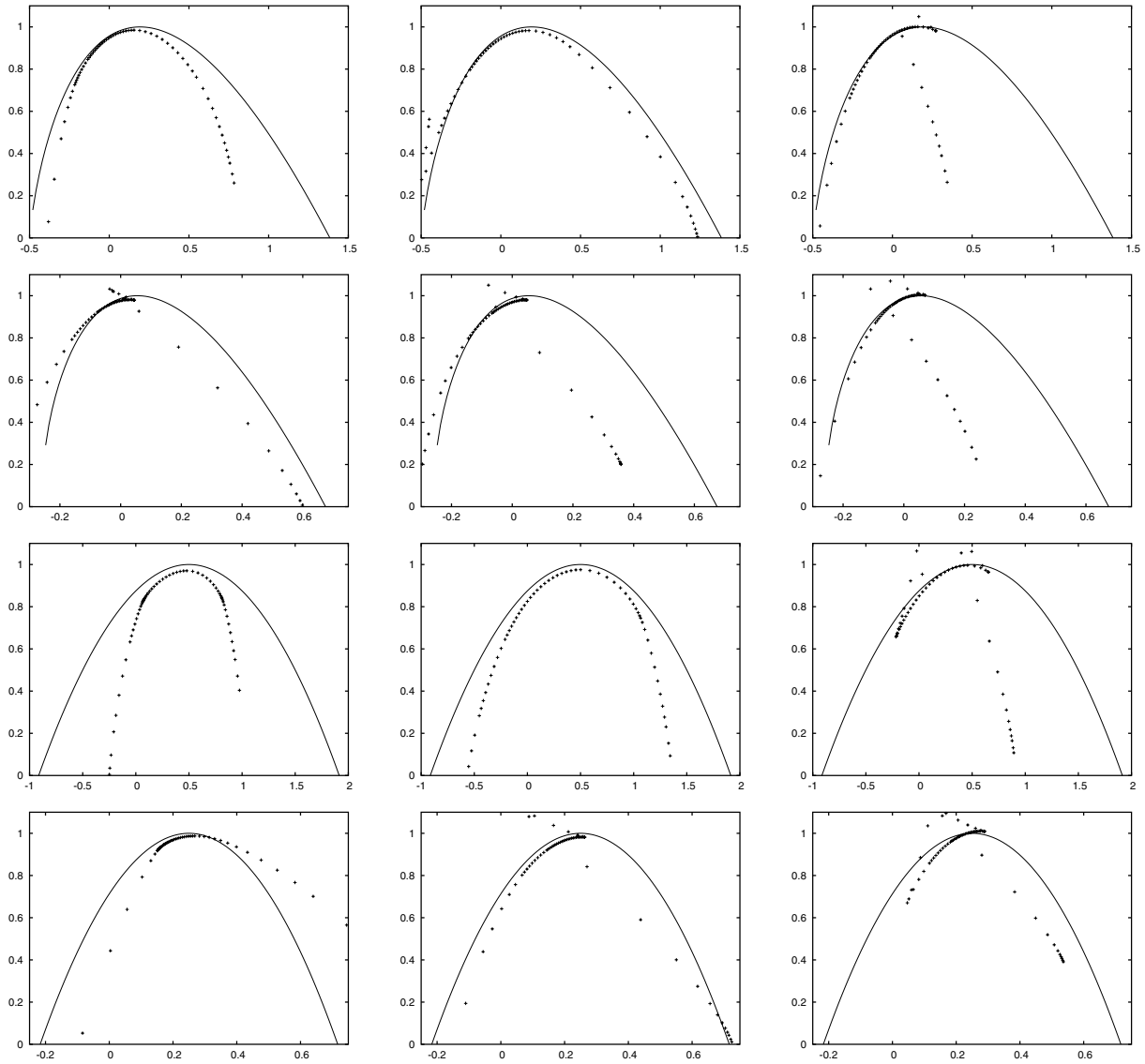

Fig. 4. Singularity spectra with WTMM method. First row: ensembles PW1, PW2 and PW3; second row: PN1, PN2 and PN3; third row: NW1, NW2 and NW3; fourth row: NN1, NN2 and NN3. Solid lines: theoretical spectra.

ensemble; in the scripts given in this paper web-page you can find the specific choices for each one of the ensembles in the benchmark. Notice that this method requires from expertised use and a considerable amount of supervision.

The curves show that, as for the moment method, WTMM method works better over narrow spectra. WTMM method generally provides a good description of the central part of the singularity spectrum, but in a very limited extent. WTMM method also linearizes the right tail, which can be connected with the "mode smallness" problem. Regarding the quality of the estimation of the right tail, this can be improved by increasing the number of vanishing moments of the wavelet, but then, it would induce some other severe constraints over the method (see Section 7.2 for a discussion about the choice of the wavelet). In addition, this method has a tendency to truncate the right part, an effect probably due to undersampling that we will next explain.

One of the theoretical, severe limitations of WTMM is the necessity of having isolated singularities, because maxima lines are determined as local maxima of the wavelet transform. If the analyzed signal is uniformly bounded and the wavelet integrable, at any finite scale the wavelet projections $T_\Psi s(\vec{x}, r)$ are continuous with $\vec{x}$, so the maxima at any scale are isolated (unless the projection becomes constant). A similar

argument – although a rather mathematically sophisticated one – can be used when the signal is not uniformly bounded. The question is that, over any bounded range of the signal (i.e., an interval) and a fixed scale r we have access to a finite number of maxima lines. But we work with discretized signals, so we have a minimum, finite scale r_0, given by the resolution, at which we must stop the analysis. As a consequence, for any sampled signal we will have a non-dense pack of maxima lines which end up at some particular points, but not at every one. We will be able to determine the singularity exponent for those points which are on the tip of a maxima line. The relative density of those ending points over the whole domain is determined by the type of wavelet we are using; see discussion in Section 7.2. Jaffard proposed [44] a refinement to WTMM which consists in restricting the computation of the partition function in the interval of width r to the largest maxima. As a consequence, in the case the signal has singularities everywhere, only one WT modulus maximum for each interval of length r should be taken into account in the sum (12). Some other attempts have been made to extend WTMM in order to obtain the singularity exponents of all the points in the sampling, but the theoretical and practical difficulties are still very high; see the extensive discussion in [23].

The undersampling resulting from the extraction of maxima lines has two consequences. First, the method becomes very demanding, as it dismisses a great part of the statistics (around 90% when using the second derivative of the Gaussian) which means that the less represented components are obviously the most affected ones. Second, less singular points tend to be underrepresented. As the maxima lines behave as $r^{h(\vec{x})}$ ($h(\vec{x})$: singularity exponent associated to the ending point) as $r \searrow 0$, a local maximum of $r^{h(\vec{x})}$ with respect to $\vec{x}$ corresponds to a local minimum of $h(\vec{x})$. A point with a large singularity h_0 will be a local minimum of $h(\vec{x})$ only if the corresponding neighborhood only contains even larger singularities. As the value of h_0 grows, this possibility is more and more unlikely, leading to the observed underrepresentation of larger singularities.

To end with the discussion on WTMM method, it is noticed that for short series as those of ensembles with labels ending in 3, disregarding the fact that those ensembles are the ones containing the largest statistics, the estimates become unstable for the widest spectra (see PW 3 and NW 3 in Fig. 4). We think that this problem is connected with the necessity in WTMM of having series large enough to evenly sample all the singularities.

5.3. GMWP method on Illustration Benchmark

In Fig. 5 we present the results obtained after our implementation of GMWP method. We have used a simple Gaussian function, which does not verify the "admissibility condition" required for representation wavelets [33], but which can anyway be used as an analysis wavelet. To obtain the singularity exponents at each point we have covered a large enough range of scales which ensures both a good spatial localization and accuracy in their numerical values; for this paper we will take a range going from $r = 1$ point to $r = 64$ points. Our choice is a good compromise, but the reader is encouraged to take our code, modify these values and perform her/his own experiments. The scales are exponentially sampled, starting in $r_0 = 1$ point, the ith scale being $r_i = r_0 Q^i$ for a scaling factor $Q > 1$ and such that $r_n = 64$ points. We have chosen $n = 6$ (hence $Q = 2$) in our experiences, but other values of n (equivalently, of Q) lead to similar results.

The limitations on GMWP method come from the fact that we must require the function (more properly, distribution) $|\nabla s|$ to be integrable, in order to have a well-behaved multifractal measure μ, defined as in Eq. (13). Notice that this problem is not exclusive of this method but also affects the others. When working with physical signals we should not worry too much about that, because any physical signal is bounded over compact sets [45,46] and hence the divergences in the derivative must be integrable. Some signals in the benchmark (notably, log-Normal ones) have unbounded singularity domains, which implies that there exist points with singularities below $h = -1$ and hence the gradient is not integrable. But those signals also affect the stability of methods such as WTMM, because even the signals themselves are not locally integrable.[4] There is a way to solve the integrability issue in the case of GMWP. Experimental exponents obtained from GMWP's of discretized signals with unbounded singularity range always appear shifted by a constant value η, such that the signal becomes integrable. A simple way to obtain the value of η is to study the average GMWP at different resolu-

[4] It should be noticed that despite of the fact that the log-Normal model is often used in the scientific literature the problem posed by the lack of integrability has not been discussed in depth.

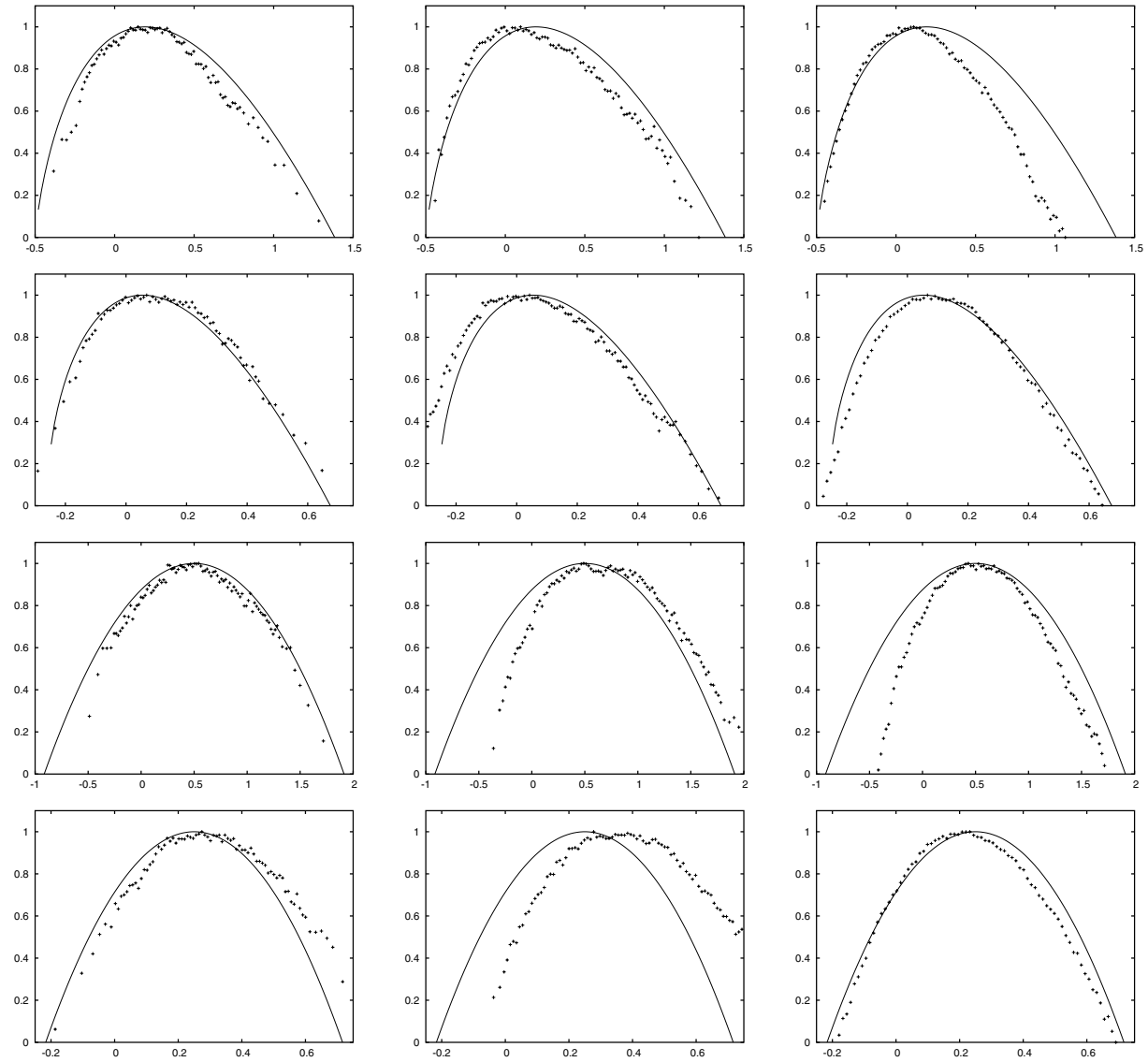

Fig. 5. Singularity spectra with GMWP method. First row: ensembles PW1, PW2 and PW3; second row: PN1, PN2 and PN3; third row: NW1, NW2 and NW3; fourth row: NN1, NN2 and NN3. Solid lines: theoretical spectra.

tions: the average GMWP should be the same at any scale, but due to the shift it evolves as power-law with exponent η. Once η is obtained, the raw singularities are shifted by $-\eta$. As you can see in Fig. 5, the results have a very good quality (it seems however the shift was not well corrected in the case of ensemble NN2). In fact, GMWP appears as the best method to obtain an estimate of the singularity spectrum (see Section 7).

5.4. GH method on Illustration Benchmark

In Fig. 6 we show the results for GH method. Not very much about the settings needs to be explained, as the main virtue of this method is its simplicity. Let us only note that the derivatives are estimated by finite differences. The main limitation of GH method comes from its tendency to linearize the right side of the spectrum. This behavior is a consequence of the ''mode smallness'' issue. The main virtue of the gradient method is its outstanding quality in the determination of the left tail of the curve. It provides a sampling on the left tail which is experimentally less dense than that of GMWP, but with greater accuracy.

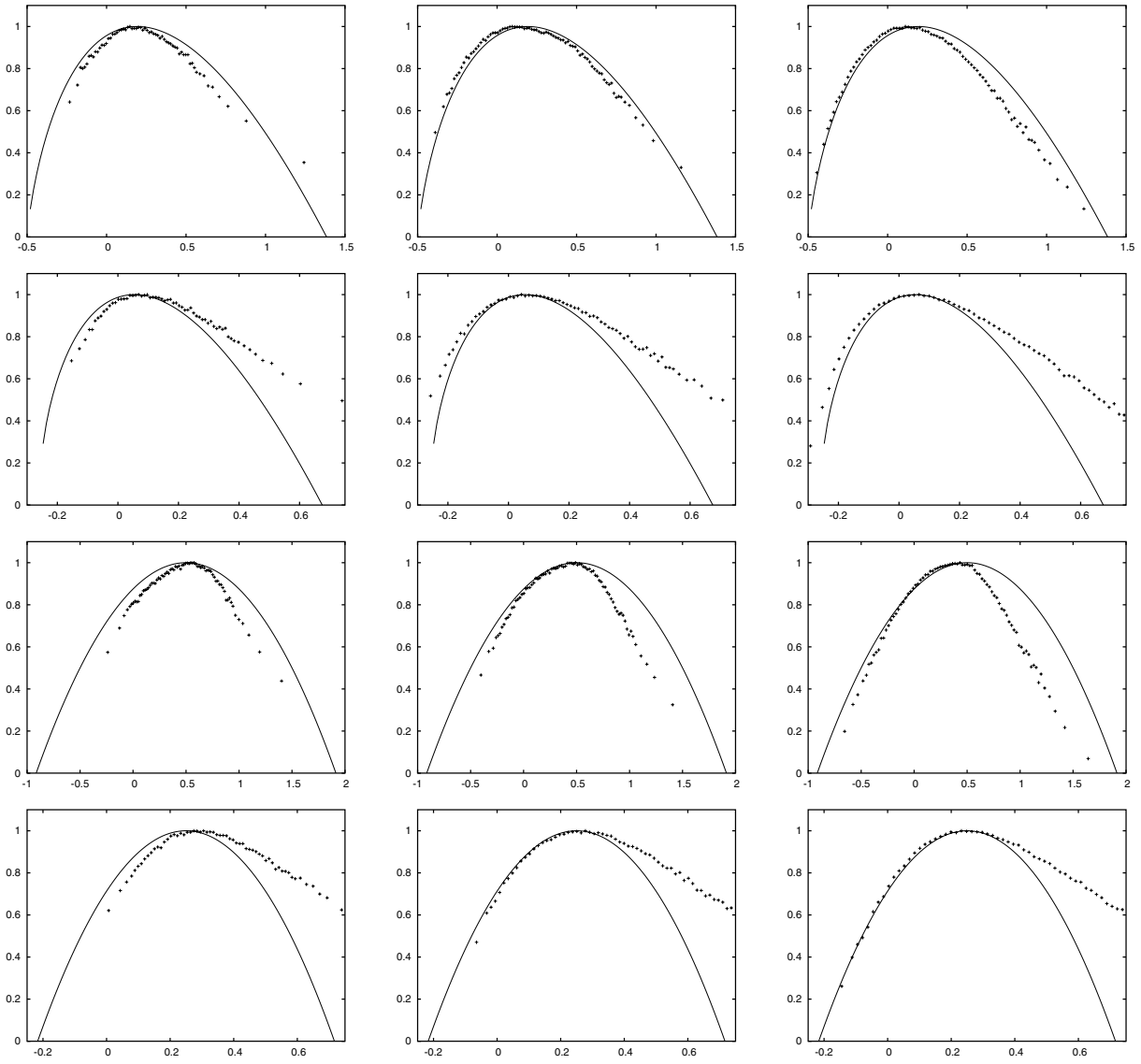

Fig. 6. Singularity spectra with GH method. First row: ensembles PW1, PW2 and PW3; second row: PN1, PN2 and PN3; third row: NW1, NW2 and NW3; fourth row: NN1, NN2 and NN3. Solid lines: theoretical spectra.

6. Results: dependence on statistics

6.1. Definition of the error measure

To perform a quantitative analysis of the quality of the outcomes we have defined as a measure of the error the difference between the estimated $D^e(h)$ and the actual singularity spectrum $D(h)$, i.e., the normalized average deviation E_e, given by

$$E_e = \frac{\langle |D(h) - D^e(h)| \rangle_{(h_{\min}, h_{\max})}}{h_{\max} - h_{\min}}, \tag{32}$$

where the average $\langle \cdot \rangle_{(h_{\min}, h_{\max})}$ is taken over the total range of singularities $(h_{\min}, h_{\max})$ that we want to describe. We need to fix the total range of singularities to allow comparisons among the results obtained with different statistics, because the range of singularities that a method can actually solve depends on the statistics. When a

method does not return a valid value of estimated spectrum for a value of h in the studied range, we assume that the estimated spectrum is 0 (which is our minimum reference value to consider a singularity manifold):

$$D^e(h^*) = 0 \quad \forall h^* \in (h_{\min}, h_{\max}) \text{ with no valid } D^e(h^*). \tag{33}$$

We fix the range of singularities by the two zero-crossings of the theoretical spectrum, $D(h_{\min}) = D(h_{\max}) = 0$. As the singularity spectra are concave, by construction $D(h) > 0\ \forall h \in (h_{\min}, h_{\max})$. So, the range of singularities that we will consider concerns all the singularity components which have positive dimension, so with probability 1 they can be found in any series (negative dimensions, as commented before, represent events happening with a certain probability and typically less than once per series, see [35]).

In addition, in the definition of normalized average deviation E, Eq. (32), we have normalized the average deviation by the size of the considered range of singularities. This should help to compare estimation errors among ensembles having spectra of different width: wider spectra imply a higher computational cost for the validation methods and greater average error, but at the same time they describe a wider range of singularities. By normalizing by the size of this range we obtain an estimation which is independent of the range width.

We will next compare the performance of the methods by tabulating the errors associated to each method, using normalized average deviations as estimates of those errors. To ease a fast visualization of the results, we have applied the following color convention for the boxes of the tables:

$E \leq 0.10$	$0.10 < E \leq 0.20$	$0.20 < E \leq 0.30$	$E > 0.30$
White box	Yellow box	Green box	Red box

6.2. Analysis of Size Benchmark

The Size Benchmark contains 80 ensembles and for each one we have applied the four validation methods. Tables 2–5 summarize the results for the ensembles in the Size Benchmark.

Table 2
Estimation errors for M method applied to Size Benchmark

	PW				*PN*			
	L= 1024	L= 4096	L= 16384	L= 65536	L= 1024	L= 4096	L= 16384	L= 65536
N=1	0.13	0.21	<u>0.20</u>	0.15	0.11	0.17	<u>0.07</u>	0.07
N=10	0.24	<u>0.22</u>	0.25	0.27	0.27	<u>0.14</u>	0.10	0.11
N=100	<u>0.25</u>	0.28	0.25	0.30	<u>0.10</u>	0.10	0.08	0.09
N=1000	0.23	0.23	0.22	0.18	0.09	0.09	0.09	0.09
N=10000	0.17	0.21	0.21	0.21	0.09	0.09	0.09	0.09
	NW				*NN*			
	L= 1024	L= 4096	L= 16384	L= 65536	L= 1024	L= 4096	L= 16384	L= 65536
N=1	0.18	0.18	<u>0.17</u>	0.17	0.25	0.25	<u>0.21</u>	0.27
N=10	0.17	<u>0.17</u>	0.18	0.18	0.25	<u>0.28</u>	0.24	0.24
N=100	<u>0.16</u>	0.17	0.17	0.19	<u>0.26</u>	0.29	0.24	0.25
N=1000	0.20	0.21	0.16	0.19	0.25	0.25	0.22	0.25
N=10000	0.24	0.19	0.16	0.19	0.24	0.23	0.26	0.26

Underlined values correspond to the spectra in the Illustration Benchmark.

Table 3
Estimation errors for WTMM method applied to Size Benchmark

	PW				*PN*			
	L= 1024	L= 4096	L= 16384	L= 65536	L= 1024	L= 4096	L= 16384	L= 65536
N=1	0.18	0.16	0.19	0.06	0.19	0.48	0.31	0.53
N=10	0.25	0.06	0.17	0.08	0.27	0.55	0.68	0.34
N=100	0.29	0.26	0.16	0.15	0.57	0.68	0.71	0.38
N=1000	0.19	0.23	0.18	0.08	0.59	0.65	0.58	0.56
N=10000	0.25	0.15	0.24	0.07	0.73	0.97	0.59	0.68
	NW				*NN*			
	L= 1024	L= 4096	L= 16384	L= 65536	L= 1024	L= 4096	L= 16384	L= 65536
N=1	0.08	0.12	0.18	0.18	0.16	0.35	0.23	0.35
N=10	0.08	0.10	0.10	0.22	0.32	0.29	0.41	0.35
N=100	0.19	0.06	0.13	0.17	0.49	0.38	0.37	0.35
N=1000	0.07	0.27	0.12	0.18	0.41	0.39	0.43	0.35
N=10000	0.08	0.28	0.27	0.24	0.41	0.87	0.45	0.35

Underlined values correspond to the spectra in the Illustration Benchmark.

Table 4
Estimation errors for GMWP method applied to Size Benchmark

	PW				*PN*			
	L= 1024	L= 4096	L= 16384	L= 65536	L= 1024	L= 4096	L= 16384	L= 65536
N=1	0.22	0.17	0.05	0.02	0.34	0.19	0.04	0.09
N=10	0.11	0.05	0.03	0.03	0.05	0.06	0.03	0.08
N=100	0.10	0.05	0.03	0.02	0.06	0.06	0.03	0.07
N=1000	0.09	0.04	0.03	0.02	0.05	0.01	0.03	0.07
N=10000	0.09	0.04	0.03	0.02	0.05	0.01	0.03	0.07
	NW				*NN*			
	L= 1024	L= 4096	L= 16384	L= 65536	L= 1024	L= 4096	L= 16384	L= 65536
N=1	0.14	0.05	0.04	0.04	0.29	0.22	0.09	0.18
N=10	0.08	0.08	0.05	0.08	0.11	0.28	0.04	0.10
N=100	0.08	0.04	0.05	0.08	0.08	0.05	0.10	0.08
N=1000	0.08	0.04	0.05	0.06	0.07	0.04	0.08	0.07
N=10000	0.09	0.04	0.05	0.06	0.09	0.05	0.06	0.07

Underlined values correspond to the spectra in the Illustration Benchmark.

Size Benchmark has been constructed in order to assess the quality of the methods depending on the parameters defining the size of the statistics (namely, N and L), and the convergence speed of the different estimates to the correct spectra, according to these parameters. We analyze the results method by method:

Table 5
Estimation errors for GH method applied to Size Benchmark

	PW				*PN*			
	L= 1024	L= 4096	L= 16384	L= 65536	L= 1024	L= 4096	L= 16384	L= 65536
N=1	0.27	0.18	0.11	0.12	0.27	0.15	0.19	0.19
N=10	0.13	0.06	0.03	0.02	0.12	0.14	0.12	0.13
N=100	0.04	0.03	0.02	0.01	0.12	0.12	0.12	0.12
N=1000	0.03	0.02	0.02	0.01	0.13	0.13	0.12	0.12
N=10000	0.03	0.02	0.02	0.01	0.13	0.12	0.12	0.12
	NW				*NN*			
	L= 1024	L= 4096	L= 16384	L= 65536	L= 1024	L= 4096	L= 16384	L= 65536
N=1	0.18	0.15	0.13	0.11	0.40	0.30	0.30	0.25
N=10	0.05	0.11	0.08	0.07	0.16	0.18	0.18	0.21
N=100	0.07	0.05	0.06	0.05	0.12	0.11	0.11	0.11
N=1000	0.05	0.05	0.05	0.04	0.10	0.10	0.10	0.10
N=10000	0.05	0.05	0.04	0.04	0.10	0.10	0.10	0.10

Underlined values correspond to the spectra in the Illustration Benchmark.

- *M method:* There is a weak tendency to error reduction when L or N is increased. However, the linearization of the right tails (for all types of spectra) and the displacement in the position of the curve (for log-Normal spectra) forces the error to saturate at rather large values. Hence, this method only gets a slightly beneficial effect of increasing the series length L or the amount N of series processed.
- *WTMM method:* It achieves a noticeable improvement with respect to M method for wide spectra, specially because this method seems to eliminate the displacement present in the M method. However, WTMM method is still affected by the linearization of the right tail, and this effect is much more important than in the M method for narrow spectra, so the results obtained by WTMM on PN and NN spectra saturate at larger values than those of M method. As a consequence, WTMM does not exhibit a decrease in the error when statistics is increased, but a convergence towards a given value result of the systematic deviation between the estimated right tail and the theoretical one.
- *GMWP method:* This method provides the best overall performance, and it is also the only method which exhibits a certain convergence to smaller errors when L or N are increased. This type of convergence is almost perfectly verified for increasing N. However, for increasing L this convergence is not always evident, specially for small values of N. This is not so surprising, because the series are generated independently at different L, and for small N the particular series obtained at a smaller length could be better estimated than those at the greater one, specially taking into account that the distributions of the wavelet coefficients in Benzi et al.'s model are very kurtotic. However, as N increases the particularities of the series should be less important and so the errors should converge to the same number. A case which deserves some comment is NN for $L = 4096$, $N = 10$, which coincides with the case NN2 in Illustration Benchmark. In this case, the estimate has an anomalously large error. This is a consequence of the bad correction of the shift induced by the non-integrability for that particular case, as shown in Fig. 5.
- *GH method:* This method has the second best overall performance, in spite of being affected by the linearization of the right tail. It also exhibits convergence with increasing L and N up to attaining a minimum error level, related with the linear right tail, that cannot be further diminished. This error level seems to be much reduced than those of M and WTMM methods.

6.3. Compared performances on Spectrum Benchmark

For Spectrum Benchmark we have generated 35 different singularity spectra. Tables 6–9 summarize the performance of the four validation methods on the Spectrum Benchmark.

Spectrum Benchmark was designed to show the sensibility of the methods under different types of spectra. The two characteristics of those spectra which seem to affect the different methods are the spectrum width and the average degree of singularity or regularity of the spectrum (that we measure as the singularity exponent associated to the manifold of maximum dimension). According to the method, we can make the following comments:

- *M method:* On the log-Normal statistics, the M method is strongly affected by the incorrect placing of the spectrum: the results appears always centered around $h = 0$. For that reason, the column referring to $h_m = 0$ is the one with the smallest value (except for $\sigma_h = 1$, because of the linearization of the right tail makes the configuration $h_m = 0.5$ more advantageous). For the log-Poisson statistics the errors are moderate, and in all instances induced by the linearization of the right tail.
- *WTMM method:* Errors on the log-Normal statistics are smaller than those of M method. The error is always a consequence of the linearization of the right tail. Wider spectra are better described than the narrow ones. For log-Poisson, errors are moderate except for the narrowest spectra, for which the linearization effect induce severe errors.
- *GMWP method:* Errors are moderate to small, except for the the narrowest log-Normal with the most singular mode (severe error) and the narrowest log-Poisson (medium error). Anyway, this method is once more the one attaining the best performance.
- *GH method:* Errors are important on the narrowest log-Normal series, and in the narrowest log-Poisson, due to the linearization effect. In the other cases the errors are moderate in average.

Table 6
Estimation error for M method over the Spectrum Benchmark

	Log-Normal						Log-Poisson			
	h_m −1	h_m −0.5	h_m 0	h_m 0.5	h_m 1		h_∞ −0.8	h_∞ −0.6	h_∞ −0.4	h_∞ −0.2
$\sigma_h = 0.33$	0.30	0.83	0.16	0.58	1.61	$D_\infty = 0$	0.15	0.20	0.19	0.26
$\sigma_h = 0.5$	0.35	0.52	0.19	0.31	0.55	$D_\infty = 0.2$	-	0.20	0.19	0.20
$\sigma_h = 1$	0.26	0.25	0.20	0.17	0.20	$D_\infty = 0.4$	-	-	0.27	0.14
$\sigma_h = 2$	0.14	0.14	0.14	0.14	0.13	$D_\infty = 0.6$	-	-	-	0.25

Table 7
Estimation error for the WTMM method over the Spectrum Benchmark

	Log-Normal						Log-Poisson			
	h_m −1	h_m −0.5	h_m 0	h_m 0.5	h_m 1		h_∞ −0.8	h_∞ −0.6	h_∞ −0.4	h_∞ −0.2
$\sigma_h = 0.33$	0.22	0.17	0.57	0.28	1.73	$D_\infty = 0$	0.17	0.17	0.29	1.08
$\sigma_h = 0.5$	0.07	0.15	0.26	0.24	0.27	$D_\infty = 0.2$	-	0.22	0.17	1.06
$\sigma_h = 1$	0.15	0.20	0.22	0.18	0.12	$D_\infty = 0.4$	-	-	0.16	0.82
$\sigma_h = 2$	0.15	0.14	0.14	0.13	0.12	$D_\infty = 0.6$	-	-	-	0.18

Table 8
Estimation error for GMWP method over the Spectrum Benchmark

Log-Normal						Log-Poisson				
	h_m −1	h_m −0.5	h_m 0	h_m 0.5	h_m 1		h_∞ −0.8	h_∞ −0.6	h_∞ −0.4	h_∞ −0.2
$\sigma_h = 0.33$	0.33	0.20	0.09	0.05	0.14	$D_\infty = 0$	0.11	0.05	0.04	0.25
$\sigma_h = 0.5$	0.14	0.11	0.10	0.02	0.15	$D_\infty = 0.2$	-	0.10	0.02	0.17
$\sigma_h = 1$	0.08	0.06	0.06	0.04	0.18	$D_\infty = 0.4$	-	-	0.05	0.09
$\sigma_h = 2$	0.05	0.04	0.03	0.04	0.09	$D_\infty = 0.6$	-	-	-	0.07

Table 9
Estimation error for GH method over the Spectrum Benchmark

Log-Normal						Log-Poisson				
	h_m −1	h_m −0.5	h_m 0	h_m 0.5	h_m 1		h_∞ −0.8	h_∞ −0.6	h_∞ −0.4	h_∞ −0.2
$\sigma_h = 0.33$	0.44	0.36	0.31	0.21	0.15	$D_\infty = 0$	0.12	0.14	0.08	0.30
$\sigma_h = 0.5$	0.26	0.22	0.19	0.13	0.12	$D_\infty = 0.2$	-	0.14	0.15	0.23
$\sigma_h = 1$	0.09	0.09	0.10	0.11	0.13	$D_\infty = 0.4$	-	-	0.16	0.17
$\sigma_h = 2$	0.06	0.07	0.08	0.08	0.09	$D_\infty = 0.6$	-	-	-	0.09

7. Discussion

7.1. Analysis of the compared performances

Let us now summarize the application domain, advantages and limitations of each method.

M method is a conceptually simple method, easy to implement but in general, rather demanding in data. In addition, the observed linearization of the right tail of the spectrum makes impossible to reduce the error below a bound, which is the error associated to approximate the right part of the curve by a straight line. Hence, any increase in the statistics can only decrease error up to the bound, then it does not longer evolve. In addition, for unbounded spectra there is another source of systematic error: spectra appear shifted. The method tends to shift the spectra and to produce a translational invariant equivalent: in fact, we have observed that for the datasets in Spectrum Benchmark all the evaluated spectra for a given value of σ_h roughly coincide, disregarding the value of h_m. This issue has no simple solution. The overall performance of this method is medium, with errors ranging from small to moderately large.

WTMM method is, by far, the one with the most involved technique, but some reasons justify its use under some circumstances. We next summarize four main reasons in favor of WTMM formalism. First reason, because of its representation capability: Maxima lines reduce the signal to an "essential" multiresolution representation. The different features show up either as scale invariant objects (unlimited maxima lines) or as fixed scale objects (maxima lines which start at the "detection" scale of the object). The argument does not favor WTMM method with respect to others, but the use of maxima lines can be justified as a "physically significant" representation. The second reason concerns its filtering capability: Non-intermittent signals have frequently large-scale contributions over imposed to the multifractal signature, which need to be filtered by (appropriate) wavelet projections. Anyway, notice that GMWP and GH methods are also capable to filter large-scale contributions. Third reason, because of the WTMM capability to correctly assess oscillating singularities: some particular spatial arrangement of singularities require to tightly follow the maxima lines in

order to obtain a correct estimate of the singularity exponents [47]. Finally, because of some mathematical studies, which have proved that the WTMM method is appropriate to obtain the singularity spectra on a special type of functions [48], while for such functions M method only gives partial access to the correct spectra.

It is thus evident that the WTMM method has strong mathematical foundations and a marked capability to filter spurious tendencies with oscillating singularities. Besides, this method implies a considerable improvement with respect to the M method, which was the historical reason for its creation. However, this method has many problems in practice and only attains a medium to poor performance. First, it is difficult to correctly tune the WTMM parameters and an important degree of expertise on the method is required. Second, the WTMM method is very demanding in data: maxima lines can represent less than 10% of the total number of points. Third, estimates are unstable partly due to problems in the sampling of the maxima lines and partly due to the necessity of attaining a high numerical accuracy and data size to evaluate high order positive and negative moments. Fourth, this method also linearizes right tails, so errors saturate in a level associated to the systematic deviation between the straight line and the actual spectrum. Concerning the sensitivity to the variables defining the ensembles, errors show no clear tendency when size of length of the series are modified. This strange behavior calls into question the assumed additivity of the partition functions: in some of the ensembles analyzed here, when we let the number of series to grow, a few new contributions which introduce maxima lines associated to high order singularities completely distort the behavior of the partition function at negative moments and the right tail of the spectra appears corrupted. There is no simple way to correct, and the problems when dealing with negative moments has been recognized as one of the main problems of moment-based methodologies. Concerning performance, the method provides a medium performance on wide spectra, while it has a poor performance on narrow spectra, mainly due to the fact that the right tail is quite often corrupted. This method should be thoroughly tested in order to understand these limitations and solve them.

GMWP method gives a precise, stable estimation of singularity spectra, with the additional capability of estimating error bars (see Section 7.3). This method has an important bonus with respect to the other: it gives access to point information ($h(\vec{x})$). This method is the most stable one and possesses the best global performance on the benchmark. Increases in size and length tend to diminish error, and the quality of the estimates is slightly better for wide spectra than for narrow spectra. It can be affected by non-integrable spectra (for instance, the case of NN spectrum for $N = 10$, $L = 4096$, which corresponds to NN2 in Illustration Benchmark).

GH method, in spite of being extremely simple and reproducible gives an excellent estimation of the left tail of the singularity spectra; however, it systematically linearizes the right tail. It allows to connect experimental properties (as intermittency) with the multifractal theory. Any increase in statistics tend to reduce the error committed by this method, up to the bound associated to the linearization of the right tail; this lower bound, however, is significantly smaller than those of M and WTMM methods, and for that GH method performs better than them. For GH method, errors are greater over narrow spectra than over wide spectra, because the linearization of the right tail induces a faster accumulation of error.

7.2. Projecting the signal or projecting the modulus of its gradient: a key issue

One of the main reasons of the improvement in performance of GWMP with respect to WTMM comes from the fact that the wavelets we have used to project the modulus of the gradient are of lower order (i.e., have smaller number of zero-crossings) than those which are required to perform WTMM. As we have seen, WTMM requires isolating maxima lines, which implies being able to resolve (distinguish) the wavelet projections from two close points. For ideal, continuous series we would never have this problem, but we need to work with real, discretized data and the actual value of wavelet projections can only be approximated, and they can be accurately estimated only for a precise range of scales. The minimum distance at which two maxima lines can appear is directly proportional to the minimum spread of the wavelet, the later being defined as the minimum size at which empirical wavelet projections are reasonably well estimated. The minimum spread is, at least, of the order of the discretization size of the series, the *series box*, but some wavelets have larger minimum spreads, see Fig. 7. The point is that we cannot allow a significant amount of the area of a positive peak and of a negative valley to be contained inside the same series box, because their contributions would almost cancel, anomalously reducing the contribution of that series box. Hence, any analysis must stop at

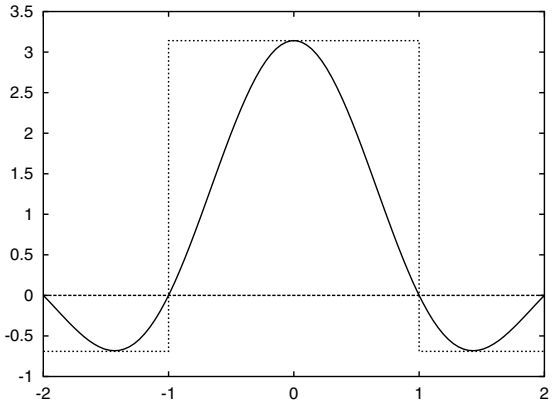

Fig. 7. A wavelet with two zero crossings will have, at least, a minimum spread of three or four series boxes.

scales greater than the minimum spread. In the case of the results shown in this paper using WTMM method, as the minimum spread of the second derivative of the Gaussian is four or five pixels, it is rather natural to observe maxima lines separated at least 10 boxes (see Fig. 8).

Let us now examine the maxima lines from WTMM method for an application example in Fig. 8. Observe the spatial regularity in the line arrangement. The WTMM is probably isolating a maxima line each time it can resolve it from the precedent one. In addition, for the analyzed statistics, maxima lines are almost x-independent, which confirms that, for this example, it is not necessary to follow maxima lines in order to compensate oscillations. Changing the signal by the gradient removes the difficulty of dealing with high-order wavelets (i.e., wavelets with many zero-crossings) at the cost of introducing the difficulty of finding a good estimate of the derivative. We used a simple definition for the derivative, the finite difference (i.e., $s'(x) = s(x+1) - s(x)$). This definition is related to the linear increment formalism originally derived for the assessment of multifractal behavior in turbulent flows [11]: the modulus of the gradient $|s'|(x)$ is approximated by the linear increment at the distance of one resolution box, $|s(x+1) - s(x)|$. At the resolution scale the relative importance of long-range correlations must be negligible if the signal can be assessed as multifractal. The use of the non-linear function "absolute value" is crucial; without it, the increments would sum up and the final results would be dominated by differences of the signal at the finite scale r, which are indeed affected by long-range correlations.

7.3. Error control in histogram-based methods

GMWP and GH methods offer the advantage of enabling a certain control on the sampling error incurred by fluctuations in the statistics. By a simple reasoning on the construction of the empirical histograms, we can associate an error bar to each point in the singularity spectrum. The argument goes as follows.

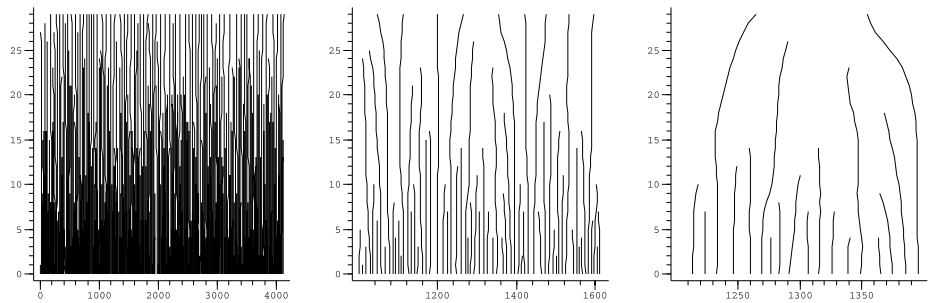

Fig. 8. Maxima lines for ensemble PW1 (log-Poisson $h_\infty = -0.5$, $d - D_\infty = 1$, in a single 16,384-point series), in three different zoom ranges.

Given a random variable h, a histogram box can be interpreted as a continuous range of values $I_\alpha = [h_\alpha, h_{\alpha+1})$ with associated probability p_α. Disregarding the other boxes in the histogram (that is, taking them as independent, which is reasonable if the number of boxes is large) the probability for a single point of being inside the box I_α is p_α. If we have a sample of N independent realizations of h, $\{h_i\}_{i=1,\ldots,N}$, we can construct a simple estimator of p_α as $\hat{p}_\alpha = N_\alpha/N$, where N_α is the number of values of h falling inside I_α. The random variable N_α, by construction, has a binomial distribution of parameter p_α. Hence, if we denote by $\langle\cdot\rangle$ the expectation values then we have that $\langle N_\alpha\rangle = N\,p_\alpha$ and $\sigma^2_{N_\alpha} = \langle N_\alpha^2\rangle - \langle N_\alpha\rangle^2 = Np_\alpha(1-p_\alpha) \approx Np_\alpha$, the last approximation being valid when $p_\alpha \ll 1$. If $Np_\alpha > 30$ we can approximate the binomial by a Gaussian. We will assume that N is large, that p_α is very small and that N_α is very close to its mean value, so we concentrate in cases such that $N_\alpha > 30$. Under those conditions, $\hat{p}_\alpha$ has a normal distribution of mean p_α and variance p_α/N. Now, it is very easy to assign a significant error bar to the measurements by confidence intervals. We will seek a $\sim$99% confidence, what implies taking 3 standard deviations around the mean in a Gaussian, so we state that

$$\hat{p}_\alpha \in \left(p_\alpha - 3\sqrt{\frac{p_\alpha}{N}}, p_\alpha + 3\sqrt{\frac{p_\alpha}{N}}\right) \quad \text{with confidence of } 99\%, \tag{34}$$

which we simplify to

$$\hat{p}_\alpha \in p_\alpha\left(1 - 3\sqrt{\frac{1}{Np_\alpha}}, 1 + 3\sqrt{\frac{1}{Np_\alpha}}\right) \tag{35}$$

and now we make the approximation $Np_\alpha \approx N_\alpha$ to finally obtain

$$\hat{p}_\alpha \in (\hat{p}^0_\alpha - \delta\hat{p}_\alpha, \hat{p}^0_\alpha + \delta\hat{p}_\alpha) = p_\alpha\left(1 - 3N_\alpha^{-\frac{1}{2}}, 1 + 3N_\alpha^{-\frac{1}{2}}\right), \tag{36}$$

where the confidence interval for $\hat{p}_\alpha$ has been expressed in terms of an uncertainty radius $\delta\hat{p}_\alpha$. Notice that the center of the confidence interval, $\hat{p}^0_\alpha$, is exactly p_α. Something interesting about Eq. (36) is that the uncertainty radius is proportional to the value of p_α, and so it can be expressed as a percentage error over p_α. Assuming that $N_\alpha^{-\frac{1}{2}}$ is small enough, we can obtain a relation analogous to Eq. (36) for the estimation of p_α from the value of $\hat{p}_\alpha$ and N_α

$$p_\alpha \in (p^0_\alpha - \delta p_\alpha, p^0_\alpha + \delta p_\alpha) = \hat{p}_\alpha\left(1 - 3N_\alpha^{-\frac{1}{2}}, 1 + 3N_\alpha^{-\frac{1}{2}}\right), \quad 3N_\alpha^{-\frac{1}{2}} \ll 1 \tag{37}$$

from which we conclude that the confidence interval for the estimation of the parameter p_α is centered around $p^0_\alpha = \hat{p}_\alpha$ and has an uncertainty radius $\delta p_\alpha = 3\hat{p}_\alpha N_\alpha^{-\frac{1}{2}}$. In this derivation we have assumed that $Np_\alpha \sim N_\alpha$ and an explicit use of the following constraints: $p_\alpha \ll 1$, $N \gg 30$, $N_\alpha \geqslant 30$ and $3N_\alpha^{-\frac{1}{2}} \ll 1$. Under such requirements, the interval defined in Eq. (37) is significant up to a 99% of confidence. We take this estimate as a first estimate of the error bars, and defer the study of intervals with small N_α to a specific analysis. Once the error bars for the histogram are known, it is quite simple to propagate the error on Eq. (17) to find the error bars associated to any spectrum obtained from the histogram formula. Thus if we know the error in a value of the histogram, $\rho(h) \pm \delta\rho(h)$, then by using Eq. (17) we get the error in the spectrum $D(h) \pm \delta D(h)$ as

$$\delta D(h) = \frac{1}{\log r_0}\frac{\delta\rho(h)}{\rho(h)}. \tag{38}$$

From Eq. (37) we have that $\delta\rho(h) = 3\rho(h)N_h^{-\frac{1}{2}}$, which substituted in the preceding equation yields

$$\delta D(h) = \frac{1}{\log r_0} 3N_h^{-\frac{1}{2}}. \tag{39}$$

To give some intuition on the figures, let us suppose that for a given value of h we record $N_h = 81$ events. In this case, $3N_h^{-\frac{1}{2}} = 0.33$, what means a percentual variation of 33% in the value of the probability. In terms of error in the singularity spectrum, for a 1024-record series the value of $\delta D(h)$ would be $\approx$0.05, and for a 16384-record series the uncertainty falls to $\delta D(h) \approx 0.035$. Conversely, to attain a precision of $\delta D(h) = 0.01$ would imply a sampling of $N_h = 1874$ events for a 1024-record series, and of $N_h = 956$ for a 16384-record series. The values indicate that we can have a moderately good precision, of order 0.05, in the determination of

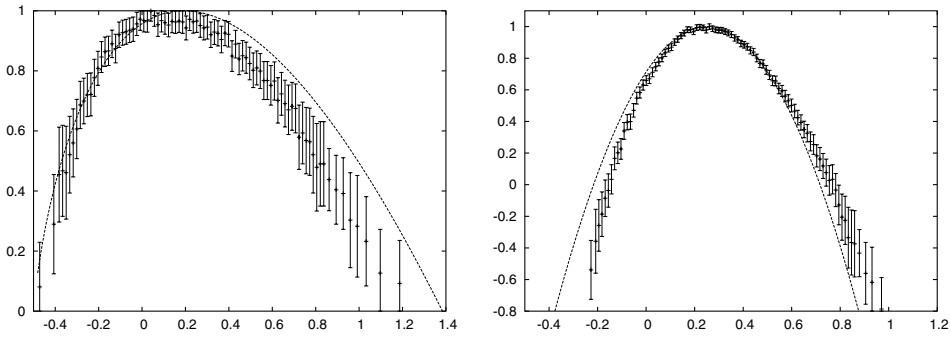

Fig. 9. Examples on the estimation of error bars in a validation method based on histogram estimation. The graphs correspond to the estimated singularity spectra, including error bars, for ensembles PW1 and NN3 using GMW. The continuous lines are the theoretical spectra.

the spectra even for moderately long series, but they also imply that a large increase in the statistics only leads to moderate improvements in accuracy. In Fig. 9 we show two examples on the estimation of error bars for the singularity spectra in a histogram-based validation method (GMWP). As expected, error bars decrease in the parts with greater probability and increase for the least probable singularity values, which would require larger and larger datasets in order to allow an accurate estimation of the associated fractal dimension. The advantage of obtaining explicit error bars lies in the fact that they give some intuition on the quality of the estimate. It should be noticed, however, that they can only inform about the deviations which are of random nature; any systematic source of error will not be counted.

8. Conclusions

In this paper we have tested the performance of four different methods employed in the evaluation of multifractal spectra from experimental data. Some of the methods allow for several possible different implementations. We have concentrated in one particular, general enough realization for each one of the methods involved. Doing so we pretend to give a precise idea on the particularities, scope and limitations of each methodology, beyond its precise work around. In the same spirit, we have discussed the causes of the observed deviations, trying to explain them as consequences of the method under study. Our goal has been to make explicit the requirements on the data under analysis that should direct our choice of one method before another.

The four validation methods we have tested can be named as moment (M) method, wavelet transform modulus maxima (WTMM) method, gradient modulus wavelet projection (GMWP) method and gradient histogram (GH) method. Each one is adapted to a different aspect of the theory and/or data type. We have produced a benchmark of synthetic multifractal 1D series as test data ensembles. The synthesized signals have known spectra, which makes it very easy to compare experimental and theoretical spectra. We have thus systematically made an assessment of each method for different data lengths, ensemble sizes and type of theoretical spectra.

Concerning the performance depending on the properties of the benchmark, we have seen that all the methods always give better estimates of the left part of the spectra than of the right part. Besides, as the spectrum width increases the quality on its determination decreases (which is really dramatic for methods such as WTMM). We have also observed that the high kurtosis of the histograms associated to multifractal variables is at the origin of the ''mode smallness'' issue, which is responsible to the marked tendency of all the validation methods except GMWP to linearize the right part of the singularity spectra. In addition, when dealing with unbounded spectra, new problems arise, as the presence of constant shifts due to the lack of integrability of the measured quantities. Concerning the compared performance, the main conclusion of our study is that GMWP seems to be the method with the best overall performance and stability; in addition, this method provides error bars on the estimation.

Concerning the individual performances, we see that *M method* has a medium performance. The method is very demanding, and requires large datasets to infer the tails. The problems with M method are that it has a tendency to linearize the right tail and it shifts unbounded spectra. Concerning *WTMM method*, it has uneven

performance. This method is even more demanding than M method. The quality on the estimate is very affected by series length and so the results are very unstable when data series are short, even with large datasets. The problems specific to WTMM include the linearization of the right tail and that the implementations may produce non-convex curves. In the results presented here, we have used one particular, freely accessible implementation (*LastWave*), but we have also tested other software tools (WaveLab, FracLab, etc.) obtaining slightly different results but with the same main difficulties. We have even implemented our own version of the method in the code we offer along this paper, see [9], and the results obtained were very similar. The problems with WTMM method come from the cost of having a technique suitable to deal with oscillating singularities. It seems that the beautiful mathematical formulation needs to be refined in several ways to eliminate some artifacts. Moreover, present implementations introduce a lot of parameters that require a careful, precise tuning. Concerning *GMWP method*, it has a good overall performance. It does not linearize the right tail of spectra. As a bonus, it allows to estimate error bars (this is a feature that we have not used in this paper). Finally, concerning *GH method*, it has very good performance on the left tail of the spectra, and very bad on the right tail, which is always linearized. As GMWP method, GH method also allows for error bar estimation.

About the possibilities of improving the existing methods, it should be remarked that in bounded multifractals the left tail of the singularity spectrum contains enough information to determine the whole spectrum. The left tail is connected with the positive p moments and if the positive moments do not diverge with p too fast the whole distribution can be retrieved from them [49]. As a matter of fact, bounded multifractals correspond to real, physical signals while unbounded multifractals are essentially mathematical objects. So, in practical applications it will be always more critical to attain a good estimation of the left tail of the spectrum than of the right tail. All the methods validated in this paper provide a better determination of the left tail than of the right part. Hence, it should be possible to process the information on the left tail and to correct any deviation on the right tail. Such a procedure could be used to improve all the methods discussed in this work, and this is an interesting, open possibility.

Acknowledgments

A. Turiel is supported by a Ramón y Cajal contract from the Spanish Ministry of Education and Science. J. Grazzini is funded by a Marie-Curie post-doctoral fellowship from the European Commission (Contract No. HPMD-CT-2001-00121). This work has been partially funded by the project BFM2003-08258-C02-02.

References

[1] A. Arneodo, C. Baudet, F. Belin, R. Benzi, B. Castaing, B. Chabaud, R. Chavarria, S. Ciliberto, R. Camussi, F. Chilla, B. Dubrulle, Y. Gagne, B. Hébral, J. Herweijer, M. Marchand, J. Maurer, J.F. Muzy, A. Naert, A. Noullez, J. Peinke, F. Roux, P. Tabeling, W. vandeWater, H. Willaime, Structure functions in turbulence, in various flow configurations, at Reynolds number between 30 and 5000, using extended self-similarity, Europhys. Lett. 34 (6) (1996) 411–416.
[2] A. Turiel, G. Mato, N. Parga, J.P. Nadal, The self-similarity properties of natural images resemble those of turbulent flows, Phys. Rev. Lett. 80 (1998) 1098–1101.
[3] P. Ivanov, L. Amaral, A. Goldberger, S. Havlin, M. Rosenblum, Z. Struzik, H. Stanley, Multifractality in human heartbeat dynamics, Nature 399 (1999) 461–465.
[4] B.B. Mandelbrot, A. Fisher, L. Calvet, A multifractal model of asset returns, Cowles Foundation Discussion Paper No. 1164, 1997.
[5] A. Turiel, N. Parga, The multi-fractal structure of contrast changes in natural images: from sharp edges to textures, Neural Comput. 12 (2000) 763–793.
[6] A. Turiel, N. Parga, Multifractal wavelet filter of natural images, Phys. Rev. Lett. 85 (2000) 3325–3328.
[7] A. Turiel, C. Pérez-Vicente, Multifractal geometry in stock market time series, Physica A 322 (2003) 629–649.
[8] R. Benzi, L. Biferale, A. Crisanti, G. Paladin, M. Vergassola, A. Vulpiani, A random process for the construction of multiaffine fields, Physica D 65 (1993) 352–358.
[9] The webpage associated to this paper, http://www.icm.csic.es/geo/gof/people/turiel/SUPP_INFO/multifractal_comparison.html.
[10] A.N. Kolmogorov, Dissipation of energy in a locally isotropic turbulence, Dokl. Akad. Nauk. SSSR 32 (1941) 141.
[11] U. Frisch, Turbulence, Cambridge University Press, Cambridge, MA, 1995.
[12] E.A. Novikov, Infinitely divisible distributions in turbulence, Phys. Rev. E 50 (1994) R3303.
[13] G. Parisi, U. Frisch, On the singularity structure of fully developed turbulence, in: M. Ghil, R. Benzi, G. Parisi (Eds.), Turbulence and Predictability in Geophysical Fluid Dynamics, Proceedings of the International School of Physics E. Fermi, North Holland, Amsterdam, 1985, pp. 84–87.

[14] R. Riedi, I. Scheuring, Conditional and relative multifractal spectra, Fractals. Interdisciplin. J. 5 (1997) 153–168.
[15] A. Turiel, A. del Pozo, Reconstructing images from their most singular fractal manifold, IEEE Trans. Im. Proc. 11 (2002) 345–350.
[16] K. Falconer, Fractal Geometry: Mathematical Foundations and Applications, Wiley, Chichester, 1990.
[17] J.F. Muzy, E. Bacry, A. Arneodo, Wavelets and multifractal formalism for singular signals: application to turbulence data, Phys. Rev. Lett. 67 (1991) 3515–3518.
[18] J. Grazzini, A. Turiel, H. Yahia, Entropy estimation and multiscale processing in meteorological satellite images, in: Proceedings of ICPR 2002, vol. 3, 2002, pp. 764–768.
[19] S. Mallat, A theory for multiresolution signal decomposition: the wavelet representation, IEEE Trans. Pattern Anal. Mach. Intelligence 11 (1989) 674–693.
[20] S. Mallat, S. Zhong, Characterization of signals from multiscale edges, IEEE Trans. Pattern Anal. Mach. Intelligence 14 (1992) 710–732.
[21] E. Bacry, J.F. Muzy, A. Arneodo, Singularity spectrum of fractal signals from wavelet analysis: exact results, J. Stat. Phys. 70 (1993) 635–673.
[22] A. Arneodo, F. Argoul, E. Bacry, J. Elezgaray, J.F. Muzy, Ondelettes, multifractales et turbulence, Diderot Editeur, Paris, France, 1995.
[23] Z.R. Struzik, Determining local singularity strengths and their spectra with the wavelet transform, Fractals 8 (2) (2000) 163–179.
[24] S. Mallat, S. Zhong, Wavelet transform maxima and multiscale edges, in: R.M. B. et al. (Eds.), Wavelets and their Applications, Jones and Bartlett, Boston, 1991.
[25] A. Arneodo, Wavelet analysis of fractals: from the mathematical concepts to experimental reality, in: G. Erlebacher, M.Y. Hussaini, L. Jameson (Eds.), Wavelets. Theory and Applications, ICASE/LaRC Series in Computational Science and Engineering, Oxford University Press, Oxford, 1996, p. 349.
[26] A. Chhabra, R. Jensen, Direct determination of the f(alpha) singularity spectrum, Phys. Rev. Lett. 62 (1989) 1327–1330.
[27] A. Arnéodo, E. Bacry, J.F. Muzy, The thermodynamics of fractals revisited with wavelets, Physica A 213 (1995) 232–275.
[28] P. Kestener, A. Arnéodo, Three-dimensional wavelet-based multifractal method: the need for revisiting the multifractal description of turbulence dissipation data, Phys. Rev. Lett. 91 (19) (2003) 194501.
[29] E. Bacry et al., LastWave, http://www.cmap.polytechnique.fr/~bacry/LastWave.
[30] D. Donoho et al., WaveLab, http://www-stat.stanford.edu/~wavelab/.
[31] J. Lévy Vehel et al., FracLab, http://fractales.inria.fr/index.php?page=fraclab.
[32] 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.
[33] I. Daubechies, Ten lectures on waveletsCBMS-NSF Series in Applied Mathematics, Capital City Press, Montpelier, Vermont, 1992.
[34] A. Turiel, C. Pérez-Vicente, Multifractal measures: definition, description, synthesis and analysis. A detailed study, in: J.-P. Nadal, A. Turiel, H. Yahia (Eds.), Proceedings of the "Journées détude sur les méthodes pour les signaux complexes en traitement d'image", 2004, pp. 41–57.
[35] B. Mandelbrot, Random multifractals: negative dimensions and the resulting limitations of the thermodynamic formalism, Proc. Roy. Soc. A 434 (1991) 79–88.
[36] R. Benzi, S. Ciliberto, C. Baudet, G.R. Chavarria, C. Tripiccione, Extended self similarity in the dissipation range of fully developed turbulence, Europhys. Lett. 24 (1993) 275–279.
[37] R. Benzi, S. Ciliberto, C. Tripiccione, C. Baudet, F. Massaioli, S. Succi, Extended self-similarity in turbulent flows, Phys. Rev. E 48 (1993) R29–R32.
[38] R. Benzi, S. Ciliberto, C.B.G.R. Chavarria, On the scaling of three dimensional homogeneous and isotropic turbulence, Physica D 80 (1995) 385–398.
[39] A. Nevado, A. Turiel, N. Parga, Scene dependence of the non-Gaussian scaling properties of natural images, Network 11 (2000) 131–152.
[40] Z.S. She, E. Leveque, Universal scaling laws in fully developed turbulence, Phys. Rev. Lett. 72 (1994) 336–339.
[41] B. Castaing, The temperature of turbulent flows, Journal de Physique II 6 (1996) 105–114.
[42] J.F. Muzy, J. Delour, E. Bacry, Modelling fluctuations of financial time series: from cascade process to stochastic volatility model, Euro. Phys. J. B 17 (2000) 537–548.
[43] B. Dubrulle, Intermittency in fully developed turbulence: log-Poisson statistics and generalized scale covariance, Phys. Rev. Lett. 73 (1994) 959–962.
[44] S. Jaffard, Some mathematical results about multifractal for functions, in: C.K. Chui, L. Montefusco, L. Puccio (Eds.), Wavelets: Theory, Algorithms and Applications, Academic Press, San Diego, CA, 1994, pp. 325–361.
[45] W. Rudin, Real and Complex Analysis, McGraw Hill, New York, USA, 1987.
[46] S. Mallat, A Wavelet Tour of Signal Processing, Second ed., Academic Press, New York, 1999.
[47] A. Arneodo, E. Bacry, S. Jaffard, J. Muzy, Singularity spectrum of multifractal functions involving oscillating singularities, J. Fourier Anal. Appl. 4 (2) (1998) 159–174.
[48] S. Jaffard, Multifractal formalism for functions. I. Results valid for all functions, SIAM J. Math. Anal. 28 (4) (1997) 944–970.
[49] T. Carleman, Sur le problème des moments, Comptes rendus Acad. Sci. Paris 174 (1922) 1680.