Fractals, Vol. 8, No. 2 (2000) 163–179

DETERMINING LOCAL SINGULARITY STRENGTHS AND THEIR SPECTRA WITH THE WAVELET TRANSFORM

ZBIGNIEW R. STRUZIK
Centre for Mathematics and Computer Science (CWI),
Kruislaan 413, 1098 SJ Amsterdam, The Netherlands
E-mail: Zbigniew.Struzik@cwi.nl

Received September 2, 1999; Accepted December 28, 1999

Abstract

We present a robust method of estimating the effective strength of singularities (the effective Hölder exponent) locally at an arbitrary resolution. The method is motivated by the multiplicative cascade paradigm, and implemented on the hierarchy of singularities revealed with the wavelet transform modulus maxima (WTMM) tree. In addition, we illustrate the direct estimation of the scaling spectrum of the effective singularity strength, and we link it to the established partition function-based multifractal formalism. We motivate both the local and the global multifractal analysis by showing examples of computer-generated and real-life time series.

1. INTRODUCTION

The WTMM-based multifractal formalism developed by Arneodo et al. in the early 1990s[1,2] has almost reached the status of a standard. It has been extensively used to test many natural phenomena and has contributed to substantial progress in each domain in which it has been applied.[3–6] Nevertheless, the respective partition function-based methodology is intrinsically statistical in nature. It provides only global estimates of scaling (of the moments of relevant quantity). While this is often a required property, there are cases when local information about scaling provides more relevant information than the global spectrum. This is particularly true for time series where scaling properties are non-stationary, whether it be due to intrinsic changes in the signal scaling characteristics or even boundary effects. In addition to this, the weighted singularity selection provided by the moments in the partition function method also leads to the smoothing of the obtained singularity

spectra. This can hide interesting and relevant information in the shape of the spectrum. This information can be accessed when the spectrum of the singular structures is evaluated directly from the histograms of the singularity strength. In a traditional setting, the estimation of both local singularity strengths and their spectra has the problem of being very unstable and prone to gross numerical errors.

However, with the help of the wavelet transform multiscale decomposition, we are able to provide stable procedures for both the local exponent and its global spectrum. In particular, we address the problem of the estimation of the *local* scaling exponent through the paradigm of the multiplicative cascade. We reveal the hierarchy of the scaling branches of the cascade with the WTMM tree, which has proven to be an excellent tool for the purpose.[2,7] Contrary to the intrinsically unstable local slope of the maxima lines, this estimate is robust and provides a stable, effective singularity strength, local in scale and position — the effective Hölder exponent estimate.[8] From this, we make an attempt to derive the multifractal spectra directly from log-histogram scaling evaluation, linking the local analysis with the global multifractal spectra approach. Almost as stable as the global scaling estimates from the partition function method, the direct histogram of the effective singularity strength provides considerably more information about the relative density of local scaling exponents, and it may prove to be an interesting alternative in multifractal spectra estimation.

The structure of the paper is as follows. In Sec. 2, we focus on the relevant aspects of the wavelet transformation (WT), in particular the ability to characterize scale-free behavior through the Hölder exponent. Together with the hierarchical scale-wise decomposition provided by the wavelet transform, this will enable us to reveal the scaling properties of the tree of the multiplicative cascading process. In Sec. 3, we introduce a technical model enabling us to estimate the scale-free characteristic (the effective Hölder exponent) for the branches of such a process. In Sec. 4, we use the derived effective Hölder exponent for the local temporal description of the time series characteristics at a given resolution (scale). This is followed by an analysis of distributions of local h and the (scaling) evolution of the log-histogram and its relation to the standard partition function-based multifractal formalism. We motivate both the local and multifractal analysis by showing examples of generated and real-life time series. Section 5 closes the paper with conclusions and suggestions for future developments.

2. CONTINUOUS WAVELET TRANSFORM AND ITS MAXIMA USED TO REVEAL THE STRUCTURE OF THE TIME SERIES

The recently introduced WT[9,10] provides a way of analyzing the local behavior of functions. In this, it fundamentally differs from global transforms like the Fourier Transformation. In addition to locality, it possesses the often very desirable ability of filtering the polynomial behavior to some pre-defined degree. Therefore, correct characterization of time series is possible, in particular, in the presence of *non-stationarities* like global or local trends or biases. One of the main aspects of the WT which is of great advantage for our purpose is the ability to reveal the *hierarchy* of (singular) features, including the scaling behavior.[2]

Conceptually, the WT is a convolution product of the time series with the scaled and translated kernel — the wavelet $\psi(x)$, usually a nth derivative of a smoothing kernel $\theta(x)$. Usually, in the absence of other criteria, the preferred choice is the kernel which is well-localized both in frequency and position. In this report, we chose the Gaussian $\theta(x) = \exp(-x^2/2)$ as the smoothing kernel, which has optimal localization in both domains.

The scaling and translation actions are performed by two parameters; the scale parameter s "adapts" the width of the wavelet kernel to the *microscopic resolution* required, thus changing its frequency contents, and the location of the analyzing wavelet is determined by the parameter b:

$$Wf(s, b) = \frac{1}{s}\int_{-\infty}^{\infty} dx\, f(x)\psi\left(\frac{x-b}{s}\right) \qquad (1)$$

where $s, b \in \mathbf{R}$ and $s > 0$ for the continuous version (CWT).

In Fig. 1, we show the wavelet transform of a random walk sample decomposed with the Mexican hat wavelet — the second derivative of the Gaussian kernel. From the definition, the transform retains all of the temporal locality properties — the position axis is in the forefront of the 3D plot. The standard way of presenting the CWT is using the logarithmic scale, therefore the scale axis pointing "into the depths" of the plot is log(s). The third

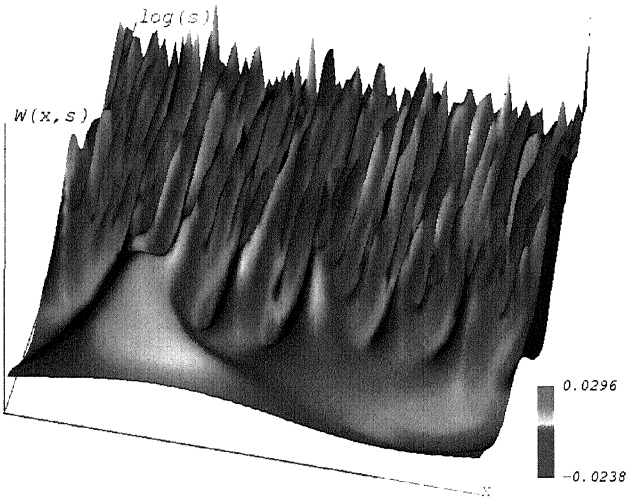


Fig. 1 Continuous Wavelet Transform representation of the random walk (Brownian process) time series. The wavelet used is the Mexican hat — the second derivative of the Gaussian kernel. The coordinate axes are: position x, scale in logarithm $\log(s)$, and the value of the transform $W(s, x)$.

vertical axis denotes the magnitude of the transform $W(s, b)$.

The 3D plot shows how the wavelet transform reveals more and more detail while going towards smaller scales, i.e. towards smaller $\log(s)$ values. Therefore, the wavelet transform is sometimes referred to as the "mathematical microscope",[2] due to its ability to focus on weak transients and singularities in the time series. The wavelet used determines the optics of the microscope; its magnification is varied by the scale factor s.

2.1 Assessing Singular Behavior with the Wavelet Transformation

Quite frequently it is the singularities, the rapid changes, discontinuities and frequency transients, and not the smooth, regular behavior which are interesting in the time series. Let us, therefore, demonstrate the wavelet's excellent suitability to address singular aspects of the analyzed time series in a *local* fashion. The singularity strength is often characterized by the Hölder exponent.

If there exists a polynomial P_n of degree $n < h$, such that

$$|f(x) - P_n(x - x_0)| \leq C|x - x_0|^h, \qquad (2)$$

the supremum of all h, such that the above relation holds, is termed the Hölder exponent $h(x_0) \in (n, n+1)$ of the singularity at x_0. P_n can often be associated with the Taylor expansion of f around x_0, but Eq. 2 is valid even if such an expansion does not exist.[13] The Hölder exponent is therefore a function defined for each point of f, and it describes the local regularity of the function (or distribution) f.

Let us take the wavelet transform $W^{(n)}f$ of the function f in $x = x_0$ with the wavelet of at least n vanishing moments, i.e. orthogonal to polynomials up to degree n:

$$\int_{-\infty}^{+\infty} x^m \psi(x) dx = 0 \quad \forall m,\, 0 \leq m < n\,.$$

For the sake of illustration, let us assume that the function f can be characterized by Hölder exponent $h(x_0)$ in x_0, and f can be locally described as:

$$f(x)_{x_0} = c_0 + c_1(x - x_0) + \cdots + c_n(x - x_0)^n + C|x - x_0|^{h(x_0)}\,.$$

Its wavelet transform $W^{(n)}f$ with the wavelet with at least n vanishing moments now becomes:

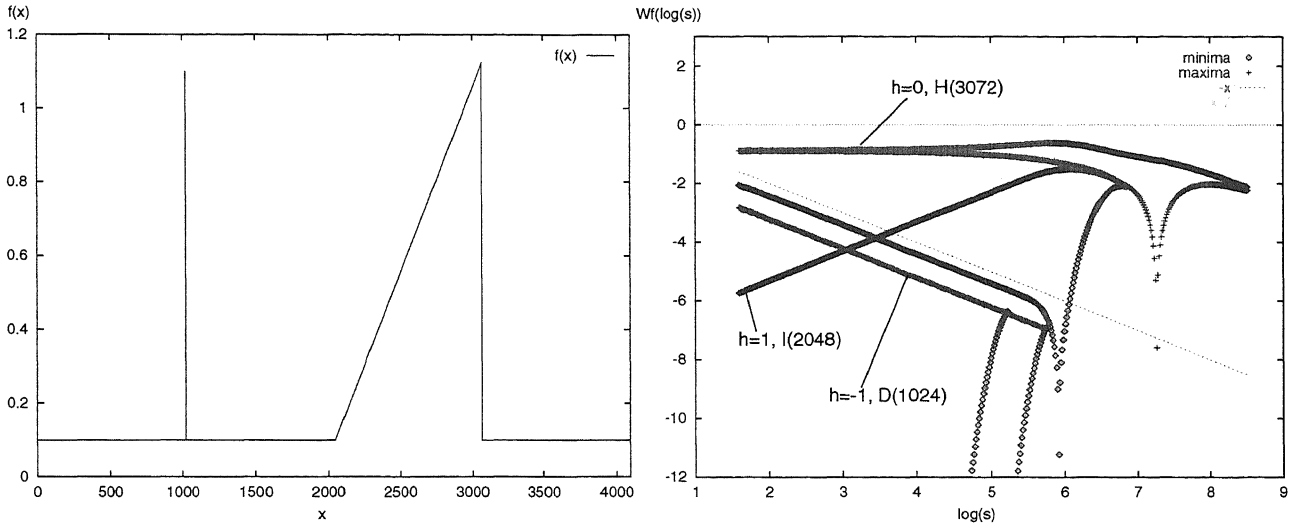
$$W^{(n)}f(s, x_0) = \frac{1}{s}\int C|x - x_0|^{h(x_0)}\psi\left(\frac{x - x_0}{s}\right)dx$$
$$= C|s|^{h(x_0)}\int |x'|^{h(x_0)}\psi(x')dx'\,.$$

Therefore, we have the following power law proportionality for the wavelet transform of the (Hölder) singularity of $f(x_0)$:

$$W^{(n)}f(s, x_0) \sim |s|^{h(x_0)}\,.$$

Note: One should bear in mind that the above relation is an approximate case for which exact theorems exist.[11] In particular, we will restrict the scope of this paper to cusp singularities for which the local and pointwise Hölder exponents are equal.[12] Thus we will not take into consideration the chirp or "oscillating" singularities [e.g. $x^\alpha \sin(1/x^\beta)$] requiring two exponents.[13,14] Nevertheless, it is sufficient for our purpose to state that the continuous wavelet transform can be used for characterizing the cusp singularities in the time series even if masked by the polynomial bias.

It can be shown[15] that for cusp singularities, the location of the singularity can be detected, and the related exponent can be recovered from the scaling of the Wavelet Transform, along the so-called *maxima line*, converging towards the singularity. This



(a) (b)

Fig. 2 **(a)** The test signal consisting of the Dirac pulse $D(1024)$, the change in slope — integrated Heaviside step $I(2048)$, and the Heaviside step $H(3072)$. **(b)** The log-log plot of the maxima, together with their respective logarithmic derivatives, corresponding to all three singularities: $D(1024)$, $I(2048)$ and $H(3072)$. Lines of theoretical slope are also indicated; these are $-x$ for $D(1024)$, x for $I(2048)$ and a constant for $H(3072)$. The wavelet used is the Mexican hat.

is a line where the wavelet transform reaches local maximum (with respect to the position coordinate). Connecting such local maxima within the continuous wavelet transform "landscape" gives rise to the entire tree of maxima lines. As shown in the following subsection, it appears that restricting oneself to the collection of such maxima lines provides a particularly useful representation of the entire CWT.

Let us consider the following set of examples of simple singular structures, see Fig. 2(a); a single Dirac pulse at $D(1024)$, the saw tooth consisting of an integrated Heaviside step function at $I(2048)$, and the Heaviside step function for $S(3072^+)$, where $+$ denotes the right-hand limit. The Hölder exponent of a Dirac pulse is -1 by definition. For cusp singularities, the process of integration and differentiation respectively adds and subtracts one from the exponent. We, therefore, have $h = 0$ for the right-sided step function $S(3072^+)$ and $h = 1$ for the integrated step $I(2048)$.

These values can also be verified in the scaling of the corresponding maxima lines. We obtain the (logarithmic) slopes of the maxima values very closely following the correct values of these exponents [see Fig. 2(b)]. This, of course, suggests the possibility of the estimation of the Hölder exponent of cusp singularities from the slope of the maxima lines approaching these singularities. An important limitation is, however, the requirement for the singularities to be *isolated* for this procedure to work. Note that the scaling of the maxima lines becomes stable in the log-log plot in Fig. 2(b), only below some critical scale s_{crit}, below which the singularities effectively become isolated for the analyzing wavelet. Indeed, the distance between the singular features in the test time series in Fig. 2(a) equals 1024, which is in the order of three standard deviations of the analyzing wavelet at $\log(s_{\text{crit}}) = 5.83 = \log(1024/3)$. This example largely simplifies the issue since the singular structures are of the same size, resulting in one characteristic scale at which they appear in the wavelet transform. Also, generally, the scaling of the maxima lines for other than the presented simple examples will not follow a straight line even for isolated singularities. Still, the rate of decrease of (the supremum of) the related wavelet transform maximum will be consistent, thus allowing the estimation of h.

2.2 Wavelet Transform Modulus Maxima Representation

The continuous wavelet transform described in Eq. 1 is an extremely redundant representation,

much too costly for most practical applications. This is the reason why other, less redundant representations, are frequently used. Of course, in going from high to low redundancy (or even orthogonality), certain (additional) design criteria are necessary. For our purpose of analysis of the local features of time series, one critical requirement is the translation shift invariance of the representation; nothing other than the boundary coefficients of the representation should change, if the time series is translated by some Δx.

A useful representation satisfying this requirement and of much less redundancy than the CWT is the representation making use of the local maxima of the WT as suggested in the previous section. Such maxima interconnected along scales form the so-called Wavelet Transform Modulus Maxima (WTMM) representation, see Fig. 3, first introduced by Mallat.[16] In addition to translation invariance, the WTMM also possesses the ability to characterize fully the local singular behavior of time series, as illustrated in the previous subsection.

Moreover, the wavelet transform and its WTMM representation can also be shown to be invariant with respect to the rescaling/renormalization operation.[2,7,14,17] This property makes it an ideal tool for revealing the renormalization structure of the (hypothetical) multiplicative process underlying the analyzed time series.

Suppose we have the time series f invariant with respect to some renormalization operation $\mathcal{R}$:

$$f = \mathcal{R}f\,.$$

The wavelet transform of f will, for a certain class of $\mathcal{R}$, in particular for multiplicative cascades, show the invariance with respect to an operator $\mathcal{T} \leftrightarrow \mathcal{R}^{-1}$. This can be recovered from the invariance of the wavelet transform of F:

$$W(f) = \mathcal{T}W(f)$$

and in particular from the invariance of (the hierarchy of) the WTMM tree.[7,17]

2.3 Multifractal Formalism on the WTMM Tree

The WTMM tree has been used for defining the partition function-based multifractal formalism.[2] It uses the moments q of the measure distributed on the WTMM tree to obtain the dependence of the scaling function $\tau(q)$ on the moments q:

$$\mathcal{Z}(s,\, q) \sim s^{\tau(q)}\,.$$

The $\mathcal{Z}(s,\, q)$ is the partition function of the qth moment of the measure distributed over the wavelet transform maxima at the scale s considered:

$$\mathcal{Z}(s,\, q) = \sum_{\Omega(s)} (Wf\omega_i(s))^q \tag{3}$$

where $\Omega(s) = \{\omega_i(s)\}$ is the set of all maxima $\omega_i(s)$ at the scale s, satisfying the constraint on their local logarithmic derivative in scale.[18] (The local slope bound used throughout this paper is $|\check{h}| \leq 2$.)

Intuitively, since the moment q has the ability to select a desired range of values: small for $q < 0$, or large for $q > 0$, the scaling function $\tau(q)$ globally captures the distribution of the exponents $h(x)$ — weak exponents are addressed with large negative q, while strong exponents are suppressed and effectively filtered out. For the large positive q, the opposite takes place (and strong exponents are addressed while weak exponents are effectively filtered out).

This dependence may be linear, indicating that there is only one class of singular structures and related exponent, or it can have a slope non-linearly changing with q. In the latter case, the local tangent slope to $\tau(q_*)$ will give the corresponding exponent, i.e. $h(q_*)$, with its related dimension marked on the ordinate axis $C = D(h(q_*))$, where $\tau(q_*) = h(q_*)q_* + C$. The set of values C, i.e. dimensions $D(h(q_*))$ for each value of h selected with q_* is the so-called spectrum of the singularities $D(h)$ of the fractal signal. Formally, the transformation from $\tau(q)$ to $D(h)$ is referred to as the Legendre transformation:

$$\frac{d\tau(q)}{dq} = h(q)$$

$$D((h(q)) = qh(q) - \tau(q)\,.$$

Note that even though the method uses the maxima tree containing full local information about the singularities, this is lost at the very moment the partition function is computed. Therefore, there is no explicit local information present in the scaling estimates; τ, h or D, and all these are global statistical estimates. This is also where the strength of the partition function method lies — global averages are much more stable than local information and in some cases all that it is possible to obtain.

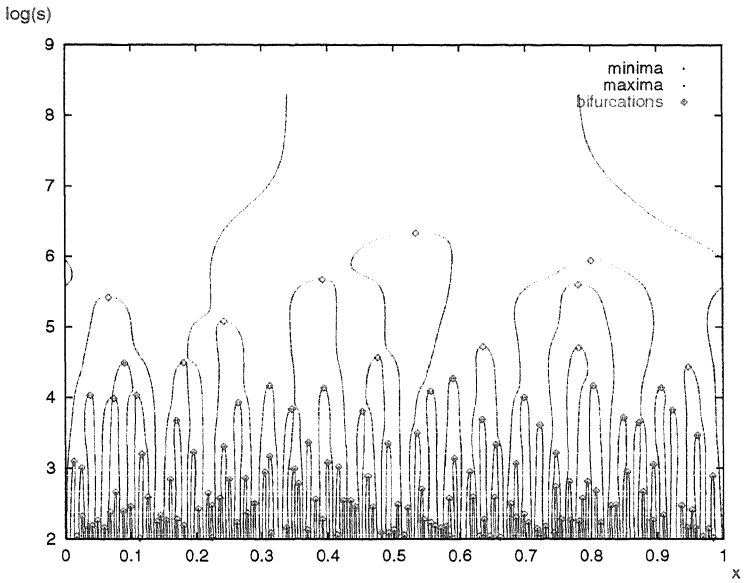
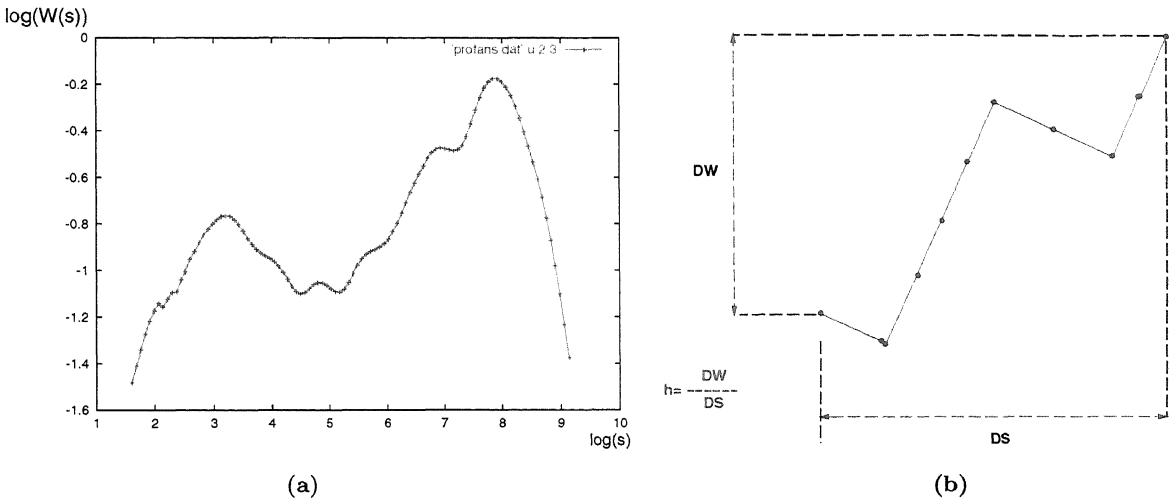


Fig. 3 WTMM representation of the time series and the bifurcations of the WTMM tree. Mexican hat wavelet.



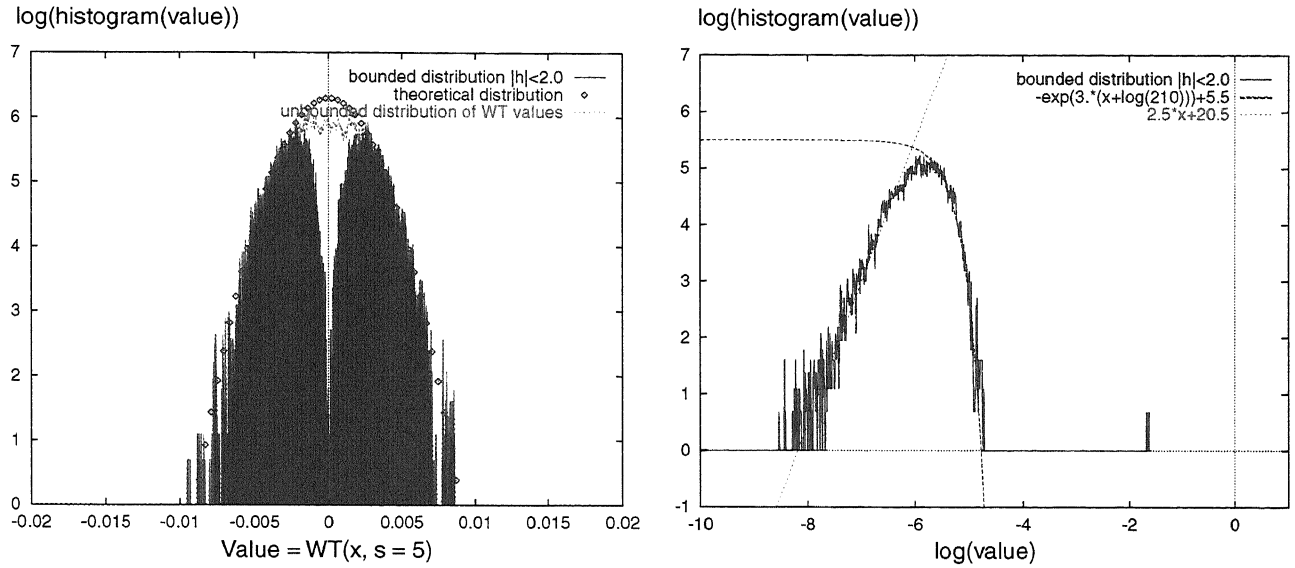
(a) (b)

Fig. 4 (a) It is impossible to evaluate the scaling exponent for an arbitrary maximum line participating in a complex process: a real-life example of a maximum line. (b) The local effective Hölder exponent estimate takes the effective difference in the logarithm of the density of the process with respect to the logarithm of the scale difference gained along the process path.

Indeed, it is generally not possible to obtain local estimates of the scaling behavior other than in the case of isolated singular structures from the WT. A typical example of the evolution of the maximum line along scale is shown in Fig. 4(a). It is not possible to evaluate the slope of the plot, not even on the selected range of scales. This is why we introduced[8] an approach circumventing this problem while retaining local information — a local effective Hölder exponent in which we model the singularities as created in some kind of a collective process of a very generic class.

3. ESTIMATION OF THE LOCAL, EFFECTIVE HÖLDER EXPONENT USING THE MULTIPLICATIVE CASCADE MODEL

We have shown in the previous section that the wavelet transform, and in particular its maxima



(a) (b)

Fig. 5 **(a)** Distribution of WT values of fBm of $H = 0.6$ with and without local slope bounding in log-normal coordinates. **(b)** Bounded distribution in log-log scale. Bound $|\breve{h}| < 2$, scale $s = 5$. Sample length 16386.

lines, can be used in evaluating the Hölder exponent in isolated singularities. In most real-life situations, however, the singularities in the time series are not isolated but densely packed. The logarithmic rate of increase or decay of the corresponding wavelet transform maximum line is usually not stable but fluctuates wildly, in addition often making estimation impossible due to divergence problems when the value of the WT along the maximum line approaches zero.

As a remedy for the problem of diverging maxima values in log-log plots, we used the procedure of bounding the local Hölder exponent[18] to preprocess the maxima. The crux of the method lies in the explicit calculation of the bounds for the (positive and negative) slope locally in scale. The parts of the maxima lines for which the slope exceeds the bounds imposed are simply not considered in the calculations.[18] For example, compare in Fig. 2 the example of log-log slopes above the critical scale s_{crit}, where the singularities can no longer be considered as isolated. In particular, note that the local slope near scale $\log(s) = 6$ and $\log(s) = 7$ reaches $\pm\infty$. Such diverging slopes are thresholded and removed by applying the bounding procedure. In this example and throughout this paper, we use $|\breve{h}| < 2$ bound on the local slope $\breve{h}$ of each maximum. The output of this procedure is, therefore, the set of non-diverging values of the maxima lines corresponding to the singularities in the time series.

In Fig. 5 above, we show the effect of the procedure on the distribution of the maxima values for a fixed scale $s = 5$, for a fractional Brownian motion (fBm) record with $H = 0.6$. The unbounded distribution has a Gaussian shape as expected, which shows as a parabola in the logarithmic plot in Fig. 5(a). Bounding local slopes to $|\breve{h}| < 2$ results in a rapid decay of small values of maxima towards the limit of 0 value, see the filled histogram, thus making negative moments well-defined.[a] In Fig. 5(b), we verify in log-log coordinates that the decay of the small values follows a power law.

Even with the diverging parts of the maxima removed, generally it is still not possible to obtain local estimates of the scaling behavior other than in the case of isolated singular structures. Typically, see Fig. 4(a), the WT maxima values strongly fluctuate, with the local slope changing from point to point. While it is not possible to evaluate the slope of the plot, not even on the selected range of scales, the fluctuations actually carry information

[a]Note that removing the diverging maxima parts is only necessary for the partition function method, not for the effective Hölder exponent and the related direct spectra. Since we want to compare the two, we pre-process the data in the same way.

very relevant to the scaling properties of the process underlying the origin of the input time series.[7] As mentioned above, the WT has the ability to capture the renormalization process (hypothetically) pertinent to the time series analyzed. The behavior of the maxima lines closely follows both the hypothetical production (multiplicative) rule and its branching process. Based on this, a local effective Hölder exponent approach has been suggested[8] to provide a model-based approximation of the local scaling exponent, i.e. singularity strength. The model used is exactly of the type which can be revealed by the WT — the multiplicative cascade process, a straightforward generalization of the binomial multiplicative process.

3.1 Multiplicative Cascade Model

Let us take the well-known example of the Besicovitch measure on the Cantor set.[19] The set of transformations B_i, $i \in \{1, 2\}$ describing the Besicovitch construction can be expressed as:

$$B_i f(x) = p_i f\left(\frac{x + b_i}{c_i}\right)$$

with the normalization requirement:

$$p_1 + p_2 = 1\,. \tag{4}$$

Additionally, we put conditions ensuring non-overlapping of the transformations:

$$\frac{1 + b_1}{c_1} < \frac{0 + b_2}{c_2}$$

while all the respective values b_1/c_1, b_2/c_2, c_1^{-1}, c_2^{-1} are from the interval $(0, 1)$.

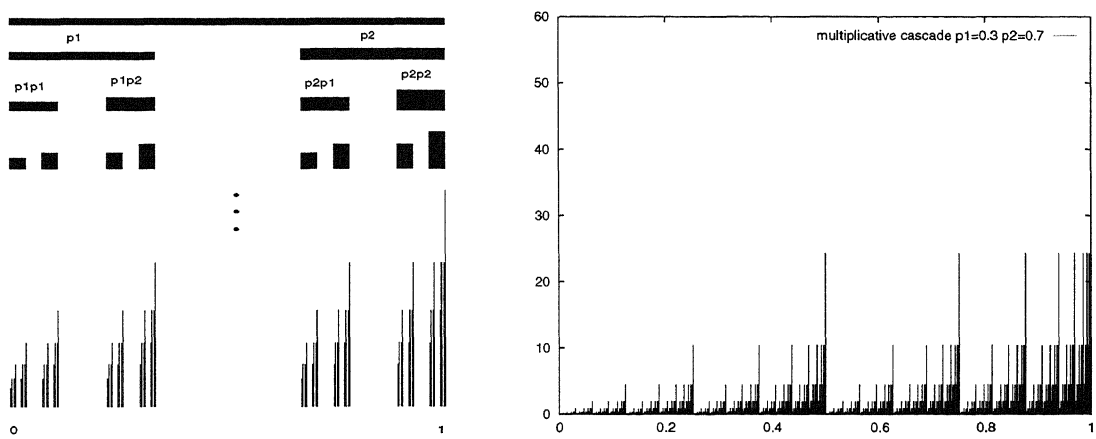
For equal ratios, $p_1 = p_2 = 1/2$ and $c_1 = c_2 = 3$ with $b_1 = 0$ and $b_2 = 2$, we recover the middle-third, homogeneous distribution of the measure on the Cantor set. We have the non-homogeneous, multifractal Besicovitch measure for non-equal p_i, see Fig. 6(a). Finally, for non-equal p_i, regardless of the normalization Eq. 4 and with $c_1 = c_2 = 2$ with $b_1 = 0$ and $b_2 = 1$, we have the multiplicative cascade on (0.1) interval, see Fig. 6(b).

Each point of this cascade is uniquely characterized by the sequence of weights $(s_1, \ldots, s_n)$ taking values from the (binary) set $\{1, 2\}$, and acting successively along a unique process branch leading to this point. Suppose that we denote the density of the cascade at the generation level F_i (i running from 0 to max) by $\kappa(F_i)$, we then have

$$\kappa(F_{\max}) = p_{s_1}, \ldots, p_{s_n}\kappa(F_0) = P_{F_0}^{F_{\max}}\kappa(F_0)$$

and the local exponent is related to the product $P_{F_0}^{F_{\max}}$ of these weights:

$$h_{F_{\max}}^{F_0} = \frac{\log(P_{F_0}^{F_{\max}})}{\log((1/2)^{\max}) - \log((1/2)^0)}\,.$$



(a) (b)

Fig. 6 **(a)** The Besicovitch measure on the Cantor set, generations F_0 through F_3 and the generation F_6. The distribution of weights is $p_1 = 0.4$ and $p_2 = 0.6$. The standard middle third Cantor division is retained. **(b)** Similar construction but on 0.1 support instead of the Cantor set, leading to multiplicative cascade. $p_1 = 0.3$ and $p_2 = 0.7$, generation F_{13}.

In any experimental situation, the weights p_i are not known and h_i has to be estimated. This can be simply done using the fact that for the multiplicative cascade process of the kind just described, the effective product of the weighting factors is reflected in the difference of logarithmic values of the densities at F_0 and $F_{\max}$ along the process branch:

$$h_{F_{\max}}^{F_0} = \frac{\log(\kappa(F_{\max})) - \log(\kappa(F_0))}{\log((1/2)^{\max}) - \log((1/2)^0)} .$$

The densities along the process branch can be estimated with the wavelet transform, using its remarkable ability to reveal the entire process tree of a multiplicative process.[7] It can be shown that the densities $\kappa(F_i)$ can be estimated from the value of the wavelet transform along the maxima lines corresponding to the given process branch. The estimate of the effective Hölder exponent becomes:

$$\hat{h}_{s_{lo}}^{s_{hi}} = \frac{\log(Wf\omega_{pb}(s_{lo})) - \log(Wf\omega_{pb}(s_{hi}))}{\log(s_{lo}) - \log(s_{hi})}$$

where $Wf\omega_{pb}(s)$ is the value of the wavelet transform at the scale s, along the maximum line ω_{pb} corresponding to the given process branch. Scale s_{lo} corresponds with generation $F_{\max}$, while s_{hi} corresponds with generation F_0, see Fig. 4(b).

For the estimation of h, we need s_{hi} and $Wf\omega_{pb}(s_{hi})$. We can, of course, pick any of the roots of the sub-trees of the entire maxima tree in order to evaluate exponents of the partial process or sub-cascade. But for the entire sample available we must use the entire tree and for this purpose, we can only do as well as taking the sample length to correspond with s_{hi}, i.e.:

$$s_{hi} \equiv s_{SL} = \log(\text{SampleLength}) .$$

Unfortunately, the wavelet transform coefficients at this scale are heavily distorted by finite size effects. This is why we estimate the value of $Wf\omega_{pb}(s_{hi})$ using the mean h exponent (Fig. 7).

3.2 Estimation of the Mean Hölder Exponent

For a multiplicative cascade process, a mean value of the cascade at the scale s can be defined as:

$$\mathcal{M}(s) = \frac{\mathcal{Z}(s, 1)}{\mathcal{Z}(s, 0)} \tag{5}$$

where the $\mathcal{Z}(s, q)$ is the partition function, Eq. 3 of

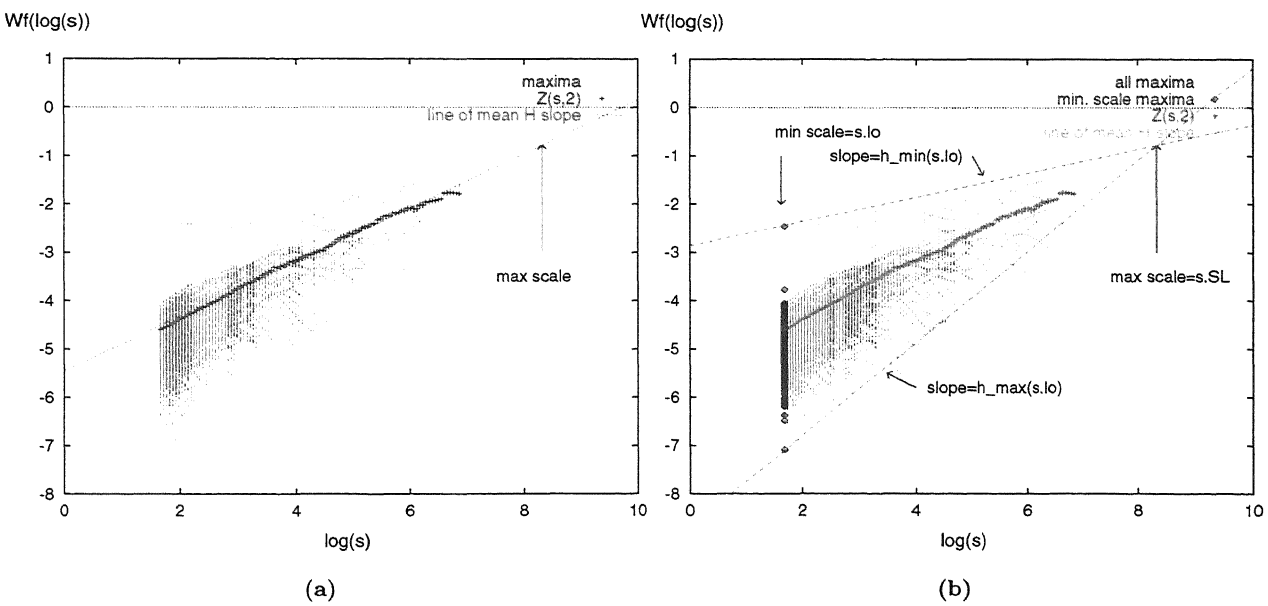


Fig. 7 (a) The projection of the maxima lines of the WT along time. The mean value of the Hölder exponent can be estimated from the log-log slope of the line shown. Also, the beginning of the cascade at the maximum scale s_{hi} is indicated. (b) The maxima at the smallest scale considered are shown in the projection along time. The effective Hölder exponent can be evaluated for each point of the maximum line at s_{lo} scale. Two extremal exponent values are indicated, for minimum and maximum slope.

the qth moment of the measure distributed over the wavelet transform maxima at the scale s considered. This mean gives the direct possibility of estimating the mean value of the local Hölder exponent as a linear fit to $\mathcal{M}$:

$$\log(\mathcal{M}(s)) = \bar{h} \log s + C\,. \tag{6}$$

We will not, however, use Definition 5 since we want the Hölder exponent to be the local version of the Hurst exponent. This compatibility is easily achieved when we take the second moment in the partition function to define the mean $\bar{h}'$:

$$\mathcal{M}'(s) = \sqrt{\frac{\mathcal{Z}(s,\,2)}{\mathcal{Z}(s,\,0)}}\,.$$

Therefore, we estimate our mean Hölder exponent $\bar{h}'$ from Eq. 6 substituting $\mathcal{M}$ with $\mathcal{M}'$. The estimate of the local Hölder exponent, from now on to be denoted $\hat{h}(x_0,\,s)$ or just $\hat{h}$, now becomes:

$$\hat{h}^{s_{SL}}_{s_{lo}} \cong \frac{\log(\,Wf(s_{lo})) - (\bar{h}' \log s + C)}{\log(s_{lo}) - \log(s_{SL})}\,.$$

4. EMPLOYING THE EFFECTIVE HÖLDER EXPONENT IN LOCAL AND GLOBAL SPECTRA ESTIMATION

An estimated local $\hat{h}(x_0,\,s)$ can be depicted in the temporal fashion, for example with a background color, as we have done in Fig. 8. The first example time series is a computer-generated sample of fractional Brownian motion with $H = 0.6$. It shows almost monochromatic behavior; centered at $H = 0.6$ the color green is dominant. There are, however, several instances of darker green and light blue, indicating locally smooth components.

The second example time series is a record of the S&P 500 index from the time period [1984–1988]. There are significant fluctuations in color in this picture, with the green color centered at $H = 0.5$, indicating both smoother and rougher components. In particular, one can observe an extremal red value at the '87 crash coordinate, followed by very rough behavior (a rather obvious fact, but to the best of our knowledge not reported to date in the rapidly growing coverage of this time series record).[20]

The third example is a real-life biological time series and comes from aphids. This is the temporal record of electrical resistance, a "penetratiogram", reflecting the penetration of the tongue of the aphid through the plant cell wall. We attempt to characterize the different regions of the time series, visible as a number of hierarchical "pits" of certain depths within the signal. With green focused at the mean Hurst exponent $H = 0.5$, the result quite convincingly shows the patchy difference in characterization of the pits at (two) different levels of pit hierarchy. Stripes of a different color spectrum indicate a high level of non-stationarity of $\hat{h}$ distribution. Note that the obvious amplitude difference does not influence the color in the plot due to the fact that constant offset is filtered out by the wavelet used — the color is due to a genuine difference in the local scaling exponent.

The last example shows a record of heartbeat intervals recorded from a healthy human heart. Contrary to the two previous examples which show a high degree of localization (or non-stationarity) of the exponent strength, this plot shows an intricate structure of interwoven singularities at various strengths. This behavior has been recently reported[6] to correspond with the multifractal behavior of the heartbeat. The green is centered at $\hat{h} = 0.1$.

Note that these examples are only meant for illustration purposes. A detailed discussion of the implications of the local $\hat{h}$ analysis applied will appear elsewhere.

4.1 Log-Histograms of the Effective Hölder Exponent

Let us for now, instead of selecting one $\hat{h}_\varepsilon(s)$ value band across scales and analyzing its scaling, group the estimated local scale-wise $\hat{h}(x_0,\,s)$ into histograms for each scale value, using bin size ε centered at h.

We will analyze histograms of $\hat{h}$, taking the logarithm of the measure in each histogram bin. This conserves the monotonicity of the original histogram, but allows us to compare the log-histograms with the spectrum of singularities $D(h)$. There is a direct correspondence between our log-histograms and the $D(h)$ through the scaling of the measure $\mu_\varepsilon(s)$ in the bin of size ε of the histogram. Estimation of the rate of growth of this measure, would in fact be an identical procedure to the evaluation of the scaling estimate for each ε wide band of $\hat{h}$, as discussed in the previous subsection.

There is, however, a substantial amount of information in the log-histogram at a particular scale

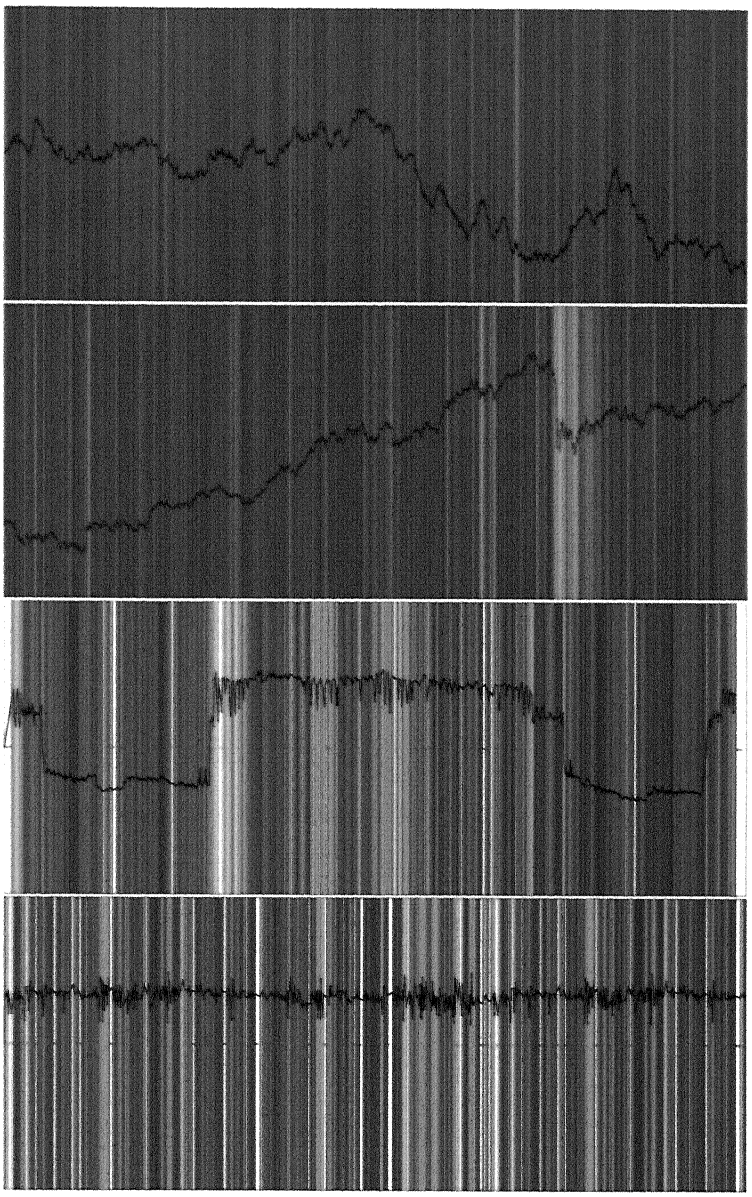


Fig. 8 Example time series with local Hurst exponent indicated in color. From top to bottom: fBm with $H = 0.6$, a record of S&P 500 index, penetratiogram of aphids, and the last is the record of healthy heart interbeat intervals. The background color indicates the Hölder exponent locally, centered at the Hurst exponent at green, color goes towards blue for higher $\hat{h}$ and towards red for lower $\hat{h}$. Local slope bounds for all the plots $|\hat{h}| \leq 2$.

which can be analyzed without performing the scaling analysis. The log-histogram shows the relative probability density of the Hölder exponent per scale. Assuming that the scaling of log-histogram is linear in log-log scale, the shape of the log-histogram remains invariant across scales and converges towards the shape of the $D(h)$, except for "normalization" of the maximum of $P(\hat{h})$, which corresponds with the scaling of the 0th moment in the partition function method. (Note that the scaling assumption can be verified both by analyzing the scaling of the moments and by the scaling of the $\hat{h}$ bin contents which will be done in the following subsection.) This is why in Fig. 9, we show the non-normalized histograms for different scale values, and compare them to the $D_m(h)$

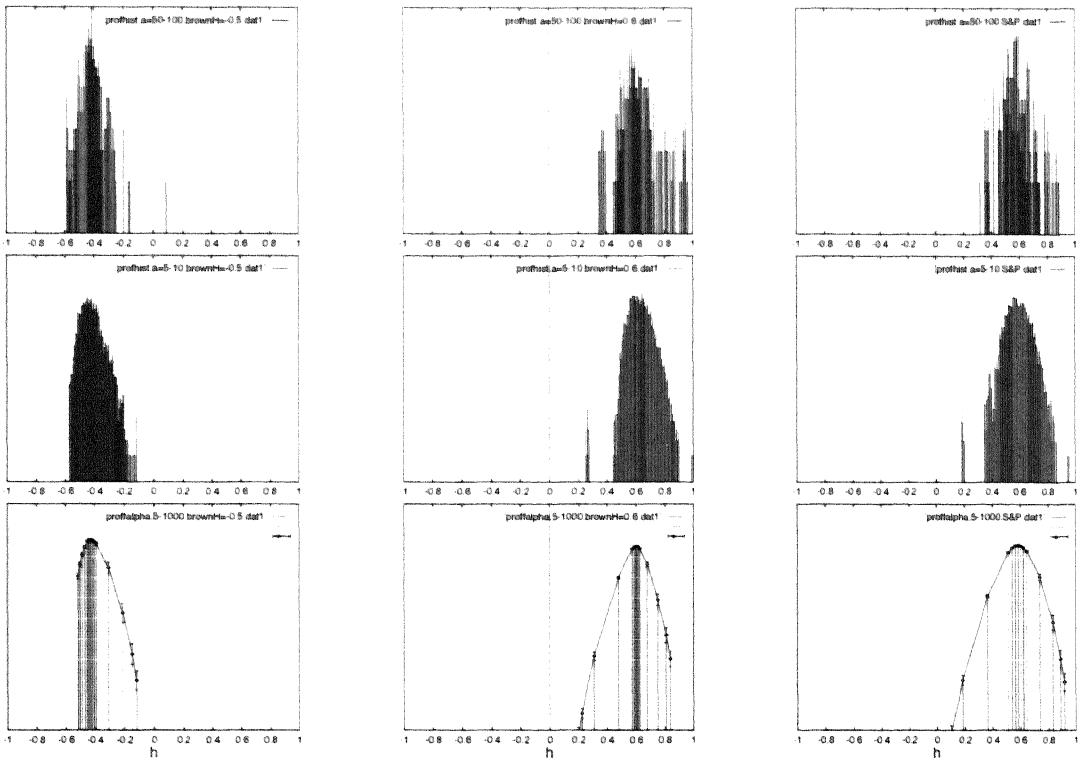


Fig. 9 Two sets of $\hat{h}$ histograms for respective scales $\log(s) \simeq 3.9$, $\log(s) \simeq 1.6$, in top first and center row. Below, in the bottom row, the $D_m(h)$ spectrum obtained with the partition function method (moments $-5 \leq q \leq 5$). Left column: for 4096 samples of white noise. Center: 4096 samples of fractional Brownian motion with $H = 0.6$. Right: S&P 500 index, first 4096 samples from Fig. 8. Local slope bounds for all the plots $|\hat{h}| \leq 2$.

spectrum obtained with the partition function method.

For three example time series, we show in Fig. 9 log-histograms of the exponent $\hat{h}$ at different scales. The time series considered are a white noise sample, a fractional Brownian motion with $H = 0.6$, and a record of the S&P index. Starting at the top, the row of histograms is made for the scale $\log(s) \simeq 3.9$, followed by histograms for $\log(s) \simeq 1.6$. The upper histograms show considerable fragmentation. Several modes become visible, and for all the scales above this scale, the fluctuations will dominate the distribution and consistent statistical behavior will become dispersed. On the contrary, while descending in scale, the bulk of consistent behavior converges to one distribution. The consistent statistical behavior is also captured in the $D_m(h)$ spectrum obtained with the partition function method shown in the bottom row of Fig. 9.

Several aspects of the $D_m(h)$ versus $P(\hat{h})$ already discussed by others[22–24] are visible in the plots. The $P(\hat{h})$ evidently contains more information than $D_m(h)$, in particular, $D_m(h)$ is a convex hull over the $P(\hat{h})$. This is particularly visible in the second (and third) sample, where a pair of extremal $\hat{h}$ values, disconnected from the distribution, corrupts the $D_m(h)$ spectrum. It can be verified through analyzing the corresponding maxima that these values are the end of the sample artifacts and thus do not belong to the distribution. Such maxima can, of course, be removed prior to $D_m(h)$ evaluation, but we kept them for the purpose of illustration: the $P(\hat{h})$ evidently shows that these values disconnect from the bulk of the distribution, while $D_m(h)$ is inherently unable to do so.

4.2 Scale-wise Evolution of the Effective Hölder Exponent

In addition to one scale plot showing the color spectrum of singular behavior, we can also see the scale position locations where the effective Hölder

exponent is near a particular value. We show an example *band* of $\hat{h}_\varepsilon(s)$ of width $\varepsilon = 0.02$, by selecting $\hat{h} = -0.5 \pm 0.01$ in Fig. 10(a) for the record of white noise. The number of locations that fall within the band range visibly grows with scale and this growth determines the dimension $D(h)$ which can be associated with the particular $\hat{h}$, at the band resolution ε.

Such $D(h)$ can be estimated for the entire range of h, resulting in the so-called *spectrum of singularities*. It is a standard way of visualizing the distribution of singularities — it gives the (fractal) dimension $D(h)$ of the supporting set of singularities for each exponent value h in the time series.

$$D(\hat{h}) = \dim(\{x_0\} : Tf(x - x_0) \sim |x - x_0|^{h(x_0)})$$

$$\sim \lim_{\varepsilon \to 0} \lim_{s_{lo} \to 0} \frac{\log(\mu_\varepsilon(\hat{h}(s_{lo})))}{\log(s_{lo})}$$

where μ_ε is the measure of the total number of locations (selected maxima) that fall within the band of size ε at a particular scale location s_{lo}.

One additional modification to the scheme is the linear correction on the ε band width. It is meant to compensate for the decreasing scale range when shorter scale spans are considered. The ε_{mod} used for the calculation of the spectra is thus $\varepsilon_{\text{mod}} \log(s_{lo}/s_{hi})$.

The dimension $D(\hat{h})$ is evaluated in the standard way from the scaling of the log-log plot, as in Fig. 10(b). Obviously the width of the band of the exponents is a parameter of choice, but this does not affect the slope of the log-log line within the realistic range of the ε. Three examples of log-log plots for three values of $\varepsilon = 0.005$, $\varepsilon = 0.01$ and $\varepsilon = 0.025$ are shown in Fig. 10(b). The slope of the log-log plot remains practically independent of the band width, this is especially true towards the small scale limit. The vertical shift in log-log plots is of course due to the decrease/increase of the maxima points within the selected band with its width decreased/increased. Similarly the variance of the log-log plot visibly increases with the decreasing band width.

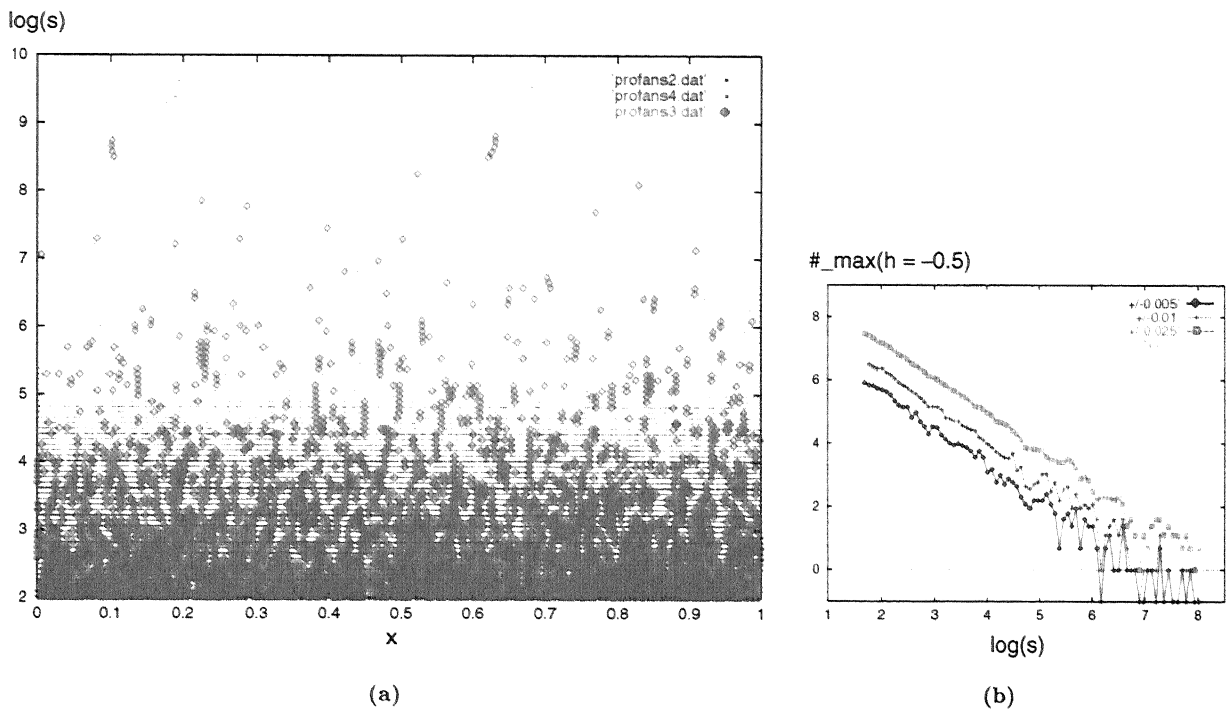


Fig. 10 (a) WTMM representation of a sample of white noise. The maxima are highlighted where the effective Hölder exponent reaches a particular value of $\hat{h} = -0.5 \pm 0.01$, i.e. $\varepsilon = 0.02$. (b) For three values of the band width $\varepsilon = 0.01$, $\varepsilon = 0.02$, $\varepsilon = 0.05$, the logarithm of the sum of the highlighted maxima is shown for each scale with respect to the $\log(s)$ axis. Consistent scaling of -1 rate is shown for *supremum* of $\varepsilon = 0.02$ plot.

4.3 Direct Spectra from the (Bands of) the Effective Hölder Exponent on the WTMM Tree

Calculation of direct spectra from the ε bands of the Hölder exponent simply accounts for covering the entire range of the local effective Hölder exponents detected on the maxima tree. In Fig. 11, we show the entire $D(\hat{h})$ spectra evaluated for the white noise sample of 16 k length and for the record of healthy human heartbeat intervals of equal length. We used the band width of $\varepsilon = \pm 0.01$. The width of the spectrum of white noise is non-zero, as is inevitable for the finite length sample; still the heartbeat sample clearly shows a considerably wider spectrum, confirming recently reported findings.[6]

Due to the fact that it relies on selecting a very narrow band of exponents, this procedure is, however, inherently sensitive to the choice of parameters such as the band width and the density of sampling of the scale axis. While the latter equally affects the partition function method, it does not pose any serious limitation since it can be increased at will,

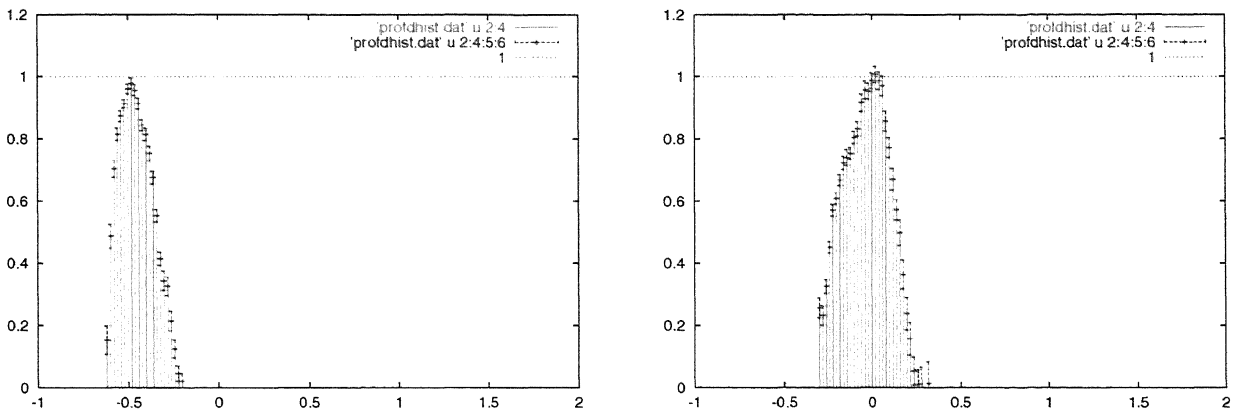


Fig. 11 The multifractal spectrum calculated directly from the scaling of the histogram of the Hölder exponent on the maxima tree. **(a)** For the white noise input signal. **(b)** For the heartbeat intervals record. Band width 0.01.

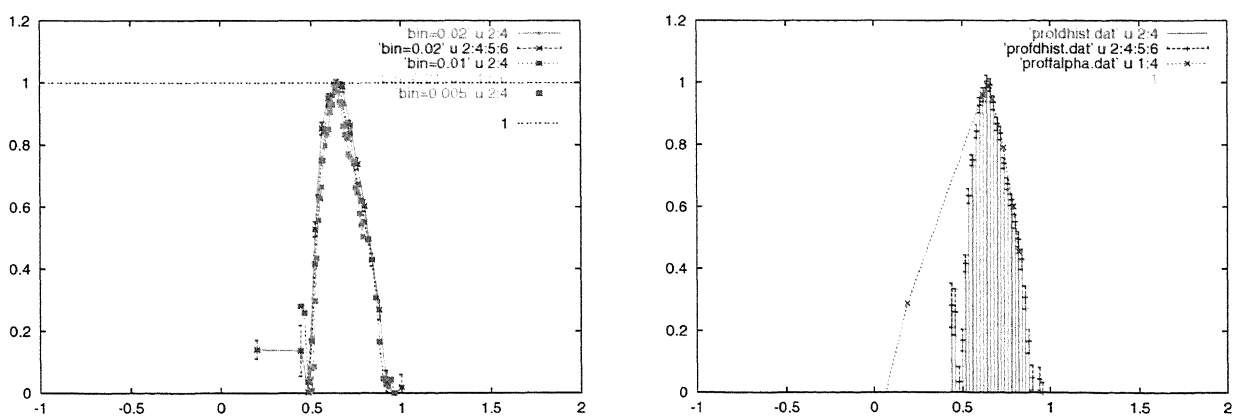


Fig. 12 **(a)** The direct multifractal spectrum shows reasonable stability with respect to varied band width ε. Shown are overlapped cases $\varepsilon = 0.02$, $\varepsilon = 0.01$ and $\varepsilon = 0.005$. **(b)** The direct multifractal spectrum of the fBm sample of $H = 0.6$ is not corrupted by the "outliers" — the end of the sample singularities, as is the standard partition function method spectrum included for comparison.

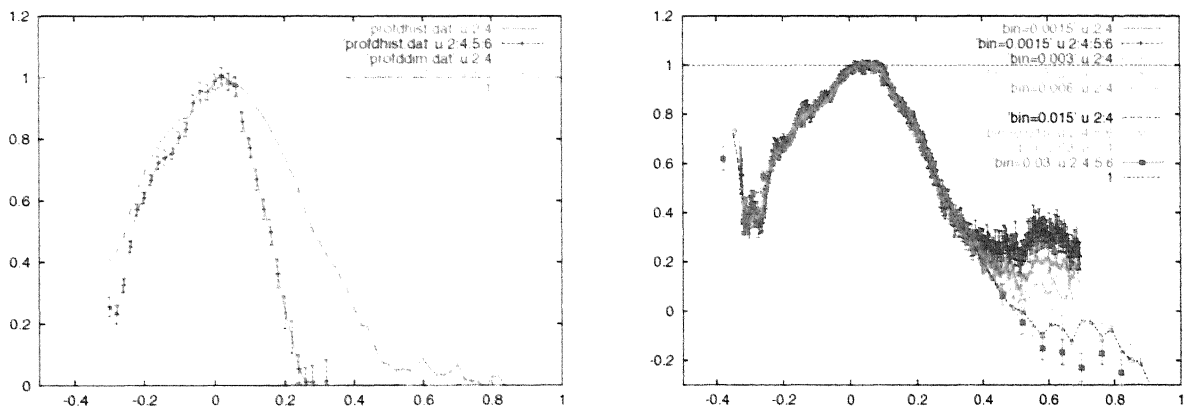
only adding to computation costs. In the case of the band width ε, only inherent to the direct method, we would, however, need to be assured of some elementary degree of stability. The experiments indicate that the spectrum obtained remains stable for a wide choice of ε without loss of quality. An example is shown in Fig. 12(a), where spectra are calculated with $\varepsilon = 0.02$, $\varepsilon = 0.01$ and $\varepsilon = 0.005$.

At the cost of slightly lower stability, we obtain the advantages of the direct spectrum calculation. The spectrum better captures local variations in the scaling of the h bands, where the partition function method provides only rough, "outline" information about the $D(h)$ spectrum. In particular, the partition function spectrum can be dramatically corrupted by outliers (e.g. the end of the sample singularities, resulting from the linear trend present in the sample). The direct method seems to be much less prone to such behavior. An experimental verification of this is shown in Fig. 12. Both types of spectra are calculated for the 16 k record of slightly correlated fBm of $H = 0.6$. This sample contains some effective linear trend in it, which results in the "trivial" end of the sample singularities (see Fig. 9). The partition function method is inherently unable to distinguish these singularities, resulting in a wide spectrum which can easily be suspected of being multifractal in origin. On the contrary, the direct spectrum quickly falls away for singularities lower than $H = 0.6$.

4.4 Direct Spectra from the (Bands of) the Effective Hölder Exponent on the Entire CWT

We have already mentioned that the procedure of direct estimation of $D(\hat{h})$ is inherently unstable due to the selecting of a very narrow band of exponents, and thus a small subset of the maxima lines per scale level. This is the reason why it provides considerably less stable scaling estimates than the partition function method, which is actually at the other extreme, taking *all* the maxima as the support for the measure of which the moments are calculated.

It seems possible to take a middle path in order to calculate more stable scaling estimates of the $D(\hat{h})$ in the direct way from the scaling of the "selected" maxima parts. This can be done by weighted selection,[21] replacing the histogram box centered at $\hat{h}$ and of ε width, with a smooth, say Gaussian, kernel of ε standard deviation, centered at $\hat{h}$. We have attempted this in a slightly different way, making use of the redundant information contained in the original CWT (as opposed to the WTMM used thus far). The comparison of the direct spectra obtained with both WTMM and the CWT suggest that the CWT may contain some information lacking in the WTMM, this is especially evident in the smooth part of the spectrum. Indeed,



(a) (b)

Fig. 13 (a) The multifractal spectrum calculated directly from the scaling of the ε band of the Hölder exponent on the entire CWT as compared with the same on the WTMM tree only. While the left part of the spectrum shows similar behavior, the right part, related to smooth behavior, departs largely from the WTMM case, possibly capturing some missing information. (b) CWT direct spectra show very good stability with respect to varied band width ε. Shown are overlapped cases $\varepsilon = 0.03$, $\varepsilon = 0.015$, $\varepsilon = 0.006$, $\varepsilon = 0.003$ and $\varepsilon = 0.0015$. Note the smaller h range used.

the maxima lines primarily restrict the representation to the strongest cusp singularities, potentially leaving out the intermediate, relatively smoother behavior. The CWT direct spectra show excellent stability with respect to the ε band width variation. In Fig. 13, we went down to spectacular $\varepsilon = 0.0015$ resolution and observed the main body of the spectrum conserved with only the background noise slowly increasing.

5. CONCLUSIONS

We have presented a method of estimating an effective Hölder exponent locally for an arbitrary resolution. The method is motivated by the multiplicative cascade paradigm, and implemented on the hierarchy of the wavelet transform modulus maxima tree. Contrary to the intrinsically unstable local slope of the maxima lines, this estimate is robust and provides a stable, effective Hölder exponent, local in scale and position.

We have presented a number of real-life examples using this local exponent estimate. The color exponent panels included show the intricate scale-free structure of the time series. The exact implications of this structure, of course depend on the application.

In addition, we have illustrated the possibility of the direct estimation of the scaling spectrum of the effective Hölder exponent and linked it to the established partition function multifractal formalism. Almost as stable as the global scaling estimates from the partition functions method, the direct histogram of the effective Hölder exponent provides considerably more information about the relative density of local scaling exponents, and may prove to be an interesting alternative in multifractal spectra estimation. In particular, we have considered the direct estimation of the singularity spectra from the entire CWT. In addition, to improved stability, these spectra can possibly reveal more complete singularity contents in comparison with WTMM-based spectra.

ACKNOWLEDGMENTS

The author would like to thank Maria Haase and Jacques Lévy Véhel for discussions during Fractal '98 on the behavior of negative moments in the multifractal spectra. Thanks also to Hans van den Berg for providing the aphids data (originally produced by Dr. Freddie Tjallingii from the Department of Entomology at Wageningen Agricultural University) and for related discussions. Plamen Ivanov, thanks for providing the heartbeat data[25] and for discussions. Shlomo Havlin deserves special thanks for suggesting we include a color plot of the h exponent in Ref. 6, which was the starting point of this work. Arno Siebes, thanks for supporting this work. This work has been carried out with the financial support from the Impact project.

REFERENCES

1. A. Arneodo, E. Bacry and J. F. Muzy, "Wavelets and Multifractal Formalism for Singular Signals: Application to Turbulence Data," *PRL* **67**, 3515–3518 (1991).
2. A. Arneodo, E. Bacry and J. F. Muzy, "The Thermodynamics of Fractals Revisited with Wavelets," *Physica* **A213**, 232–275 1995; J. F. Muzy, E. Bacry and A. Arneodo, "The Multifractal Formalism Revisited with Wavelets," *Int. J. Bifurc. Chaos* **4**, 245–302 (1994).
3. A. Arneodo, A. Argoul, J. F. Muzy, M. Tabard and E. Bacry, "Beyond Classical Multifractal Analysis Using Wavelets: Uncovering a Multiplicative Process Hidden in the Geometrical Complexity of Diffusion Limited Aggregates," *Fractals* **1**, 629–646 (1995).
4. A. Arneodo, E. Bacry, P. V. Graves and J. F. Muzy, "Characterizing Long-Range Correlations in DNA Sequences from Wavelet Analysis," *PRL* **74**, 3293–3296 (1995).
5. P. Ch. Ivanov, M. G. Rosenblum, C.-K. Peng, J. Mietus, S. Havlin, H. E. Stanley and A. L. Goldberger, "Scaling Behaviour of Heartbeat Intervals Obtained by Wavelet-Based Time Series Analysis," *Nature* **383**, 323–327 (1996).
6. P. Ch. Ivanov, M. G. Rosenblum, L. A. Nunes Amaral, Z. R. Struzik, S. Havlin, A. L. Goldberger and H. E. Stanley, "Multifractality in Human Heartbeat Dynamics," *Nature* **399**, 461–465 (1999).
7. Z. R. Struzik, "The Wavelet Transform in the Solution to the Inverse Fractal Problem," *Fractals* **3**(2), 329–350 (1995); Z. R. Struzik, *From Coastline Length to Inverse Fractal Problem: The Concept of Fractal Metrology*, Thesis (University of Amsterdam, 1996).
8. Z. R. Struzik, "Local Effective Hölder Exponent Estimation on the Wavelet Transform Maxima Tree," in *Fractals: Theory and Applications in Engineering*, eds. M. Dekking, J. Lévy Véhel, E. Lutton, C. Tricot (Springer Verlag, 1999), pp. 93–112.
9. I. Daubechies, *Ten Lectures On Wavelets* (S.I.A.M., 1992).
10. M. Holschneider, *Wavelets — An Analysis Tool* (Oxford Science Publications, 1995).

11. S. Jaffard and Y. Meyer, "Wavelet Methods for Pointwise Regularity and Local Oscillations of Functions," *Memoirs of AMS*, Vol. 123, No. 587 (1996).
12. B. Guiheneuf and J. Lévy Véhel, "2-Microlocal Analysis and Application in Signal Processing," in *Proc. Int. Wavelets Conference, Tangier* (1998).
13. S. Jaffard, "Multifractal Formalism for Functions: I. Results Valid for all Functions, II. Self-Similar Functions," *SIAM J. Math. Anal.* **28**(4), 944–998 (1997).
14. A. Arneodo, E. Bacry and J. F. Muzy, "Oscillating Singularities in Locally Self-Similar Functions," *PRL* **74**(24), 4823–4826 (1995).
15. S. G. Mallat and W. L. Hwang, "Singularity Detection and Processing with Wavelets," *IEEE Trans. Information Theory* **38**, 617–643 (1992).
16. S. G. Mallat and S. Zhong, "Complete Signal Representation with Multiscale Edges," *IEEE Trans. PAMI* **14**, 710–732 (1992).
17. A. Arneodo, E. Bacry and J. F. Muzy, "Solving the Inverse Fractal Problem from Wavelet Analysis," *Europhysics Letters* **25**(7), 479–484 (1994).
18. Z. R. Struzik, "Removing Divergences in the Negative Moments of the Multi-Fractal Partition Function with the Wavelet Transformation," *CWI Report, INS-R9803* (1998). See also *Fractals and Beyond — Complexities in the Sciences*, ed. M. M. Novak (World Scientific, 1998), pp. 351–352.
19. K. Falconer, *Fractal Geometry — Mathematical Foundations and Applications* (John Wiley, 1990).
20. Y. Liu, P. Cizeau, P. Gopikrishnan, M. Meyer, C.-K. Peng and H. E. Stanley, *Volatility Studies of the S&P 500 Index*, Preprint (1998); P. Cizeau, Y. Liu, M. Meyer, C.-K. Peng and H. E. Stanley, "Volatility Distribution in the S&P 500 Stock Index *Physica*, **A245**, 441–445 (1997).
21. Ch. Canus, J. Lévy Véhel and C. Tricot, "Continuous Large Deviation Multifractal Spectrum: Definitions and Estimation," in *Fractals and Beyond — Complexities in the Sciences*, ed. M. M. Novak (World Scientific, 1998), pp. 117–128.
22. J. Lévy Véhel and R. Vojak, "Multifractal Analysis of Choquet Capacities: Preliminary Results," in *Advances in Applied Mathematics* **20**(1), 1–43 (1998).
23. J. Lévy Véhel, "Numerical Computation of the Large Deviation Multifractal Spectrum," in *CFIC, Rome* (1996).
24. J. Lévy Véhel and R. H. Riedi, "Fractional Brownian Motion and Data Traffic Modeling: The Other End of the Spectrum," in *Fractals in Engineering*, eds. J. Lévy Véhel, E. Lutton and C. Tricot (Springer Verlag, 1997); R. H. Riedi and J. Lévy Véhel, *TCP Traffic is Multifractal: A Numerical Study*, INRIA Research Report, No. 3129 (1997).
25. *Heart Failure Database* (Beth Israel Deaconess Medical Center, Boston, MA).