

Local Effective Hölder Exponent Estimation on the Wavelet Transform Maxima Tree

Zbigniew R. Struzik

Centre for Mathematics and Computer Science (CWI)
Kruislaan 413, 1098 SJ Amsterdam - THE NETHERLANDS
email: Zbigniew.Struzik@cwi.nl

Abstract. We present a robust method of estimating an 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 tree. In addition, we illustrate the possibility of the direct estimation of the scaling spectrum of the effective Hölder exponent, and we link it to the established partition functions based multifractal formalism. We motivate both the local and the global multifractal analysis by showing examples of computer generated and real life time series.

keywords

multifractal analysis, wavelet transform, Hölder exponent

1 Introduction

The application of the wavelet transform (modulus maxima) representation of a signal to multi-fractal analysis has almost reached the status of a standard. The formalism developed by Arneodo et al in the early nineties [1, 2] 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, 4, 5]. Nevertheless the respective methodology is intrinsically statistical in nature and 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.

We, therefore, address the problem of 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 wavelet transform modulus maxima tree, which has proven to be an excellent tool for the purpose [2, 6]. Contrary to the intrinsically instable local slope of the maxima lines, this estimate is robust and provides a stable, effective Hölder exponent, local in scale and position. From this an attempt can be made to derive the multifractal spectra directly from log-histogram scaling evaluation, linking the local analysis with the global multifractal spectra approach. Not 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 with some added on stabilisation, may prove to be an interesting alternative in multifractal spectra estimation.

The structure of the paper is as follows. In section 2, we focus on the relevant aspects of the wavelet transformation, in particular the ability to characterise scale-free behaviour through the Hölder exponent. Together with the hierarchical scale-wise decomposition provided by the wavelet transform, it will enable us to reveal the scaling properties of the tree of the multiplicative cascading process. In section 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 section 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 functions 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 Wavelet Transformation (WT), see e.g. Ref. [7, 8], provides a way of analysing the local behaviour 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 behaviour to some predefined degree. Therefore, correct characterisation 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 behaviour. [2]

Conceptually, the wavelet transformation is a convolution product of the time series with the scaled and translated kernel - the wavelet $\psi(x)$, usually a n -th derivative of a smoothing kernel $\theta(x)$. Usually, in the absence of other criteria, the preferred choice is the kernel, which is well localised 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 localisation 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 analysing wavelet is determined by the parameter b :

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

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

In figure 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

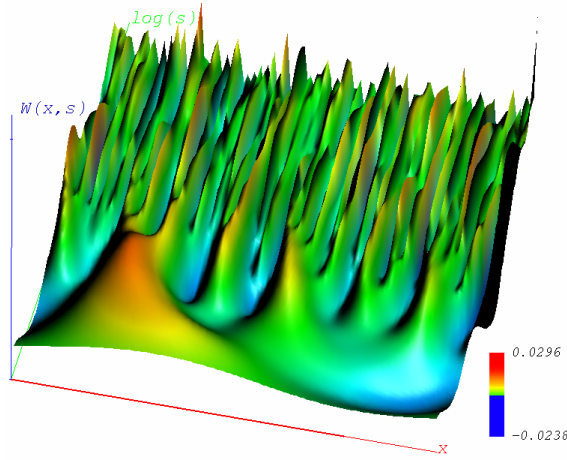


Fig. 1. Continuous Wavelet Transform representation of the random walk (Brownian process) time series. The wavelet used is the Mexican hat - the second derivative of the Gaussian kernel. The coordinate axes are: position x , scale in logarithm $\log(s)$, and the value of the transform $W(s, x)$.

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 ‘in depth’ of the plot is $\log(s)$. The third 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’, 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 varies with the scale factor s .

2.1 Assessing Singular Behaviour with the Wavelet Transformation

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

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

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

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 expansion does not exist [11]. 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 characterised by Hölder exponent $h(x_0)$ in x_0 , and f can be locally described as:

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

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

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

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 [9]. In particular, we will restrict the scope of this paper to Hölder singularities for which the local and pointwise Hölder exponents are equal [10]. Thus we will not take into consideration the ‘oscillating singularities’ (e.g. $x^\alpha \sin(1/x^\beta)$) requiring two exponents [11, 12]. Nevertheless, it is sufficient for our purpose to state that the continuous wavelet transform can be used for characterising the Hölder singularities in the time series even if masked by the polynomial bias.

It can be shown [13] that for Hölder 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 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 we show 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 figure 2 left; 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-handed limit. The Hölder exponent of a Dirac pulse is -1 by definition. For Hölder 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 $I(2048)$.

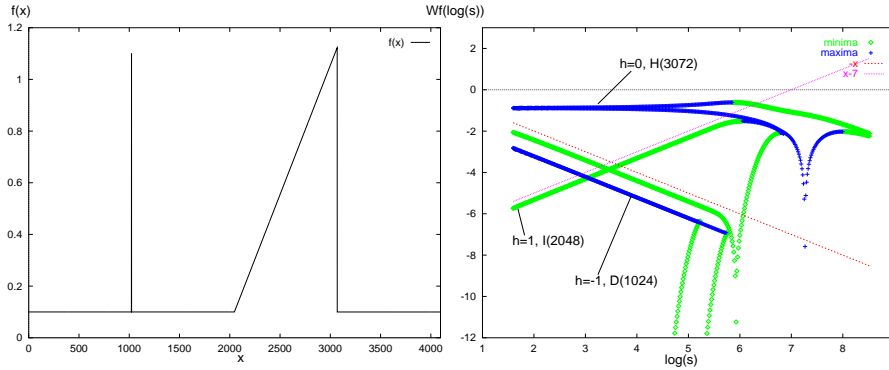


Fig. 2. Left: the test signal consisting of the Dirac pulse $D(1024)$, the change in slope - integrated Heaviside step $I(2048)$, and the Heaviside step $H(3072)$. Right: the log-log plot of the maxima, together with their respective logarithmic derivatives, corresponding to all three singularities: $D(1024)$, $I(2048)$ 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.

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 figure 2 right. This, of course, suggests the possibility of the estimation of the Hölder exponent of (Hölder) 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 figure 2 right only below some critical scale s_{crit} , below which the singularities effectively become isolated for the analysing wavelet. Indeed, the distance between the singular features in the test time series in figure 2 left equals 1024, which is in the order of three standard deviations of the analysing wavelet at $(\log(s_{crit}) = 5.83 = \log(1024/3))$. 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 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 redundancy 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 .

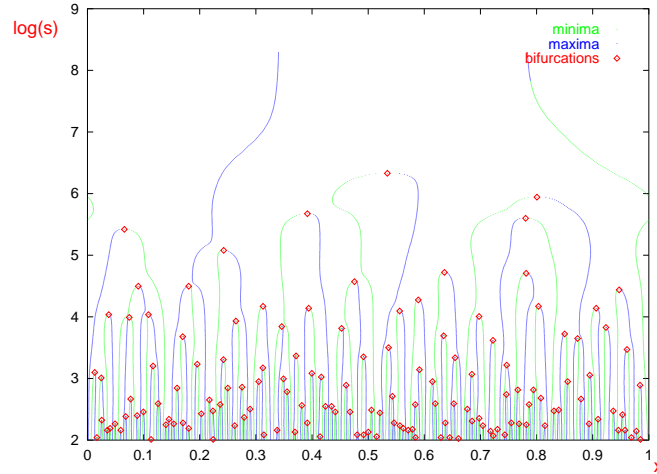


Fig. 3. WTMM representation of the time series and the bifurcations of the WTMM tree. Mexican hat wavelet.

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, first introduced by Mallat [14]. In addition to translation invariance, the WTMM also possesses the ability to characterise fully the local singular behaviour 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/renormalisation operation [6, 2, 15, 12]. This property makes it an ideal tool for revealing the renormalisation structure of the (hypothetical) multiplicative process underlying the analysed time series.

Suppose we have the time series f invariant with respect to some renormalisation 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 [15, 6].

The aforementioned properties of the maxima lines representation make it particularly useful for our purpose.

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 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, 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 estimation problems, we will use the characterisation with the model based approximation of the local scaling exponent, which we will refer to as an *effective* Hölder exponent of the singularity.

In order to estimate this exponent in real life time series with dense singular behaviour, we need to approach the problem of diverging maxima values in log-log plots and the problem of slope fluctuations. We used the procedure of bounding the local Hölder exponent as described in [16] to pre-process 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 calculations. (For the technical details of local slope calculation and bounding, we refer the reader to [16].) E.g. compare in figure 2 the example 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 $|\tilde{h}| < 2$ bound on the local slope \tilde{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 figure 4 below, 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 4 left. Bounding local slopes to $|\tilde{h}| < 2$ 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.¹ In figure 4 right, we verify in log-log coordinates that the decay of the small values follows a power law.

Even though instead of fluctuating wildly between $+\infty$ and $-\infty$, the WT values are now more tempered, they still fluctuate, with the local slope changing from point to point (within bounds). Of course, this is why it is not possible to evaluate the Hölder exponent by a linear fit in log-log plot, something we can do for isolated cases giving a stable maximum value decay/increase. Therefore, we resort to the second assumption, in which we model the singularities as created in some kind of a collective process of a very generic class. For the estimation of the local Hölder exponent in such time se-

¹ One can argue about which is better: removing the apparently relevant information from the distribution of the maxima values or lacking the negative moments... Since we do not have proof that the information removed is redundant, we do not provide a definitive answer here.

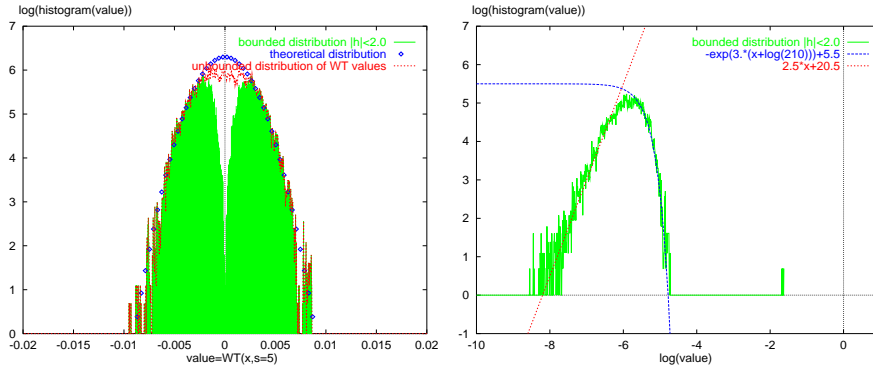


Fig. 4. Left: distribution of WT values of fBm of $H=0.6$ with and without local slope bounding in log-normal coordinates. Right: bounded distribution in log-log scale. Bound $|\check{h}| < 2$, scale $s = 5$. Sample length 16386.

ries, we will use a multiplicative cascade model. This will allow us to construct a stable estimate of a local $h(x_0)$ exponent. The multiplicative cascade model is a generalisation of a binomial multiplicative process, otherwise known as the Besicovitch binomial process.

3.1 Multiplicative Cascade Model

Let us take the well known example of the Besicovitch measure on the Cantor set, see e.g. [17]. 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 normalisation requirement:

$$p_1 + p_2 = 1. \quad (3)$$

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 measure on the Cantor set. We have the Besicovitch measure for non-equal p_i , with the other settings above retained. Finally, for non-equal p_i , regardless of normalisation Eq 3 and with $c_1 = c_2 = 2$ with $b_1 = 0$ and $b_2 = 1$, we have the multiplicative cascade on $(0..1)$ interval.

Each point of this cascade is uniquely characterised by the sequence of weights $(s_1 \dots s_n)$ taking values from the (binary) set $\{1, 2\}$, and acting successively along a

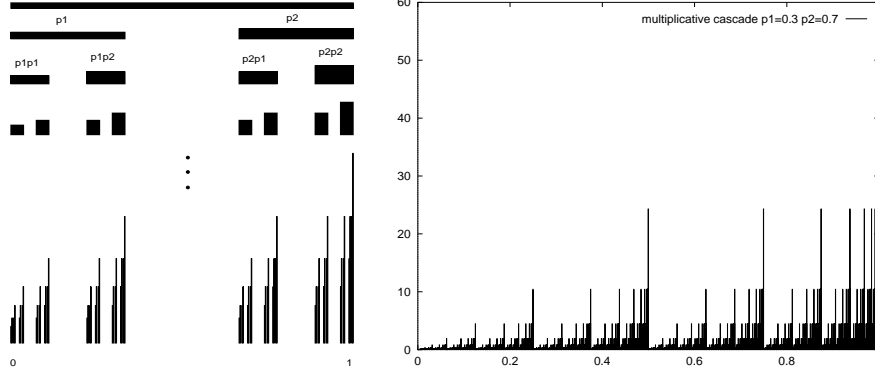


Fig. 5. Left: the Besicovitch measure on the Cantor set, generations F_0 through F_3 and the generation F_6 . The distribution of weights is $p_1 = 0.4$ and $p_2 = 0.6$. The standard middle third Cantor division is retained. Right: similar construction but on 0.1 support instead of the Cantor set, leading to multiplicative cascade. $p_1 = 0.3$ and $p_2 = 0.7$, generation F_{13} .

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

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

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

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

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 [6]. 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_{l_o} corresponds with generation F_{max} , while s_{hi} corresponds with generation F_0 .

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.

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:

$$\mathcal{M}(s) = \frac{\mathcal{Z}(s, 1)}{\mathcal{Z}(s, 0)} , \quad (4)$$

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

$$\mathcal{Z}(s, q) = \sum_{\Omega(s)} (Wf_{\omega_i}(s))^q , \quad (5)$$

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 [16]. This mean gives the direct possibility of estimating the mean value of the local Hölder exponent as a linear fit to \mathcal{M} :

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

We will not, however, use the definition 4 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 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:

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

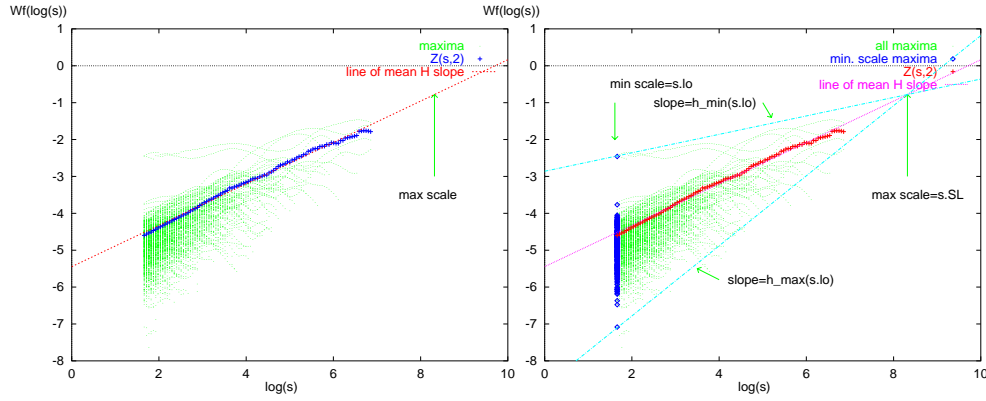


Fig. 6. Left: the projection of the maxima lines of the WT along time. The mean value of the Hölder exponent can be estimated from the log-log slope of the line shown. Also, the beginning of the cascade at the maximum scale s_{hi} is indicated. Right: the maxima at the smallest scale considered are shown 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.

4 Employing the Effective Hölder Exponent in Local and Global Spectra Estimation

Such an estimated local $\hat{h}(x_0, s)$ can be depicted in the temporal fashion, for example with a background colour, as we have done in figure 7. The first example time series is a computer generated sample of fractional Brownian motion with $H = 0.6$. It shows almost monochromatic behaviour, centred at $H = 0.6$ the colour 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&P500* index from time period [1984-1988]. There are significant fluctuations in colour in this picture, with the green colour centred at $H = 0.5$, indicating both smoother and rougher components. In particular, one can observe an extremal red value at the crash '87 coordinate, followed by very rough behaviour (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), see e.g. [18].

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 characterise 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 equal $H = 0.5$, the result quite convincingly shows the patchy difference in characterisation of the pits at (two) different levels of pit hierarchy. Stripes of a different colour spectrum indicate a high level of non-stationarity of \hat{h} distribution. Note that the obvious amplitude difference does not influence the colour in the plot due the fact that constant offset is filtered out by the wavelet used - the colour 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 localisation (or non-stationarity) of the exponent strength, this plot shows an intricate structure of interwoven singularities at various strengths. This behaviour has been recently reported [19] to correspond with the multifractal behaviour of the heartbeat. The green is centred 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 Scale-wise Evolution of the Effective Hölder Exponent

In addition to one scale plot showing the colour spectrum of singular behaviour, we can also see the scale position locations where the effective Hölder exponent is near a particular value. We show an example *band* of $\hat{h}_\epsilon(s)$ of width $\epsilon = 0.02$, by selecting $\hat{h} = -0.5 \pm 0.01$ in figure 8 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 visualising 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.

$$D(\hat{h}) = \dim(\{x_0\} : Tf(x - x_0) \sim |x - x_0|^{h(x_0)}) \sim \lim_{\epsilon \rightarrow 0} \lim_{s_{l_0} \rightarrow 0} \frac{\log(\mu_\epsilon(\hat{h}(s_{l_0})))}{\log(s_{l_0})},$$

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_{l_0} .

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. Therefore, it provides considerably less stable scaling estimates than the commonly used *partition functions*, or *Legendre transform*, method, see Ref [2]. The partition functions method is actually at the other extreme, taking *all* the maxima as the support for the measure of which the moments are calculated. This is done through the partition function \mathcal{Z} , Eq. 5. The $D(h)_m$ spectrum is then obtained from the scaling of the partition function using the Legendre transformation: $D_m(h) \stackrel{\mathcal{L}}{\leftrightarrow} \tau(q)$, if $\mathcal{Z}(s, q) \sim s^{\tau(q)}$. In this construction, the moment parameter q has the purpose of ‘selecting’ an adequate range of Hölder exponents from the global quantity \mathcal{Z} . As a result, the partition function method provides only rough, ‘outline’ information about the $D_m(h)$ spectrum.

It seems possible to take a middle path in order to calculate more stable scaling estimates of the $D(h)$ in the direct way from the scaling of ‘selected’ maxima parts. This can be done by weighted selection, replacing the histogram box centred at \hat{h} and of ϵ width, with a smooth, say Gaussian, kernel of ϵ standard deviation, centred at \hat{h} .

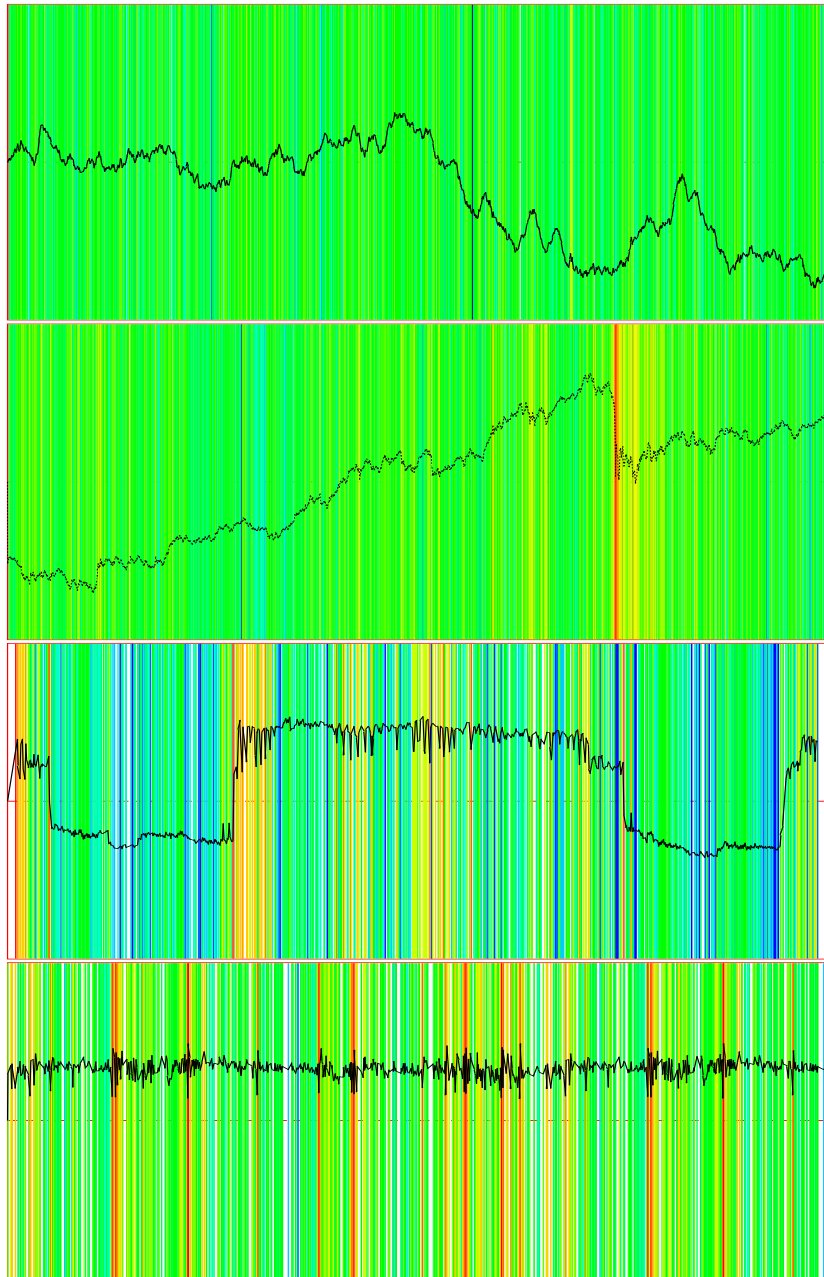


Fig. 7. Example time series with local Hurst exponent indicated in colour. From top to bottom: fBm with $H = 0.6$, a record of S&P500 index, penetratogram of aphids, and the last is the record of healthy heart interbeat intervals. The background colour indicates the Hölder exponent locally, centred at the Hurst exponent at green, colour goes towards blue for higher \hat{h} and towards red for lower \hat{h} . Local slope bounds for all the plots $|\hat{h}| \leq 2$.

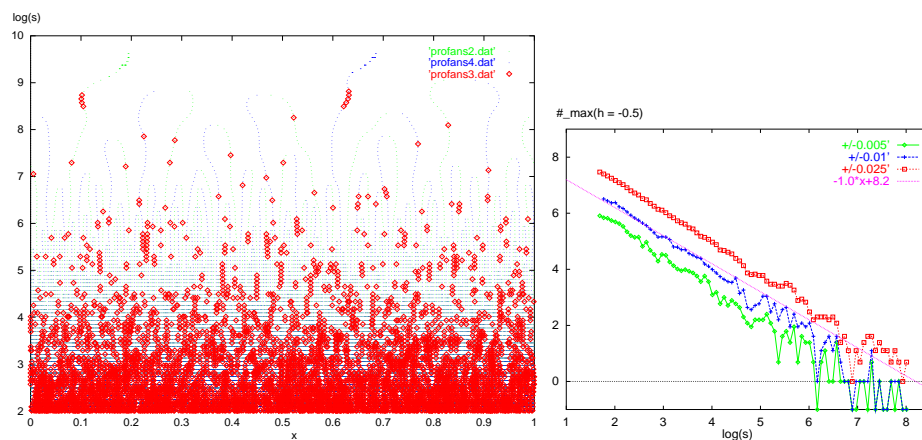


Fig. 8. Left: WTMM representation of a sample of white noise. The maxima are highlighted where the effective Hölder exponent reaches a particular value of $\hat{h} = -0.5 \pm 0.01$, i.e. $\epsilon = 0.02$. Right: for three values of the bin size $\epsilon = 0.01, \epsilon = 0.2, \epsilon = 0.05$, the 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 $\epsilon = 0.02$ plot.

While we will not pursue this approach further here, leaving it to a separate treatment, it seems that a similar idea has successfully been applied to large deviation multifractal spectra estimation on dyadic partitions as reported in [20]. For a description of this multifractal spectrum, see [21, 22]

4.2 Log-histograms of the Effective Hölder Exponent

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

We will analyse 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 $\mu_\epsilon(s)$ 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 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 which can be analysed 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 ‘normalisation’ of the maximum of $P(\hat{h})$, which corresponds with the scaling of the zeroth moment in the partition functions method. (Note that the scaling assumption can be verified both by analysing the scaling of the moments and by the

scaling of the \hat{h} bin contents.) This is why in figure 9, we show the non-normalised histograms for different scale values, and compare them to the $D_m(h)$ spectrum obtained with the partition functions method.

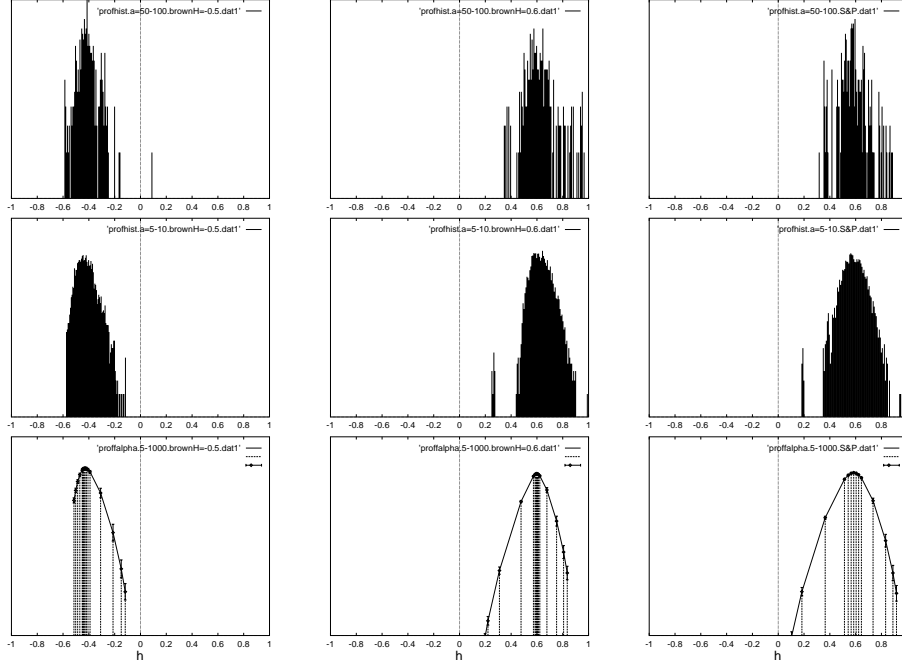


Fig. 9. Two sets of \hat{h} histograms for respective scales $\log(s) \simeq 3.9$, $\log(s) \simeq 1.6$, in top first and centre row. Below, in the bottom row, the $D_m(h)$ spectrum obtained with the partition functions method (moments $-5 \leq q \leq 5$). Left column: for 4096 samples of white noise. Centre: 4096 samples of fractional Brownian motion with $H = 0.6$. Right: S&P 500 index, first 4096 samples from figure 7. Local slope bounds for all the plots $|\check{h}| \leq 2$.

For three example time series, we show in figure 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 behaviour will become dispersed. On the contrary, while going down with the scale, the bulk of consistent behaviour converges to one distribution. The consistent statistical behaviour is also captured in the $D_m(h)$ spectrum obtained with the partition functions method shown in the bottom row of figure 9.

Several aspects of the $D_m(h)$ versus $P(\hat{h})$ already discussed in [23] 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 analysing 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.

Both the distributions $P(\hat{h})$ and the $D_m(h)$ spectra are evaluated using the local slope bound of $|\check{h}| \leq 2$, on a relatively short samples of 4096 data points. In figure 10, we show both the distribution $P(\hat{h})$ and the $D_m(h)$ spectrum evaluated with maximally tight and relaxed bounds and a relatively long length of 2^{15} samples of white noise. We can verify that for the maximally tight bounds, both the distribution and the $D_m(h)$ spectrum are well clustered around the $\hat{h} = H$ value, indicating that the bounding procedure does not remove any information relevant to recovering this part of the distribution/spectrum. Also, the width of the spectrum is slightly reduced when compared to the results from 8 times shorter data such as used in figure 9. The spectrum remains within the theoretical predictions $1 - H - h$ indicated with the line plot. The same data set processed with maximally relaxed bounds gives the $D_m(h)$ spectrum with its part for the negative moments exceeding the theoretical bounds. The histogram, however, fills the entire space granted by the theoretical bounds.

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 colour 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 direct estimation of scaling spectrum of the effective Hölder exponent and linked it to the established partition functions multifractal formalism. Not 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 with some added stabilisation through the appropriate choice of a smoothing kernel, may prove to be an interesting alternative in multifractal spectra estimation.

Acknowledgments

The author would like to thank Maria Haase and Jacques Lévy Véhel for discussions during Fractal '98 on the behaviour of negative moments in the multifractal spectra.

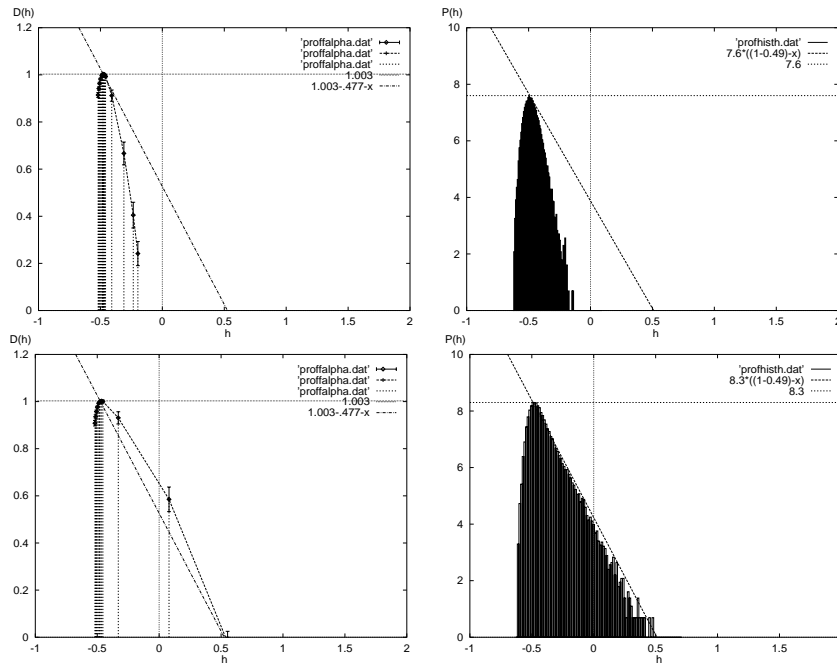


Fig. 10. White noise sample, length 32.768 data points. Left: the $D_m(h)$ spectrum obtained with the partition functions method (moments $-5 \leq q \leq 5$). Right: the corresponding distribution $P(\hat{h})$ for $\log(s) = 1.61$. Top row is obtained using the local slope bound of $|\hat{h}| \leq 2$, lower row with relaxed local slope bound. Theoretical predictions $C(1 - H - h)$ are indicated with the line plot.

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 [24] and for discussions. Shlomo Havlin deserves special thanks for suggesting we include a colour plot of the h exponent in [19], 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 (1991): Wavelets and Multifractal Formalism for Singular Signals: Application to Turbulence Data. *PRL*, **67**, No 25, 3515–3518.
2. A. Arneodo, E. Bacry and J.F. Muzy (1995): The Thermodynamics of Fractals Revisited with Wavelets. *Physica A*, **213**, 232–275.
J.F. Muzy, E. Bacry and A. Arneodo (1994): The Multifractal Formalism Revisited with Wavelets. *International Journal of Bifurcation and Chaos* **4**, No 2, 245–302.
3. A. Arneodo, A. Argoul, J.F. Muzy, M. Tabard and E. Bacry (1995): Beyond Classical Multifractal Analysis using Wavelets: Uncovering a Multiplicative Process Hidden in the Geometrical Complexity of Diffusion Limited Aggregates. *Fractals* **1**, 629–646.

4. A. Arneodo, E. Bacry, P.V. Graves and J.F. Muzy (1995): Characterizing Long-Range Correlations in DNA Sequences from Wavelet Analysis. *PRL*, **74**, No 16, 3293–3296.
5. P.Ch. Ivanov, M.G. Rosenblum, C.-K. Peng, J. Mietus, S. Havlin, H.E. Stanley and A.L. Goldberger (1996): Scaling Behaviour of Heartbeat Intervals Obtained by Wavelet-based Time-series Analysis. *Nature*, **383**, 323–327.
6. Z.R. Struzik (1995): The Wavelet Transform in the Solution to the Inverse Fractal Problem. *Fractals* **3** No.2, 329–350.
Z.R. Struzik (1996): From Coastline Length to Inverse Fractal Problem: The Concept of Fractal Metrology. *Thesis*, University of Amsterdam.
Z.R. Struzik (1997): Fractals under the Microscope or Reaching Beyond the Dimensional Formalism of Fractals with the Wavelet Transform. *CWI Quarterly*, **10**, No 2, 109–151.
7. I. Daubechies (1992): *Ten Lectures on Wavelets*. S.I.A.M..
8. M. Holschneider (1995): *Wavelets - An Analysis Tool*. Oxford Science Publications.
9. S. Jaffard and Y. Meyer (1996): Wavelet methods for pointwise regularity and local oscillations of functions. *Memoirs of AMS*, 123(587).
10. B. Guiheneuf and J. Lévy Véhel (1998): 2-Microlocal Analysis and Application in Signal Processing, *In proceedings of International Wavelets Conference, Tangier*.
11. S. Jaffard (1997): *Multifractal Formalism for Functions: I. Results Valid for all Functions, II. Self-Similar Functions*. *SIAM J. Math. Anal.*, 28(4): 944–998.
12. A. Arneodo, E. Bacry and J.F. Muzy (1995): Oscillating Singularities in Locally Self-Similar Functions. *PRL*, **74**, No 24, 4823–4826.
13. S.G. Mallat and W.L. Hwang (1992): Singularity Detection and Processing with Wavelets. *IEEE Trans. on Information Theory* **38**, 617–643.
14. S.G. Mallat and S. Zhong (1992): Complete Signal Representation with Multiscale Edges. *IEEE Trans. PAMI* **14**, 710–732.
15. A. Arneodo, E. Bacry and J.F. Muzy (1994): Solving the Inverse Fractal Problem from Wavelet Analysis. *Europhysics Letters*, **25**, No 7, 479–484.
16. Z.R. Struzik (1998): Removing Divergences in the Negative Moments of the Multi-Fractal Partition Function with the Wavelet Transformation. *CWI Report*, **INS-R9803**. Also in ‘Fractals and Beyond - Complexities in the Sciences’, M.M. Novak, Ed., World Scientific, 351–352.
17. K. Falconer (1990): *Fractal Geometry - Mathematical Foundations and Applications*, John Wiley.
18. Y. Liu, P. Cizeau, P. Gopikrishnan, M. Meyer, C.-K. Peng and H.E. Stanley (1998): Volatility Studies of the S&P 500 Index. *Preprint*.
P. Cizeau, Y. Liu, M. Meyer, C.-K. Peng and H.E. Stanley (1997): Volatility Distribution in the S&P 500 Stock Index. *Physica A*, **245**, 441–445.
19. P.Ch. Ivanov, M.G. Rosenblum, L.A. Nunes Amaral, Z.R. Struzik, S. Havlin, A.L. Goldberger and H.E. Stanley (1998): Multifractality in Human Heartbeat Dynamics. *Preprint*.
20. Ch. Canus, J. Lévy Véhel and C. Tricot (1998): Continuous Large Deviation Multifractal Spectrum: Definitions and Estimation. ‘Fractals and Beyond - Complexities in the Sciences’, M.M. Novak, Ed., World Scientific, 117–128.
21. J. Lévy Véhel and R. Vojak (1998): Multifractal Analysis of Choquet Capacities: Preliminary Results *In Advances in Applied Mathematics*, Vol. 20, No. 1, pp. 1–43, January.
22. J. Lévy Véhel (1996): Numerical Computation of the Large Deviation Multifractal Spectrum, *In CFIC, Rome*.
23. J. Lévy Véhel and R. H. Riedi (1997): Fractional Brownian Motion and Data Traffic Modeling: The Other End of the Spectrum. *Fractals in Engineering*, J. Lévy Véhel, E. Lutton, C. Tricot, Eds., Springer Verlag.

R. H. Riedi and J. Lévy Véhel (1997): TCP Traffic is Multifractal: a Numerical Study. *INRIA Research Report*, No. 3129.

24. *Heart Failure Database* (Beth Israel Deaconess Medical Center, Boston, MA).