FRACTAL DIMENSION ESTIMATION: EMPIRICAL MODE DECOMPOSITION VERSUS WAVELETS

Paulo Gonçalves

INRIA - ENS Lyon, LIP, RESO
46 allée d'Italie, 69364 Lyon, France.
Paulo.Goncalves@inria.fr

Patrice Abry, Gabriel Rilling, Patrick Flandrin

CNRS UMR 5672 - ENS Lyon, Physics Dept.,
46 allée d'Italie, 69364 Lyon, France.
{firstname.lastname}@ens-lyon.fr

ABSTRACT

We address the problem of fractal dimension estimation of a discrete sample path. After recalling the multiplicity of possible definitions, we focus on the regularity dimension and on the regularization dimension, and report on the common ingredients that underlie these definitions: a scale transform of the signal, and a geometric or statistical measure on the scaled signal. Then, we propose to interchange wavelet transforms, ordinarily used as the scale transform, with empirical mode decomposition (EMD), a recently proposed signal-adaptive transform. The adaptivity of this latter yields estimation performance that overhauls usual wavelet-based techniques. To support our claim, we obtain comprehensive results from a Monte Carlo simulation on fractional Brownian motions.

Index Terms— fractal dimension, regularity exponents, wavelet transform, EMD

1. MOTIVATION

Fractal dimensions, as indicators of the (ir-)regularity of function graphs, went beyond the theoretical limits of mathematics, to become in the last two decades a reckoned measurement tool in signal and image processing. Measuring the regularity of a signal path, or of an object contour, is now often used to perform data characterization and classification in a host of real world applications, such as biomedical, economics, mechanics, tribology, … and even so, it still is a vivid research area with both theoretical and practical issues.

Mathematically, one can define the fractal dimension of a function f in many different ways, and associate them to as many regularity measurements. But in all cases, each definition relies on a particular *scale* or *resolution* dependent transformation f_a of f, and on how a certain quantity $\mathcal{M}_f(a)$ computed on f_a evolves with a. Then, the fractal formalism posits that $\mathcal{M}_f(a)$ behaves as a power law with respect to the analysis scale a, and that the regularity strength symply relates to the power law exponent.

A theoretically and pratically sensible choice for the global measure $\mathcal{M}_f(a)$ is to consider the finite order moments of f_a, or more simply to focus on the sole variance Var f_a (i.e. energy of f_a), when normal data assumption is a valid hypothesis. It is likely that this choice is closely related to the success of wavelet transforms (WT) that have proven relevant tools to play the role of the scale transformation f_a, and that, as such, have been widely used in the fractal community [1]. More recently though, an adaptive scale dependent decomposition, referred to as *empirical mode decomposition* (EMD), has been gaining interest in applications, and was already considered for regularity estimation purpose [2]. At the same time, alternative measures were explored. For instance, in [3], a measure corresponding to the algebraic length of the rectifiable[1] function graph Γ_a of f_a is proposed.

In section 2, we recall the definitions of two different fractal dimensions: the *regularity dimension* and the *regularization dimension*. Both imply different regularity definitions and different global measures $\mathcal{M}_f(a)$, that we will strive to clarify. In a second step, we intend to compare the statistical performance, benefits, drawbacks and computational costs of several estimation procedures resulting from the various combinations and declensions of scale transforms and global measurements. WT and EMD are briefly introduced in a common perspective in Section 3, and the estimation procedures themselves are defined in Section 4. In the course, we particularily emphasize a novel approach that exploits EMD assets to estimate the regularization dimension. To carry these comparisons through, we conducted extensive Monte Carlo simulations on large set of fractional Brownian motions (fBm), a paradigm for Gaussian self-similar processes. This reference stochastic process is a pertinent choice because all different fractal dimensions are equal and controlled by a unique self-similarity parameter H. Results are reported in Section 5 and demonstrate the clear benefit of combining EMD with the regularization dimension.

2. FRACTAL DIMENSIONS

There exist many different definitions for the fractal dimension of a function of $f \in \mathbb{R}^d$ [4]. However, all of them describe the way a given measure $\mathcal{M}$ applied to the function evolves as the observational scale varies.

The regularity dimensions relate to the global or to the pointwise regularity of a continuous, but non-differentiable function[2] $f \in \mathbb{R}^d$. More precisely, the global regularity of f over a closed bounded ball K of $\mathbb{R}^d$ is defined as the largest $\alpha_g \in]0,1[$ such that:

$$\exists C > 0, \forall x, y \in K, |f(x) - f(y)| < C|x-y|^{\alpha_g}. \tag{1}$$

Here, the variation $|f(x) - f(y)|$ plays the role of the measure $\mathcal{M}$ on f, and definition (1) shows that for all observation scales $|x - y|$, this measure is bounded by a power law of exponent α_g. Then, in the case of a function $f : \mathbb{R} \to \mathbb{R}$, the corresponding fractal dimension of its graph Γ is simply $\dim_g(\Gamma) = 2 - \alpha_g$. Similarly, to define the pointwise regularity α_p of f at point x_0, we need to confine the observations to a ball $B(x_0, \rho)$ centered on x_0 and of radius ρ, and

[1] The graph Γ of f is not a rectifiable curve, otherwise its fractal dimension would be inevitably equal to 1.

[2] For the sake of simplicity, we consider only non-differentiable functions, but all results are more generally established for C^N functions, in which case $N < \alpha_g < N + 1$.

consider the *oscillation* $O_f(x_0,\rho) = \sup_{x,y\in B(x_0,\rho)} |f(x)-f(y)|$ as our new local measure on f. The pointwise regularity is then defined as:

$$\alpha_p = \overline{\lim_{\rho\to 0}} \frac{\log(O_f(x_0,\rho))}{\log\rho}. \quad (2)$$

Here again, for 1-D functions, the pointwise fractal dimension is $\dim_p(\Gamma) = 2-\alpha_p$, and it conveys the convergence rate of the oscillation measure when the observational scale goes towards zero.

The regularization dimension was introduced in [3]. Let Γ be the graph of a bounded function f defined on some closed bounded interval $K\subset\mathbb{R}$. Let f_a be the projection of f on some approximation space at lower resolution $a>0$ (such that $\lim_{a\to 0} f_a = f$), and define $\mathcal{L}(a) = \int_K \sqrt{1+f_a'(x)^2}\,dx$, the finite length of the corresponding graph Γ_a. Then, the regularization dimension reads:

$$\dim_{\mathcal{R}}(\Gamma) = 1 + \overline{\lim_{a\to 0}} \frac{\log(\mathcal{L}(a))}{\log(a)}, \quad (3)$$

and it derives from the same principle that underlies all fractal dimensions, that is to fit a measure (here, the trajectory length) with a polynomial model, to reflect how fast it varies as resolution increases.

Many other definitions of fractal dimension exist, but it is remarkable that for a large class of functions, all these dimensions are equal to each other. This is notably the case for fractional Brownian motions (fBm's) $B_H(t)$ with Hurst exponent $0<H<1$, for which we have almost surely $\dim_H(B_H) = \dim_g(B_H) = \dim_p(B_H) = \dim_{\mathcal{R}}(B_H) = 2-H$, and $\forall t,\ \alpha_g = \alpha_p(t) = H$.

3. MULTIRESOLUTION DECOMPOSITIONS

Discrete wavelet transform. Let the pair $\{\phi,\psi\}$ form a multiresolution analysis of $L^2(\mathbb{R})$, the space of signals with finite energy. The set $\{\phi_{jk}(t) = 2^{-j/2}\phi(2^{-j}t-k),\ k=...,-1,0,1,...\}$ and the set $\{\psi_{jk}(t) = 2^{-j/2}\psi(2^{-j}t-k),\ k=...,-1,0,1,...\}$ are orthonormal basis of the approximation space $V_j\subset L^2(\mathbb{R})$, and of the detail space $W_j\subset L^2(\mathbb{R})$ respectively, moreover $V_{j-1} = V_j\oplus W_j$, and $W_k\perp W_j,\ \forall k\neq j$. Consequently, any signal $s\in L^2(\mathbb{R})$ can be decomposed into the sum:

$$\begin{aligned} s(t) &= \sum_{k=-\infty}^{k=\infty} \langle s,\phi_{Jk}\rangle\phi_{Jk}(t) + \sum_{j=-\infty}^{j=J}\sum_{k=-\infty}^{k=\infty}\langle s,\psi_{jk}\rangle\psi_{jk}(t) \\ &= A_J(t) + \sum_{j=-\infty}^{j=J} D_j(t), \end{aligned} \quad (4)$$

where $A_J(t)$ is the approximation of s at scale $a_J = 2^J$, and $D_j(t)$ the detail at scale $a_j = 2^j$.

Empirical mode decomposition. EMD is a recent technique [5] introduced to analyze non-stationary and non-linear time series in a totally adaptive way. In contrast to standard kernel based approaches (e.g. wavelet decompositions), EMD is a fully data-driven method that recursively decomposes a complex signal into a variable but finite number of zero-mean with symmetric envelopes AM-FM components called Intrinsic Mode Functions (IMF). The resulting signal expansion is similar in kind to a wavelet decomposition (4), $s(t) = R(t) + \sum_{k=1}^M C_k(t)$. where R is the residue, the signal approximation at the lowest resolution (i.e. trend), and C_k is the $k-$th IMF, the signal detail at characteristic scale a_k. To proceed, an iterative algorithm locally identifies in the signal the fastest oscillations and isolates them in the first IMF. Each successive IMF is then obtained iterating the same sifting process on the remaining lower trend.

This appealing analyzing tool is reversible by construction, and gives rise to a natural "scale" decomposition that goes beyond classic spectral analysis and its Fourier modes [6]. In particular, it was shown that applied to a fractional Gaussian noise (the increment process of a fBm), EMD spontaneously behaves as a constant-Q filter bank, similar in kind to a wavelet decomposition. However, EMD noticeably differs from a wavelet multiresolution analysis for no dyadic structure is imposed, but instead a natural characteristic scale organization emerges. In the sequel we will refer to the k-th characteristic EMD scale a_k, as the mean time interval separating two successive zero-crossing points of IMF $C_k(t)$.

4. ESTIMATORS

Listing all existing fractal dimension estimators is definitely beyond the scope of this work. Here, we deliberately focus on wavelet-based estimators that have proven among the most efficient and commonly used approaches to evaluate Hurst exponents from discrete fBm sample paths.

Our contribution is then twofold: Firstly, we use a wavelet multiresolution analysis to measure the curve length entering the definition (3) of the regularization dimension, and we compare the results with the original estimator proposed in [3]. Secondly, we consider a new multiresolution decomposition technique, namely the Empirical Mode Decomposition, that we propose as a substitute to the wavelet tool used in a wide variety of Hurst estimators.

Regularity dimension estimators. As the wavelet ψ of a multiresolution analysis is a zero-mean oscillating function, the projector on W_j is deemed a pseudo-differential operator, and the coefficient $d_{jk} = \langle s,\psi_{jk}\rangle$ is interpretable as a local variation of s at scale $a = 2^j$ and around time $t = 2^jk$. Thus, it is possible to use these wavelet coefficients to compute the different measures defining the global and pointwise regularities in (1) and (2). The variation $|s(x)-s(y)|$ can be simply replaced by the magnitude of wavelet coefficients at the appropriate scale $2^j\approx|x-y|$, and the oscillation $O_s(t\approx k2^j,\rho\approx 2^j)$, by the supremum of all wavelet coefficients at finer scales and lying within the interval $](k-1)2^j,(k+1)2^j]$. The retained coefficient is commonly refered to as the leader [7]. This way, estimating a fractal dimension simply amounts to perform a linear regression of some wavelet related quantity against the scale, in a bi-logarithmic plot. Moreover, as we know for fBm's $\alpha_p(t) = \alpha_g = H$, and the wavelet coefficients at a same scale form a stationary gaussian time series that fully justifies to built estimators of H from empirical $q-$th order moment of d_{jk}, rather than on the d_{jk}'s (or the corresponding leaders) directly:

$$\widehat{H} = \overline{\lim_{j\to-\infty}} \frac{\log(2^{-j}\sum_k |d_{j,k}|^q)}{-qj}. \quad (5)$$

In addition, applied to fBm's, estimator (5) with $q=2$ is asymptotically efficient [8].

Inspired by the same principle, in [2], an estimator similar to (5) takes advantage of the EMD adaptivity to bring in a spontaneous multiresolution scheme able at disclosing the natural scales that structure the analyzed process. So, decomposing a fBm $B_H(t)$ into its finite set of intrinsic mode functions $\{C_k(t),\ k=1,...,M\}$, the estimator reads:

$$\widehat{H} = \overline{\lim_{a_k\to 0}} \frac{\log(\operatorname{Var} C_k)}{\log a_k}, \quad (6)$$

where $\{a_k,\, k = 1, ..., M\}$ are the corresponding IMF's characteristic scales.

Regularization dimension estimators. In [3] where the regularization dimension was originally introduced, the approximation f_a was obtained by smoothing the initial function f with a scale variable low pass kernel $G_a(t) = |a|^{-1} G(a^{-1}t)$. A gaussian window G is recommended, and for a fBm $B_H(t)$, the Hurst exponent $H = 2 - \dim_{\mathcal{R}}$ is estimated according to (3) with:

$$\mathcal{L}(a) = \int_K \sqrt{1 + \left| \frac{d}{dt} (f \star G_a)(t) \right|^2} \, dt. \tag{7}$$

Here, we suggest to adapt the regularized dimension (3) to the multiresolution framework of wavelet decompositions, and consider the approximations $A_j(t) \in V_j$ as the regularized trajectories. Doing so, for a discrete fBm sample path, the graph length $\mathcal{L}(a_j)$ at scale $a_j = 2^j$, becomes:

$$\mathcal{L}(a_j) = \int_K \sqrt{1 + \left| \frac{d}{dt} \sum_{k=-\infty}^{k=\infty} \langle B_H, \phi_{j,k} \rangle \phi_{j,k}(t) \right|^2} \, dt. \tag{8}$$

Like expression (7), the proposed length (8) is parameterized by a sequence of predefined scale values $\{a_j,\, j = 1, ..., J\}$, chosen independently of the analyzed process. To alleviate this constraint, we evoke the same rationale as in [2] to resort to EMD and to partial reconstruction sums of IMFs to get a series of regularized functions f_{a_k} that live at their natural scales $\{a_k,\, k = 1, ..., M\}$. More precisely, assuming that $k = M$ corresponds to the IMF at the coarsest scale, we propose to use in (3) the following alternative length:

$$\mathcal{L}(a_k) = \int_K \sqrt{1 + \left| \frac{d}{dt} \sum_{i=k}^{i=M} C_i(t) \right|^2} \, dt, \tag{9}$$

with $B_H(t) = R(t) + \sum_{i=1}^{i=M} C_i(t)$.

At the final, we end up with 6 different estimators of the fBm's Hurst exponent H, which all ammount to perform in a *log-log plot* a linear regression of a particular scale-dependent measure against the scale: *(i)* wavelet coefficents variance or *(ii)* wavelet leaders variance in relation (5) – *(iii)* IMF's variance in relation (6) – *(iv)* smoothed graph length of expression (7) – *(v)* wavelet approximation graph length of expression (8) – and *(vi)* EMD approximation graph length of expression (9).

5. NUMERICAL SIMULATIONS

Experimental setup. To assess and compare the performances of the 6 estimators described above, they are tested on a set of 10.000 independent realizations of fBm's with the same H. Each sequence of size $N = 1024$ points is synthesized by the circulant matrix method of [9], and the test is repeated for different values of H from 0.1 to 0.9 with a 0.1 stepsize.

To account for a possible variability in the confidence intervals of the quantities to be regressed, all estimators *(i)*–*(vi)* perform a weighted least square fit of the *log*-measures against the *log*-scale. In addition, EMD-based estimators *(iii)* and *(vi)*, for which the scale a_k is itself a random variable, employ a weighted linear regression with uncertainty along the two axes.

Furthermore and according to previous findings, to cope with finite size effects and spectral aliasings, wavelet based estimators must adapt their regression scale range to the value of H. Then, in our simulations, we systematically disregard the finest and the two coarsest scales from wavelet decompositions. Regarding EMD-based estimators, this limitation does not hold, and they can operate on full IMF's decompositions. Finally, we used a Daubechies wavelet with 2 vanishing moments in estimators *(i)*, *(ii)* and *(v)*, and a gaussian kernel whose width varies from 8 to 256 points over 16 voices for the estimator *(iv)*.

Table 1. Mean Square Error of $\widehat{H}$ obtained for each estimator *(i)*–*(vi)* and for different values of the Hurst exponent H.

	(i)	*(ii)*	*(iii)*	*(iv)*	*(v)*	*(vi)*
$H = 0.1$	.0439	**.0044**	.0194	.0497	*.0057*	.0135
$H = 0.2$	.0166	**.0010**	.0099	.0156	.0058	*.0044*
$H = 0.3$	.0090	*.0051*	.0074	.0062	.0079	**.0023**
$H = 0.4$	.0060	.0059	.0083	*.0030*	.0102	**.0016**
$H = 0.5$	.0048	.0049	.0069	*.0019*	.0123	**.0016**
$H = 0.6$	.0044	.0041	.0076	**.0015**	.0139	*.0017*
$H = 0.7$	.0042	.0034	.0088	**.0016**	.0163	*.0018*
$H = 0.8$	.0040	.0031	.0117	*.0020*	.0181	**.0019**
$H = 0.9$	.0041	.0030	.0888	*.0027*	.0175	**.0012**

Results. Table 1 presents the Mean Square Errors $MSE(\widehat{H}) = (\mathbf{E}\widehat{H} - H)^2 + \mathrm{Var}(\widehat{H})$, for each estimator and each value of H. For most values of the Hurst exponent H, the estimate *(vi)* based on the partial EMD reconstruction lengths seems to perform better – or slightly worse than companion approaches based on regularization dimension. Only for small H, does the leaders technique *(ii)* seems to significantly outperforms all the others. But a closer view at the individual biases and variances, sheds a quite different light on the results. Indeed, by construction, a wavelet leader at scale $j + 1$ is necessarily larger or equal to its antecedent at finer scale j. Inevitably then, the $q-$th order moment entering the definition (5) of $\widehat{H}$ is an increasing function of the scale index j, which in turn mechanically prevents the slope $\widehat{H}$ from taking on negative values. Whereas this limitation may be beneficial for large H's (typically, $H > 0.2$), for small values of H, the distribution of the estimates being artificially truncated below zero, the resulting variance drastically reduces but at the cost of a deleterious bias. This shrinking effect is very clear on figure 1 representing the empirical $\widehat{H}$ distributions obtained with each method. Ideally, densities corresponding to different values of H should not overlap, yet with the leaders based estimator *(ii)*, the densities for $H = 0.1$ and $H = 0.2$ respecively, are totally superimposed, turning impossible to distinguish between the two exponents. This phenomenon does not appear with the other estimators, and altough a significant bias systematically occurs for small H's – even with the partial EMD reconstruction approach – the resolution strength remains fairly acceptable.

More generally, there are two different angles for looking at the estimators:

1. Comparing those which aim at estimating the regularity dimension (estimates *(i)*–*(iii)*) with those which objective is to estimate the regularization dimension (estimates *(iv)*–*(vi)*). From table 1, it is manifest that the second group of estimators performs generally better. One sensible reason for this is that, for a reasonably small sample size ($N = 1024$), estimating the measure $\mathcal{L}(a)$ proves to be more robust than estimating an empirical variance. We support our claim by assessing the power law scaling of relations (3), (5) and (6), which in practice shows less questionable for length measures than for empirical moments estimates.

2. Comparing the wavelet-based estimators (estimates *(i), (ii), (iv)*

and *(v)*) versus the EMD-based estimators (estimates *(iii)* and *(vi)*). According to [3], we include the gaussian regularized estimator *(iv)* in the group of wavelet-based estimators, for an evident connection with continuous wavelet transforms exists. From this viewpoint, conclusions seem more ambiguous, and it is not so easy to sketch a sharp frontier between the two groups. However a closer look at the first group of estimators (*(i), (ii), (iv) and (v)*) and at their corresponding MSE's, reveals that the best general results come from the gaussian kernel regularization method. Surprizingly, this is among all estimators of its group, the sole that relies on a (equivalent) continuous wavelet transform. This raises the interesting question of information redundancy and of its usefulness in small data sets scenarios [10].

Beyond the relative satisfacory performance achieved with EMD for regularization dimension, it is important to stress the pros and the cons inherent to this method:

– EMD provides with a totally adaptive multiresolution scheme that, in the case of a fBm (or any self-similar process), does not necessitate to tune the regression scale range, nor it calls for border effects correction;

– More promisingly, the great interest of EMD applied to fractal dimension estimation should arise when addressing processes with discrete rather than continuous scale invariance (e.g. Weierstrass processes, Mandelbrot multiplicative cascades, oscillating chirp driven data...) In these cases, a dyadic (or any apriori fixed) dichotomy may not necessarily be the appropriate thing to do. On the other hand, an adaptive multiresolution approach like EMD, will naturally lead the analysis along the characteristic scales of the process.

– As far as computational cost is concerned, EMD is incomparably more expensive than a discrete wavelet transform. An obstacle that can severely penalize its use with large sample sizes.

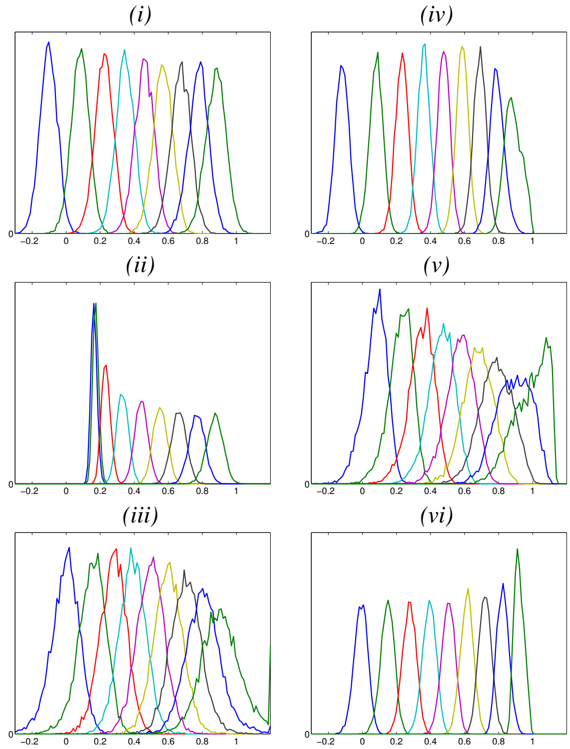

Fig. 1. For each estimator *(i)* – *(vi)* of fBm Hurst exponent, empirical distributions of $\widehat{H}$ when H varies from 0.1 (left densities on the plots) to 0.9 (right densities on the plots).

6. CONCLUSIONS AND PERSPECTIVES

To measure fractal dimension, practitioners first have to choose a scale dependent measure to suppport estimation. In applications, this choice often appears as ad hoc. Our first contribution was to reemphasize the relation between this choice and the precise dimension definition one seeks to measure. Doing so, we illustrated that in a sense, length estimation of scale dependent sample paths (hence the regularization dimension estimation) is more accurate than energy measurement (hence the regularity dimension estimation).

Our second contribution was to carry on with the comparison between wavelet transform and empirical mode decomposition when both are used to estimate the regularization dimension. The conclusion is surprizing, and clearly shows a subtantial advantage of EMD when dealing with short size signals, for which its computational cost is still not too penalizing. This encouraging aspect, along with the adaptivity of EMD, prompts its use on processes with time varying regularity, and for which a local approach is necessary. Empirical mode decomposition should naturally accommodate the scales that exists in data and thus yield better estimates of the fractal dimensions. This is under current investigations.

7. REFERENCES

[1] P. Abry, P. Gonçalves, and J. Lvy-Vhel, *Lois d'Échelle, Fractals et Ondelettes*, vol. 1 and 2 of *Information - Commande - Communication*, Herms Sciences Publications, 2002.

[2] G. Rilling, P. Flandrin, and P. Gonçalves, "Empirical mode decomposition, fractional gaussian noise and hurst exponent estimation," in *ICASSP*, 2005.

[3] J. L. Lévy Véhel and F. Roueff, "A regularization approach to fractional dimension estimation," in *proceedings of Fractals 98*, Malta, 1998.

[4] C. Tricot, *Courbes et dimension fractale*, Springer-Verlag, 1999.

[5] N.E. Huang, Z. Shen, S.R. Long, M.L. Wu, H.H. Shih, Q. Zheng, N.C. Yen, C.C. Tung, and H.H. Liu, "The Empirical Mode Decomposition and Hilbert spectrum for nonlinear and non-stationary time series analysis," *Proc. Roy. Soc. London A*, vol. 454, pp. 903–995, 1998.

[6] P. Flandrin, G. Rilling, and P. Gonçalvès, "Empirical Mode Decomposition as a filter bank," *IEEE Sig. Proc. Lett.*, vol. 11, no. 2, pp. 112–114, 2004.

[7] S. Jaffard, "Wavelet techniques in multifractal analysis," in *Fractal Geometry and Applications: A Jubilee of Benoît Mandelbrot, M. Lapidus et M. van Frankenhuijsen Eds., Proceedings of Symposia in Pure Mathematics*. 2004, vol. 72(2), pp. 91–152, AMS.

[8] P. Abry, P. Gonçalves, and P. Flandrin, "Wavelets, spectrum analysis and 1/f processes," *Wavelets and statistics*, 1995, Lecture Notes in Statistics.

[9] A.T. Wood and G. Chan, "Simulation of stationary processes in $[0, 1]^d$," *J. Comp. Graph. Stat.*, vol. 3, pp. 409–432, 1994.

[10] D. Percival and A. Walden, *Wavelet methods for time series analysis*, Cambridge series in statistical and probabilistic mathematics. Cambridge Univ. Press, 2000.