

CROSSING-TREE PARTITION FUNCTIONS

*Geoffrey Decrouez**

National Research University
Higher School of Economics
Moscow, Russia

Pierre-Olivier Amblard

GIPSAIab/CNRS UMR 5283
Université de Grenoble
Grenoble, France

ABSTRACT

A new multifractal formalism based on the crossing-tree for H -sssi processes was recently introduced [1, 2]. The crossing-tree performs an ad-hoc decomposition of a signal based on its fluctuations, and thus represents a natural tool for the multifractal analysis of time series. The estimation of the Hausdorff spectrum happens in the context of a multifractal formalism, where the spectrum is obtained from a transform of a partition function. In this contribution, we introduce a new crossing-tree partition function, which differs from the original one presented in [2]. We show numerically that the new partition function improves the stability of the estimation in many cases, compared with the original crossing-tree partition function. Estimation is further compared with state-of-the-art techniques, including wavelet and wavelet leaders.

Index Terms— H -sssi processes, crossing tree, multifractal formalism, adaptative decomposition, wavelets

1. INTRODUCTION

Motivations. Fractal and multifractal signals often occur in so-called complex systems (systems with a huge number of degrees of freedom in nonlinear interaction) as different as turbulent fluids, stock markets or the internet. An important concept underlying these signals is scale invariance: the laws that underpin the construction of the processes are the same at all scales. This has naturally led to the use of multiscale tools to analyse such signals, among which the wavelet machinery is the best known, and provide today's state-of-the art techniques for fractal and multifractal signal analysis.

In the wavelet paradigm, the notion of scale is introduced in the index set (time for signals, space for images). The notion of scale can also be defined in the amplitude space of the signal. The crossing-tree decomposition then provides the corresponding multiscale decomposition: at a given resolution, the amplitude set is cut into equal length intervals, the bound of which define the crossing levels. A crossing is defined as the path of the signal between two consecutive strictly different crossing levels. At a given resolution, the nodes of

the tree contain the consecutive crossings. A multiscale representation is then provided by iterating the process on all crossings, by studying subcrossings at a finer resolution. To present the notation, we will formalize this later in this introduction. A main difference between wavelet decomposition and the crossing tree is the fact that the latter is an adaptive technique.

The crossing tree was introduced to generalize the mid-point displacement method to construct the standard Brownian motion. In [3] it was applied to self-similar processes, to test for self-similarity and stationary increments, and to obtain an asymptotically consistent estimator of the Hölder exponent. In [4] the crossing tree was used to estimate a time-change of a self-similar process, and in [5], it was used to characterize and test if a process is a continuous local martingale. Later, it has been the basis to construct a class of monofractal and multifractal processes, see [6, 7]. This lead us to invert the point of view and use the crossing tree as an analysis tool for fractal and multifractal signals. Indeed, the crossing tree is a very general concept and can easily be computed on real data.

In [1, 2] we showed how to use the crossing duration to estimate the Hölder exponent for several classes of monofractal signals. To do so we introduced a partition function using the crossing durations. A theoretical justification for it relies on [6, 7] where multifractal formalism is proved for the so-called multifractal embedded branching processes (MEBP, processes that are constructed *via* the crossing tree). As many methods, the partition function suffers a lack of statistics for negative powers. *The aim of the present paper is to illustrate how an easy modification of the partition function allows better estimation for negative powers.*

The remaining of the paper is organized as follows. We continue this introduction by presenting the crossing tree decomposition and the class of signals we study. Section 2 will present the partition functions we use to estimate the Hölder exponents, whereas section 3 will present the practical set up as well as simulations. One aim is to compare the crossing tree approach with the state-of-the-art provided by wavelet and wavelet leaders.

The crossing tree. We consider a process $X : \mathbb{R}^+ \rightarrow \mathbb{R}$. Without loss of generality, we assume $X(0) = 0$ almost

*Thanks to the Centre for Advanced Studies for funding.

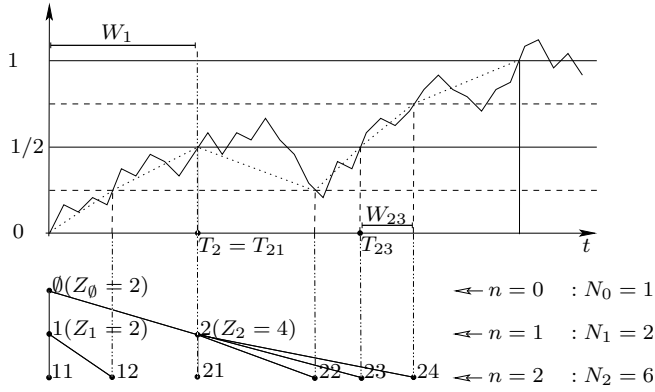


Fig. 1. Formation of the crossing tree from a sample path, and crossing tree notation. Variables are defined in the text.

surely (a.s.) and we further suppose it has continuous sample paths (a.s.). At a resolution $m \in \mathbb{Z}$, the vertical axis is cut into levels of size 2^m , and a crossing at level m is defined as the path between two successive different crossing levels. The first crossing starts at $t = 0$ and is defined as the path between $t = 0$ and the first time the signal crosses either 2^m or -2^m . Each crossing is characterized by two parameters: its duration and its direction. The level- m crossing times T_k^m are defined as

$$T_{k+1}^m = \inf\{t > T_k^m \mid X(t) \in 2^m\mathbb{Z}, X(t) \neq X(T_k^m)\},$$

where $T_0^m = 0$ and $2^m\mathbb{Z} = \{x \mid x = 2^m a \text{ for } a \in \mathbb{Z}\}$. Thus, the k -th level- m (equivalently scale 2^m) crossing $C_k^m := \{(t, X(t)) \mid T_{k-1}^m \leq t < T_k^m\}$ is the bit of sample path from T_{k-1}^m to T_k^m . Now the crossing tree is created by iterating the previous construction. Each crossing of size 2^m is decomposed into a sequence of crossings of size 2^{m-1} . The nodes of the crossing tree are crossings and the offspring of any given crossing is the corresponding set of subcrossings at the level below. An example of a crossing tree and the main notation are given in Figure 1. The crossing-tree can easily be computed for irregularly time-sampled signals, and as such the formalism developed later is adapted to this kind of data.

To code nodes in the tree we use the address space $I = \cup_{k=0}^{\infty} \mathbb{N}^k$, where \mathbb{N}^k is the set of words of k integers and $\mathbb{N}^0 = \emptyset$. For the sake of simplicity, the root of the tree and first crossing is supposed to go from 0 to ± 1 . It is labelled \emptyset and its subcrossings (each of size $1/2$) are numbered from 1 to Z_0 . The subcrossings of a crossing $\mathbf{i} = i_1 i_2 \dots i_n \in \mathbb{N}^n$ are then labelled $\mathbf{i}1, \dots, \mathbf{i}Z_i$, where Z_i is the number of subcrossings of \mathbf{i} and $\mathbf{i}j = i_1 i_2 \dots i_n j$. It is easy to show that Z_i is an even integer larger or equal to 2. We note as N_n the size of generation n . The second main parameter of a crossing \mathbf{i} is its type, either up or down, which we denote by $\sigma_{\mathbf{i}}$. The other parameter $W_{\mathbf{i}}$ is the duration of crossing \mathbf{i} . The sample path is completely described by $\{(\sigma_{\mathbf{i}}, W_{\mathbf{i}}) : \mathbf{i} \in I\}$.

Some scale invariant processes. $X(t)$ is said to be a self-similar process if there exists an $H \in (0, 1)$ such that $X(ct) = c^H X(t)$ holds for all $c > 0$ (in the finite-dimensional distribution sense). If in addition the process $X(t)$ has stationary increments, then $X(t)$ is said to be H -sssi.

The most-studied H -sssi processes are fractional Brownian motions (fBm), the only self-similar Gaussian processes with stationary increments. Their sample paths are continuous (a.s.) but non differentiable (a.s.). However, they have a degree of regularity and possess H as Hölder exponent almost everywhere. A nonGaussian generalization of fBm is provided by Hermite processes.

Let $\mathcal{B}(u)$ denote a Brownian motion. A Hermite process of order k is defined as

$$\mathcal{H}_H^k(t) = \int_{\mathbb{R}^k} \int_0^t \left(\prod_{j=1}^k (s - u_j)_+^{-(1/2+(1-H)/k)} \right) ds d\mathcal{B}(\mathbf{u}),$$

where $d\mathcal{B}(\mathbf{u}) = d\mathcal{B}(u_1) \dots d\mathcal{B}(u_k)$, for $k \geq 1$, with $H \in (1/2, 1)$, and $x_+ = \max(0, x)$. The case $k = 1$ corresponds to the case of an fBm. $\mathcal{H}_H^2(t)$ is usually referred to as the Rosenblatt process.

The Weierstrass function is defined as

$$\mathcal{W}_H(t) = \sum_{k \in \mathbb{Z}} \lambda_0^{-kH} \left(\cos(\varphi_k) - \cos(2\pi \lambda_0^k t + \varphi_k) \right),$$

where H stands for the Hölder exponent and λ_0 is a fundamental harmonic. The definition is made to impose $\mathcal{W}_H(0) = 0$. The Weierstrass function exhibits discrete scale invariance (DSI), with $\mathcal{W}_H(\lambda_0 t) = \lambda_0^H \mathcal{W}_H(t)$, in distribution. We consider here a stochastic version of this function, obtained by choosing the phases $\{\varphi_k\}_{k \in \mathbb{Z}}$ as a sequence of i.i.d. variables uniformly distributed over $[0, 2\pi]$.

2. PARTITION FUNCTIONS

The estimation of the spectrum of singularities $D(h)$, defined as the Hausdorff dimension of the set of points with given Hölder regularity h , typically occurs in the context of the multifractal formalism. The formalism relates the spectrum to a partition function $\zeta(q)$ via the (Legendre) transform

$$D(h) = \inf_{q \in \mathbb{R}} \{1 - \zeta(q) + hq\}.$$

This section reviews wavelet-based partition functions first, before introducing crossing-tree partition functions.

2.1. Wavelet-based partition functions

We first recall the wavelet decomposition of a signal $X(t)$. Let ψ be the mother wavelet. Any square integrable signal can be decomposed as

$$X(t) = \sum_{n,k \in \mathbb{Z}} c_{n,k} \psi(2^n t - k),$$

where

$$c_{n,k} = 2^n \int X(t)\psi(2^n t - k)dt.$$

We denote by $\lambda_{n,k} = [k2^{-n}, (k+1)2^{-n})$ a dyadic cube at scale n , with $k \in \mathbb{Z}$. The wavelet-based structure function of X is defined from the q -th moment of the wavelet coefficients,

$$S_{wc}(q, n) = 2^{-n} \sum_k |c_{n,k}|^q,$$

where the sum is taken over all dyadic cubes $\lambda_{n,k}$ with non vanishing coefficients. The wavelet partition function is then

$$\zeta_{wc}(q) = \liminf_{n \rightarrow +\infty} \left(\frac{\log S_{wc}(q, n)}{\log 2^{-n}} \right), \quad q \in \mathbb{R}, \quad (1)$$

which leads to the multifractal formalism [8]

$$D_X(h) = \inf_{q \in \mathbb{R}} \{1 - \zeta_{wc}(q) + hq\}. \quad (2)$$

The wavelet-based partition function (1) is known to be unstable for negative exponents, corresponding to $q < 0$. Indeed, wavelet coefficients can be arbitrary small, and a small error in their estimation can be multiplied when raised to a negative power.

To address this issue (amongst others), the wavelet leaders formalism was introduced [9]. Put $3\lambda_{n,k} = \lambda_{n,k-1} \cup \lambda_{n,k} \cup \lambda_{n,k+1}$, which corresponds to the cube centered around $\lambda_{n,k}$, three times wider. The wavelet leaders $d_{n,k}$ of a bounded function $X(t)$ are defined as

$$d_{n,k} = \sup_{\{m,i \mid \lambda_{m,i} \subset 3\lambda_{n,k}\}} |c_{m,i}|.$$

It is then natural to introduce the wavelet leader structure function

$$S_{wl}(q, n) = 2^{-n} \sum_k |d_{n,k}|^q,$$

where the sum is taken over all non vanishing wavelet leader coefficients. The scaling function is

$$\zeta_{wl}(q) = \liminf_{n \rightarrow +\infty} \left(\frac{\log S_{wl}(q, n)}{\log 2^{-n}} \right), \quad q \in \mathbb{R}, \quad (3)$$

which leads to the multifractal formalism

$$D_X(h) = \inf_{q \in \mathbb{R}} \{1 - \zeta_{wl}(q) + hq\}. \quad (4)$$

2.2. Crossing-tree partition functions

A new formalism for the study of H -sssi processes was recently introduced [1, 2]. The formalism relates the spectrum of singularities to a partition function computed from multi resolution quantities obtained from the crossing-tree of a signal.

Given t , let $\mathbf{i} \in \mathbb{N}^\infty$ be such that for each n , the size 2^{-n} crossing that contains t is $\mathbf{i}|n$. Then, our analogue of the multiresolution quantity is the crossing duration $W_{\mathbf{i}|n}$. We

say that the process $X(t)$ possesses scaling properties if time averages of the crossing durations follow a power law behaviour,

$$S_{ct}(n, q) = \frac{1}{N_n} \sum_{\mathbf{i}|n} |W_{\mathbf{i}|n}|^q \sim C'_q 2^{-n\zeta_{ct}(q)}, \quad (5)$$

as $n \rightarrow \infty$, where the sum is taken over all crossings of size 2^{-n} . We call $S_{ct}(n, q)$ the structure function and ζ_{ct} the crossing tree partition function. The partition function can be obtained from the structure function as a limit,

$$\zeta_{ct}(q) = \liminf_{n \rightarrow \infty} \frac{\log S_{ct}(n, q)}{-n \log 2}. \quad (6)$$

The crossing-tree partition function (6) was first introduced to obtain a theoretical expression of the Hausdorff spectrum of a class of processes called Multifractal Embedded Branching Processes (MEBP), see [6, 7]. An MBEP process X can be represented as the composition of a process Y with constant modulus of continuity, and the inverse of an increasing process \mathcal{M} , so that it can be written as $X = Y \circ \mathcal{M}^{-1}$. The increasing process \mathcal{M} is the integral of a multiplicative cascade defined on the boundary of the crossing tree of Y . The multifractal analysis of \mathcal{M} requires deducing the local Hölder exponent at every point of its support from a discretised version of it. To do so, the support of a multifractal measure is typically divided into dyadic cubes, and a structure function is defined from the empirical q -th moments of measures of cubes 3 times wider than the original partition, see e.g. Section 2 in [9]. This methodology was adapted in [7] to perform the multifractal analysis of \mathcal{M} . The analysis performed there requires the introduction of a novel partition function constructed on a random grid adapted to the process. The new partition function is defined in terms of the q -th moment of the sum of 3 consecutive crossing durations (instead of 3 consecutive dyadic cubes), and constitutes the starting point of the present study.

Let $\mathbf{i}|n-$ and $\mathbf{i}|n+$ denote respectively the left and right neighbours of $\mathbf{i}|n$. The corresponding crossing durations are denoted $W_{\mathbf{i}|n-}$ and $W_{\mathbf{i}|n+}$. The discussion above motivates the definition of a new crossing-tree structure function,

$$S_{ct}(n, q) = \frac{1}{N_n} \sum_{\mathbf{i}|n} |W_{\mathbf{i}|n-} + W_{\mathbf{i}|n} + W_{\mathbf{i}|n+}|^q, \quad (7)$$

and its associated partition function

$$\eta_{ct}(q) = \liminf_{n \rightarrow \infty} \frac{\log S_{ct}(n, q)}{-n \log 2}. \quad (8)$$

The introduction of the right and left neighbours in the definition of $S_{ct}(n, q)$ takes care of problems that may arise when the local Hölder exponent is discretised at time t , when t is a crossing time, since t corresponds in that case to one of the endpoints of the crossing durations.

Wavelet leaders were introduced to address the flaws of a multifractal analysis based directly on wavelet coefficients. In particular, wavelet leaders are known to estimate the spectrum of singularities with greater accuracy for $q < 0$. They are defined from the supremum of wavelet coefficients over dyadic cubes 3 times larger than the original dyadic partition, which is the key to their success. We thus expect that defining a structure function from the q -th moments of statistics defined on 3 consecutive intervals will also improve the stability properties of the structure function for $q < 0$, compared to the original $S_{ct}(n, q)$. The numerical work presented in the next section supports this claim.

Heuristic arguments presented in [2] lead to conjecture that for self-similar processes, the multifractal spectrum is related to the crossing-tree partition function via the transform

$$\zeta_{ct}(q) = \inf_h \{(q+1)/h - D(h)/h\}. \quad (9)$$

For monofractal processes with Hölder exponent H , one gets $\zeta_{ct}(q) = q/H$. A similar reasoning yields $\eta_{ct}(q) = q/H$ for this class of processes. This conjecture, proved for the Brownian motion, remains open. The numerical work presented in the next section supports this conjecture. The expression of the crossing-tree partition functions for monofractal processes should be compared with wavelet-based techniques, where we directly get from (2) and (4) that $\zeta_{wc}(q) = \zeta_{wl}(q) = qH$. The partition function is still linear in q , but the slope is inverted.

3. NUMERICAL WORK

For each process defined in the introduction, we compare the performance of the two crossing-tree partition functions. Estimation based on wavelet coefficients is also presented. The partition functions are estimated from an average of 1000 realizations of 2^{15} sample points each. The wavelet partitions functions are estimated using Daubechies' wavelets with 3 vanishing moments, from scale 2^3 to 2^{12} using Matlab routines from [10, 11]. The crossing tree partition function is estimated from scale 2^2 to 2^5 . The range of scales differs for wavelet methods and the crossing tree, since they are calculated differently: the scales chosen to analyse the crossing tree are computed from the spatial fluctuations of the signal.

We focus our attention to the estimation of the negative moments, which is the main objective of the present contribution. Figure 2 and 3 display the estimation of $\zeta_{ct}(q)$, $\eta_{ct}(q)$ and $\zeta_{wc}(q)$, $\zeta_{wl}(q)$ for q varying between -10 and 0. It is clear from Figure 3 that the partition function based on wavelet coefficients fails to work for negative q s, while the three other methods return stable results. In particular, it can be seen from Figure 2 that an estimation based on the crossing tree returns a good estimation for negative q s, and the estimation bias is further reduced when considering $\eta_{ct}(q)$ instead of $\zeta_{ct}(q)$. In most cases, the bias almost vanishes, while wavelet leaders return a small bias in some cases. The estimation was repeated

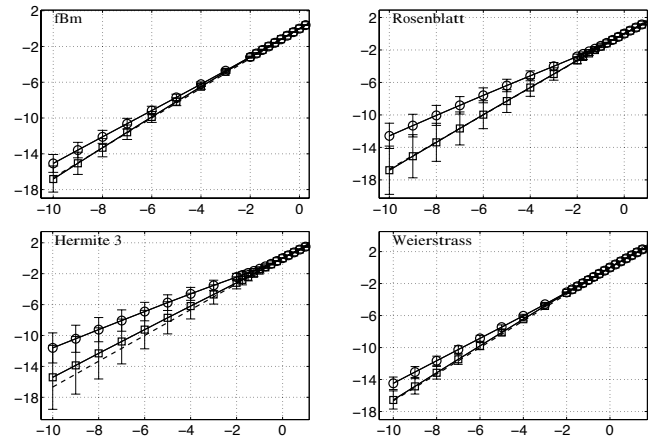


Fig. 2. Estimation of the partition function for q varying between -10 and 0 using $\zeta_{ct}(q)$ (\circ) and $\eta_{ct}(q)$ (\square). The dashed line is the theoretical line q/H . From top to bottom, left to right, fractional Brownian motion, Hermite with $k = 2$ and 3, and Weierstrass function, with $H = 0.6$.

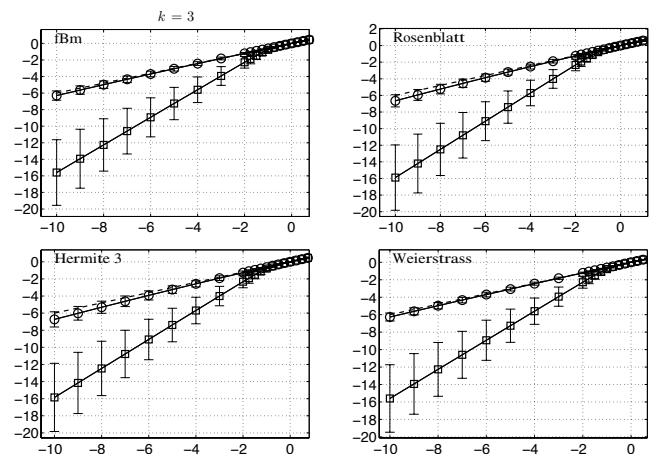


Fig. 3. Estimation of the partition function using wavelet leaders (\circ) and wavelet coefficients (\square).

for other values of $H > 0.5$, and the modified formalism systematically improves the estimation compared to the original crossing-tree partition function. Estimation for processes with $H < 0.5$ remains however challenging with the crossing tree, since the estimation of the crossing tree is all the more difficult as the process is rough (crossings are missed).

Following the seminal work of Castaing [12], we considered in [2] a polynomial expansion of the partition function, and we defined an estimator of H from the first coefficient in the expansion, estimated from the cumulant of the crossing durations. We repeated the same procedure for $\eta_{ct}(q)$, and we found that the performances of the estimators based on $\eta_{ct}(q)$ and $\zeta_{ct}(q)$ are comparable in most cases.

4. CONCLUDING REMARKS

The present contribution extends earlier work on the crossing-tree, by introducing a novel crossing-tree partition function. We showed numerically that a simple modification of the crossing-tree partition function presented in [1, 2] allows a more stable estimation for negative powers.

This work further raises many challenges (both theoretically and numerically) and open conjectures. The first step being formally proving that $\eta_{ct}(q) = q/H$ for monofractal H -sssi processes, before proving (9) in all generality. Estimation for rougher processes ($H < 0.5$) should also be further investigated, since the estimation of the crossing tree is deteriorated in this case. Finally, the crossing tree is defined so far only for one-dimensional processes. Extending its definition to two-dimensional processes might lead to a new tool to analyse images, and in particular to perform their multifractal analysis.

REFERENCES

- [1] G. Decrouez, P-O. Amblard, and O.D. Jones, "Estimation of the multifractal spectrum using the crossing tree," in *Proceedings of GRETSI, Brest, France*. Grets, 2013.
- [2] G. Decrouez and P-O. Amblard, "Analysis of h-sssi processes using the crossing tree: an alternative to wavelets," in *Proceedings of ICASSP, Brisbane, Australia*. ICASSP, 2015.
- [3] O.D. Jones and Y. Shen, "Estimating the hurst index of a self-similar process via the crossing tree," *Signal Processing Letters*, vol. 11, pp. 416–419, 2004.
- [4] W. Arroum and O.D. Jones, "A class of multifractal processes constructed using an embedded branching process," in *Complex Systems, A. Bender Ed. Article 60390F*. SPIE, 2006, vol. 6039.
- [5] O.D. Jones and D.A. Rolls, "A characterisation of, and hypothesis test for, continuous local martingales," *Electronic Comm. Prob.*, vol. 16, pp. 638–651, 2011.
- [6] G. Decrouez and O.D. Jones, "A class of multifractal processes constructed using an embedded branching process," *Annals of Applied Probability*, vol. 22, pp. 2357–2387, 2012.
- [7] G. Decrouez, B. Hambly, and O.D. Jones, "The hausdorff spectrum of a class of multifractal processes," *Stochastic Processes and their Applications.*, vol. 125, pp. 1541–1568, 2015.
- [8] G. Parisi and U. Frisch, "On the singularity structure of fully developed turbulence; appendix to," in *Fully developed turbulence and intermittency, by U. Frisch*. Proc. Int. Summer school Phys. Enrico Fermi, 1985, pp. 84–88.
- [9] S. Jaffard, B. Lashermes, and P. Abry, "Wavelet leaders in multifractal analysis," in *Applied and Numerical Harmonic Analysis*, pp. 219–264. Birkhauser Verlag Basel/Switzerland, 2006.
- [10] H. Wendt, P. Abry, and S. Jaffard, "Bootstrap for empirical multifractal analysis," *IEEE Signal Proc. Mag.*, vol. 24, no. 4, pp. 38–48, 2007.
- [11] H. Wendt and P. Abry, "Bootstrap tests for the time constancy of multifractal attributes," in *Proc. IEEE Int. Conf. Acoust., Speech, and Signal Proc. (ICASSP)*, Las Vegas, USA, 2008.
- [12] B. Castaing, Y. Gagne, and M. Marchand, "Log-similarity for turbulent flows," *Physica D*, vol. 68, pp. 387–400, 1993.