# Estimation of the multifractal spectrum using the crossing tree

Geoffrey DECROUEZ[1], Pierre-Olivier AMBLARD[1,2] *, Owen JONES[1]

[1]Dept. Mathematics & Statistics, The University of Melbourne
Parkville VIC 3010, Australie.

[2]GIPSA-lab, DIS, UMR CNRS 5216
ENSE3-BP 46 38402 Saint Martin d'Hères Cedex

dgg@unimelb.edu.au,bidou.amblard@gipsa-lab.inpg.fr,odjones@unimelb.edu.au

**Résumé –** L'arbre de branchement d'un signal a été récemment utilisé pour réaliser la synthèse de processus fractals et multifractals. L'objet de cet article est de le considérer comme un outil d'analyse. Etant confronté à un signal *a priori* multifractal, son arbre de branchement est déterminé, et les caractéristiques de l'arbre sont utilisées pour estimer le spectre multifractal du signal. Précisément, à une profondeur donnée, une fonction de partition résumant les statistiques des durées entre branchement est calculée. Pour certaines classes de processus, cette fonction de partition est une loi de puissance de la résolution d'analyse, dont l'exposant est relié au spectre multifractal. Nous illustrons cette approche sur divers signaux dont le spectre est connu théoriquement.

**Abstract –** The crossing tree of a signal has been recently used to construct monofractal and multifractal processes. In this paper, we use the crossing tree of a signal for the purpose of analysis. Given a realisation of a signal, its crossing tree is calculated, whose characteristics are used to estimate the multifractal spectrum. More precisely, a partition function is evaluated at each resolution to sum up the statistics of the crossing durations. For multifractal signals, the partition function is a power law of the resolution, whose exponent is linked to the spectrum. We illustrate numerically this approach for several classes of multifractal signals, including multiplicative cascades, fractional Brownian motion, and the Weierstrass function.

## 1 Motivation

Scale invariance and multifractal theory have been widely used over the past twenty years to analyse and describe data collected in a broad range of fields, including the study of natural phenomena in physics (in particular hydrodynamic turbulence), in computer network traffic, in financial markets for modelling volatility, but also in signal and image processing, to cite but a few.

Estimation of the multifractal spectrum of a signal $X$ typically happens in the context of the so-called "multifractal formalism". Let $T_X(a,t)$ summarize the spatial displacement of $X$ at time $t$ and at a temporal scale $a$. It is obtained from a comparison of the original process with a reference pattern $\psi$ dilated and located at different positions,

$$T_X(a,t) = a^{-1}\int X(u)\psi((u-t)/a)du\,.$$

The process $X$ is said to possess scaling properties if the time averages of $T_X(a,t_k)$ follow a power law behaviour with respect to $a$,

$$N_a^{-1}\sum_{k=1}^{N_a}|T_X(a,t_k)|^\theta \sim C_\theta a^{\zeta(\theta)} \text{ as } a\to 0\,,$$

where $N_a$ is the number of $T_X(a,t_k)$ available at scale $a$. Here $\zeta(\theta)$ is the partition function. In practice scaling properties are only observed for a limited range of scales and a limited range of $\theta$. The choice of $\psi$ plays a central role in the estimation of the partition function. Multiresolution quantities based on a wavelet decomposition of the process are the most common tool to date. For example when wavelet leaders are used we have that, for some processes at least, the set of points with Hölder exponent $\alpha$ has Hausdorff dimension $D(\alpha) = 1 + \inf_\theta(\theta\alpha - \zeta(\theta)) = 1+\zeta^*(\alpha)$, where $\zeta^*$ denotes the Legendre-Fenchel transform of $\zeta$, see [8]. More generally the RHS provides an upper bound on the multifractal spectrum. If we can write the multifractal spectrum in terms of the Legendre-Fenchel transform of $\zeta$ like this, then we say that the multifractal formalism holds. In the present study we introduce a novel set of multiresolution quantities, defined in terms of the crossing tree, which in our context can be viewed as a path-adapted multiresolution decomposition of the process.

The crossing tree is a very general concept. In [11, 12] it was applied to self-similar processes, to test for self-similarity and stationary increments, and to obtain an asymptotically consistent estimator of the Hurst index, which was shown to be an improvement on existing estimators in certain circumstances. In [1] the crossing tree was used to estimate a time-change of a self-similar process, and more generally in [10], was used to characterise and test if a process is a continuous local martingale. In [9, 3, 4, 5, 6] it was used to construct processes with scaling properties (monofractal and multifractal). *The work described here is a first attempt to use the crossing tree as an analysis tool for multifractal processes and measures.* The multiresolution quantities defined from the crossing tree are used to construct a new partition function, called the crossing-tree partition function, and to introduce a new multifractal formalism. In particular, we show that for a class of processes including multiplicative cascades, and for processes obtained as a time-change of a process of constant modulus of continuity, the Hausdorff spectrum can be written in terms of a transform of the crossing-tree partition function. In this case, the new multifractal for-

*P.O.A. est soutenu par une bourse Marie Curie International Outgoing Fellowship de l'Union Européenne.

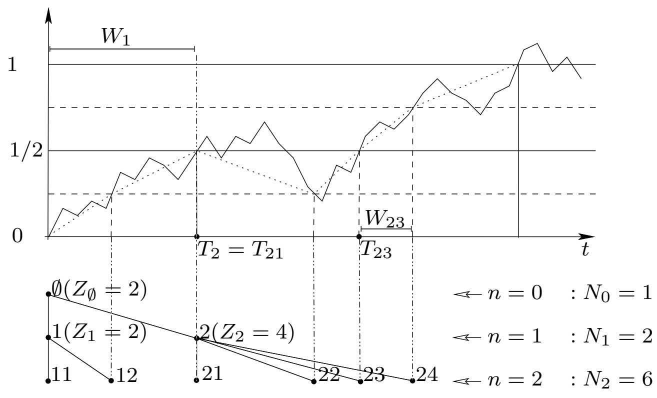


FIG. 1: Formation of the crossing tree from a sample path, and crossing tree notation. Variables are defined in the text.

malism is said to hold.

In Section 2 we define the crossing tree of a signal, we introduce the crossing-tree partition function and we describe the multifractal formalism. Section 3 illustrates numerically the formalism introduced in Section 2 for multiplicative cascades, fractional Brownian motion and the Weierstrass function.

# 2 The crossing tree partition function

We describe the crossing tree in the context of multifractal processes. Let $X : \mathbb{R}^+ \to \mathbb{R}$ be a (multifractal) process, with a.s. continuous sample paths and $X(0) = 0$. For $m \in \mathbb{Z}$ we define level-$m$ crossing times $T_k^m$ by putting $T_0^m = 0$ and $T_{k+1}^m = \inf\{t > T_k^m \mid X(t) \in 2^m\mathbb{Z},\ X(t) \neq X(T_k^m)\}$, where $2^m\mathbb{Z} = \{x \mid x = 2^m a \text{ for } a \in \mathbb{Z}\}$. The $k$-th level-$m$ (equivalently scale $2^m$) crossing $C_k^m := \{(t, X(t)) \mid T_{k-1}^m \leqslant t < T_k^m\}$ is the bit of sample path from $T_{k-1}^m$ to $T_k^m$. There is a natural tree structure to the crossings, as each crossing of size $2^m$ can be 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 are the corresponding set of subcrossings at the level below. An example of a crossing tree is given in Figure 1.

It is convenient to use the address space $I = \cup_{k=0}^\infty \mathbb{N}^k$, where $\mathbb{N}^k$ is the set of concatenations of $k$ integers and $\mathbb{N}^0 = \emptyset$, to label the crossings of the process. For simplicity we will consider the first crossing from 0 to $\pm 1$ and make this the root of our crossing tree. Label the root crossing $\emptyset$ and its subcrossings (each of size 1/2) 1 to $Z_\emptyset$. The subcrossings of a crossing $\mathbf{i} = i_1 i_2 \cdots i_n \in \mathbb{N}^n$ are then labelled $\mathbf{i}1, \ldots, \mathbf{i}Z_\mathbf{i}$, where $Z_\mathbf{i}$ is the number of subcrossings of $\mathbf{i}$ and $\mathbf{i}j = i_1 i_2 \cdots i_n j$. Necessarily $Z_\mathbf{i}$ is an even integer larger or equal to 2. Denote by $N_n$ the size of generation $n$. Each crossing $\mathbf{i}$ is one of two types, up or down, which we denote by $\sigma_\mathbf{i}$. Also let $W_\mathbf{i}$ be the duration of crossing $\mathbf{i}$, then the sample path is completely described by $\{(\sigma_\mathbf{i}, W_\mathbf{i}) : \mathbf{i} \in I\}$. Crossing-tree notation is summarized in Figure 1.

Let $\mathbf{i} \in \mathbb{N}^\infty$ be such that for each $n$, the size $2^{-n}$ crossing that contains $t$ is $\mathbf{i}|n$, where $\mathbf{i}|n$ is $\mathbf{i}$ truncated to $n$ places. Let $T_{\mathbf{i}|n}$ be the start time of crossing $\mathbf{i}|n$, then $T_{\mathbf{i}|n} \to t$ as $n \to \infty$, so formally we have

$$|X(t + W_{\mathbf{i}|n}) - X(t)| \approx 2^{-n} = W_{\mathbf{i}|n}^{-n \log 2 / \log W_{\mathbf{i}|n}} .$$

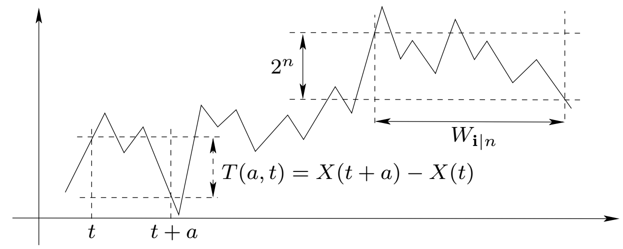


FIG. 2: Illustration of the difference between wavelet and crossing tree approaches for analysing a signal. For wavelets, local fluctuations are studied from a time decomposition, as $a \to 0$. For the crossing tree, the local behaviour is studied from a space decomposition, as $2^n \to 0$.

Thus (when everything works, for example for the Brownian motion, or more generally, for Canonical Embedded Branching Processes, see [6]) we get that $\alpha(t) = \lim_{n\to\infty} -n \log 2 / \log W_{\mathbf{i}|n}$, where

$$\alpha(t) = \liminf_{\epsilon \to 0} \frac{1}{\log \epsilon} \log \sup_{|u-t|<\epsilon} |X(u) - X(t)|$$

is the Hölder exponent of the process at time $t$. This equation gives the fundamental relationship between the multifractal spectrum and the crossing tree.

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 $T_X(2^{-n}, t)$ is just $W_{\mathbf{i}|n}$. As before, we say that the process $X$ possesses scaling properties if time averages of the crossing durations follow a power law behaviour. That is,

$$S(n, \theta) = \frac{1}{N_n} \sum_{\mathbf{i}|n} |W_{\mathbf{i}|n}|^\theta \sim C'_\theta 2^{-n\zeta_1(\theta)} , \tag{1}$$

as $n \to \infty$, where the sum is taken over all crossings of size $2^{-n}$. We call $S(n,\theta)$ the structure function and $\zeta_1$ the crossing tree partition function. The partition function can be obtained from the structure function as a limit,

$$\zeta_1(\theta) = \liminf_{n \to \infty} \frac{\log S(n,\theta)}{-n \log 2} . \tag{2}$$

The difference between partition function (2) and wavelet-based partition functions, is that it relies on an adaptive decomposition of the signal: classic methods rely on a time decomposition of the signal (usually using wavelets) whereas the proposed method relies on a space decomposition of the signal. This is illustrated in Figure 2.

Motivation for defining the crossing tree partition function (2) comes from the so-called Multifractal Embedded Branching Processes (MEBP), see [5]. MEBP constitute a large class of multifractal processes, which include random $m$-ary cascades and a class of time-changed Brownian motions. An MEBP process $X$ can be represented as the composition of a process $Y$ with constant modulus of continuity $H_Y$, 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 $Y$, where the weights of the multiplicative cascade are assumed to be i.i.d. The crossing tree of $Y$ is such that the $Z_\mathbf{i}$ are mutually independent and identically distributed, with mean $\mu$. Then

$H_Y = \log 2/\log \mu$, see [6], and

$$H_Y = \log 2/\log\left(\lim_{n\to\infty} N_n^{-1}\sum_{\mathbf{i}|n} Z_{\mathbf{i}|n}\right),$$

assuming the limit exists. It turns out that $H_Y\zeta_1$ is a natural partition function for $\mathcal{M}$ (since it is needed to divide $\log S(n,\theta)$ by $\log\mu$ in (2) instead of $\log 2$, see [4]), that the multifractal formalism holds for $\mathcal{M}$, see [7], and thus that $\mathcal{M}$ has multifractal spectrum

$$\begin{aligned} D_{\mathcal{M}}(\alpha) &= 1 + \inf_\theta(\alpha\theta - H_Y\zeta_1(\theta)) \\ &= 1 + H_Y \inf_\theta(\alpha\theta/H_Y - \zeta_1(\theta)) \\ &= 1 + H_Y\zeta_1^*(\alpha/H_Y). \end{aligned} \quad (3)$$

On the other hand, the spectrum of $\mathcal{M}^{-1}$ is given by $D_{\mathcal{M}^{-1}}(\alpha) = \alpha D_{\mathcal{M}}(1/\alpha)$, see [13, 14]. Composition with a process of constant modulus of continuity $H_Y$ transforms the spectrum as follows

$$D_X(\alpha) = D_{\mathcal{M}^{-1}}(\alpha/H_Y) = \alpha H_Y^{-1} + \alpha\zeta_1^*(1/\alpha). \quad (4)$$

Relation (4) constitutes the starting point of our multifractal analysis: it provides the exact expression of the Hausdorff spectrum of processes which can be written as the composition of a process with constant modulus of continuity and a multifractal time-change defined on its crossing tree. In a practical situation, we assume that the process under study belongs to this class, and we apply the methodology described.

# 3 Numerical study

We illustrate the methodology on two multifractal binary cascades, and on two monofractal processes: the fractional Brownian motion (fBm) and the Weierstrass function. The partition functions are estimated from an average of 1000 Monte-Carlo simulations, and errorbars correspond to plus and minus two standard deviations. Time series are of length $2^{18}$. In each case, the partition function $\zeta_1(\theta)$ is estimated from a linear regression of $\log_2 S(n,\theta)$ versus $n$. The range of scales for which the behaviour is linear is determined manually. This is illustrated in Figure 3. For all the simulations, the $\theta$ range is -10 to 10, even if such high exponents lead to poor estimation given the size of the signal considered.

**Deterministic Cascade.** The crossing tree of the binary cascade has exactly $\mu = 2$ offsprings at each node, giving $H_Y = 1$. The MEBP process reduces to $X = Y \circ \mathcal{M}^{-1} = \mathcal{M}^{-1}$, since in the previous notation, the process $Y$ corresponds to a single crossing from 0 to 1. By taking deterministic weights $W_0 = 0.25$ and $W_1 = 1 - W_0$, the crossing-tree partition function becomes

$$\zeta_1(\theta) = 1 - \log_2(W_0^\theta + W_1^\theta), \quad (5)$$

and is plotted on top left corner of Figure 4 (plain line), together with the crossing tree partition function (2), estimated over a range of scales from $2^{-1}$ to $2^{-16}$. The estimated $\zeta_1$ matches perfectly the true one. From (3), as $H_Y = 1$, the Legendre transform of $\zeta_1(\theta)$ gives the Hausdorff spectrum $D_{\mathcal{M}}(\alpha) = 1 + \zeta_1^*(\alpha)$ of $\mathcal{M}$, and that $\zeta_1$ coincides with wavelet-based partition function in this case.

**Random Cascade.** Take i.i.d. log normally distributed weights $W$, such that $\log|W|$ has mean $\mu$ and variance $\sigma^2$, so that the crossing-tree partition function is given by

$$\zeta_1(\theta) = -\frac{\sigma^2}{2\log 2}\theta^2 - \frac{\mu}{\log 2}\theta. \quad (6)$$

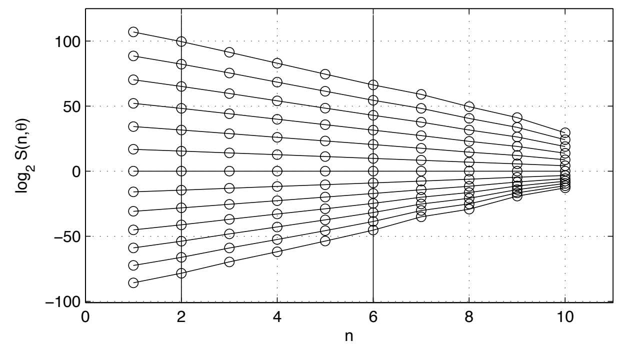


FIG. 3: Structure function in log-log plot, corresponding to one realisation of an fBm of length $2^{18}$, for $\theta$ varying from $-6$ to 6 (from top to bottom), and $n$ from 1 to 10. The crossing-tree partition function $\zeta_1(\theta)$ is estimated from a linear regression of $\log_2 S(n,\theta)$ versus $n$, for $n$ between 2 and 6.

The smallest $\alpha_m$ and largest $\alpha_M$ Hölder exponents are $\pm\sigma(2/\log 2)^{1/2} - \mu/\log 2$, see [2]. We take $\mu = -0.6\log(2)$ and $\sigma^2 = 0.05\log(2)$, which yields $\alpha_m = 0.28$ and $\alpha_M = 0.91$, corresponding to a rough process. The top right panel of Figure 4 presents $\zeta_1$ (plain line), whose expression is given in (6), together with the estimated one using the crossing tree partition function (2), where estimation is made across scales $2^{-1}$ to $2^{-12}$. There is a very good match between the two curves.

**fBm.** Consider now an fBm with $H = 0.7$. The analysis is performed across scales from $2^{-2}$ to $2^{-6}$. Results are presented in the bottom left panel of Figure 4. It is linear, as we would expect for monofractal processes, indicating that the method based on the crossing tree gives estimation consistent with the theory. The slope of the solid black line is $1/H$, where $H = 0.7$ is the Hölder exponent of the fBm. A linear regression of the estimated partition function versus $q$, gives an average estimate of $\hat{H} = 0.714$, with 95% confidence interval $[0.695,\ 0.734]$. Equation (4) then yields $D(\alpha = H) = 1$, $D(\alpha \neq H) = -\infty$, as expected.

**Weierstrass function.** We finally apply the method to the Weierstrass function, known to be a monofractal. It is defined as

$$f(t) = \sum_{k\in\mathbb{Z}} \lambda_0^{-kH}\Big(\cos(2\pi\varphi_k) - \cos(2\pi\lambda_0^k t + \varphi_k)\Big)$$

where $H$ stands for the Hölder exponent and $\lambda_0$ is a fundamental harmonic. We consider here a random version of the function, obtained by choosing the phases $\{\varphi_k\}_{k\in\mathbb{Z}}$ as a sequence of i.i.d. variables uniformly distributed over $[0,\ 2\pi]$. Note that the definition is made to impose $f(0) = 0$. We applied the crossing tree decomposition to the function for $H = 0.55$, $\lambda_0 = 1.2$ (for a sampling frequency assumed to be 1). We represent the estimated $\zeta_1(\theta)$ in the bottom right corner of Figure 4, for a range of scales from $2^{-3}$ to $2^{-6}$. The slope of the solid black line is $1/H$, where $H = 0.55$. The average estimate of $H$ is $\hat{H} = 0.560$, and a 95% confidence interval for $H$ is $[0.527,\ 0.593]$. Here as well, the partition function is linear, consistent with the theory.

**Discussion.** Both for the fbm and the Weierstrass function, the slope of the estimated partition function is $1/H$, and not $H$, as one would expect from methods relying on wavelets. We conjecture that for a monofractal process, the crossing-tree partition function is

$$\zeta_1(\theta) = \theta/H.$$

From (4), this is true for monofractal MEBP processes, and maybe holds as well for a wider class of processes. It is shown in [4] that fBm can be well approximated with MEBP (approximation is exact only in the case of a Brownian motion, for which $H = 1/2$), and it is known that the Brownian motion is included in the class of MEBP. The conjecture is thus true for the Brownian motion.

The estimate is good for both positive and negative $\theta$, corresponding respectively to the increasing and decreasing parts of the spectrum for multifractal processes. Estimation of the partition function for negative $\theta$ is a well known difficult issue, and structure functions based on wavelet coefficients are not stable for negative $\theta$. Wavelet leaders overcome this problem. A numerical study comparing the two methods will follow.

In the numerical study of the random lognormal cascade, we noted the importance of the assumption of conservation of mass when cascading the weights. Specifically, if the weights have a mean different from $1/2$, then this introduces an error in the estimation of the crossing-tree partition function. This observation follows from the assumption $\mu EW = 1$ needed to define nondegenerated MEBP processes, see [7]. The definition of the structure function in (1) needs to be slightly modified by multiplying the crossing durations at scale $n$ by the factor $a^n$, where $a = 1/(\mu EW)$. Estimates for $EW$ can be easily constructed, but are not detailed here.

The estimation of the crossing-tree partition function is more challenging when the data have Hölder coefficients less than 0.5. For very irregular processes, the discrete nature of the data makes the estimation of the crossing tree difficult at fine scales, since many crossings are missed. The poor estimation of the crossing tree results in a poor estimation of the partition function. This was noticed in particular in our numerical study when working with fBm with $H < 0.5$, where estimation for both positive and negative values of $\theta$ does not provide a good match with the theory. This effect is also visible with random cascades, where the estimated partition function is in better agreement with the theoretical curve the larger the value of the smallest Hölder exponent.

To conclude, the use of the crossing tree looks promising, but many different issues as discussed above remain to be solved. Comparison with wavelet leaders has to be performed, but as a new approach, much work obviously is needed to allow a fair comparison. But an advantage can already be put forward: the crossing tree approach works for irregularly sampled data and can thus provide an alternative to wavelets in such a situation, *e.g.* in finance applications.

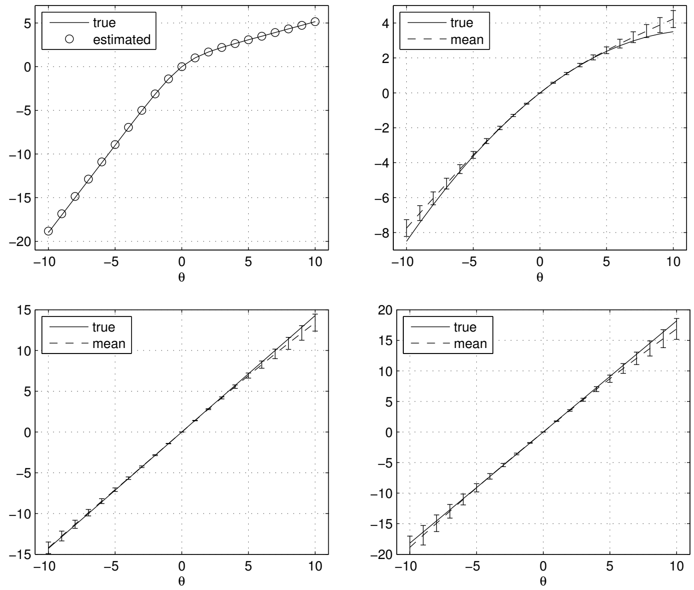


FIG. 4: Estimated and theoretical partition function for four processes. Top: deterministic cascade (left), random cascade (right). Bottom: fractional Brownian motion with $H = 0.7$ (left), and Weierstrass function with $H = 0.55$ (right). The plain lines correspond to $\zeta_1(\theta)$ for the top plots (equations (5) and (6), respectively), and to $\theta/H$ for the bottom plots.

# References


[1] ARROUM, W. AND JONES, O.D. 2006. Estimation for time-changed self-similar stochastic processes. Proceedings SPIE Vol. 6039, Complex Systems, A. Bender Ed. Article 60390F.

[2] ARNEODO, A., BACRY, E. AND MUZY, J.-F. 1998. Random cascades on wavelet dyadic trees. *J. Math Phys.*, 39(8):4142–4164.

[3] BURQ, Z. AND JONES, O.D. 2008. Simulation of Brownian motion at first-passage times. *Mathematics and Computers in Simulation* 77, pp. 64-71.

[4] DECROUEZ, G. 2009. Generation of multifractal signals with underlying branching structure. *PhD thesis.*

[5] DECROUEZ, G. AND JONES, O.D. 2012. A class of multifractal processes constructed using an embedded branching process. *Ann. of Appl. Prob. 22*, 2357–2387

[6] DECROUEZ, G., HAMBLY, B. AND JONES, O.D. 2012. The modulus of continuity of a class of monofractal processes. Available on Arxiv:1211.2895v2.

[7] DECROUEZ, G., HAMBLY, B. AND JONES, O.D. The Hausdorff spectrum of a class of multifractal processes. Submitted to the Electronic Journal of Probability.

[8] JAFFARD, S., LASHERMES, S., AND ABRY, P. 2006. Wavelet Leaders in Multifractal Analysis. *Applied and Numerical Harmonic Analysis. Birkhäuser Verlag Basel/Switzerland* pp 219–264.

[9] JONES, O.D. 2004. Fast, efficient on-line simulation of self-similar processes. *Thinking in patterns: Fractal and Related Phenomena in Nature*, M.M. Novak Ed., pp 165–176, World Scientific 2004.

[10] JONES, O.D., AND ROLLS, D.A. 2011. A characterisation of, and hypothesis test for, continuous local martingales. *Electronic Comm. Prob.* 16, pp 638–651.

[11] JONES, O.D., AND SHEN, Y. 2004. Estimating the Hurst index of a self-similar process via the crossing tree. *Signal Processing Letters* 11, pp 416–419.

[12] JONES, O.D., AND SHEN, Y. 2005. A non-parametric test for self-similarity and stationarity in network traffic. *Fractals and Engineering. New trends in theory and applications. J.Levy-Vehel and E. Lutton (Eds), Springer.*

[13] MANDELBROT, B.B. AND RIEDI, R.H. 1997 Inverse Measures, the Inversion Formula, and Discontinuous Multifractals *Advances in Applied Mathematics*, **18** 50–58

[14] RIEDI, R.H. AND MANDELBROT, B.B. 1997. Inversion Formula for Continuous Multifractals. *Advances in Applied Mathematics*, **19** 332–354.