Generation of multifractal signals with underlying branching structure

Geoffrey Decrouez

INSTITUT POLYTECHNIQUE DE GRENOBLE

N^o attribué par la bibliothèque
|__|__|__|__|__|__|__|__|__|__|

THESE EN COTUTELLE INTERNATIONALE

pour obtenir le grade de

DOCTEUR DE L'Institut polytechnique de Grenoble
et
de l'Université de Melbourne

Spécialité: Signal, Image, Parole, Télécommunications

préparée au laboratoire GIPSA-lab, Département Images et Signaux

dans le cadre de **l'Ecole Doctorale Electronique Electrotechnique, Automatique et Traitement du Signal**

et au Département des Mathématiques et Statistiques de l'Université de Melbourne

présentée et soutenue publiquement

par
Geoffrey DECROUEZ
le 12 janvier 2009

GENERATION DE SIGNAUX MULTIFRACTALS POSSEDANT UNE STRUCTURE DE BRANCHEMENT SOUS-JACENTE

DIRECTEURS DE THESE:
Pierre-Olivier AMBLARD/Owen Dafydd JONES
CO-DIRECTEURS DE THESE:
Jean-Marc BROSSIER/Konstantin BOROVKOV

JURY

M. Stéphane JAFFARD,	Président
M. Patrice ABRY,	Rapporteur
M. Benjamin Michael HAMBLY,	Rapporteur
MM. Pierre-Olivier AMBLARD/Owen Dafydd JONES,	Directeurs de thèse
MM. Jean-Marc BROSSIER/Konstantin BOROVKOV,	Co-encadrants

Acknowledgements

Firstly, I would like to thank my French and Australian supervisors and universities for making this PhD in cotutelle possible. Thanks Owen for your availability, at anytime. I enjoyed our discussions, and I greatly appreciate your help for making my arrival in Melbourne so easy. Thanks for your helpful suggestions for the writing of my thesis. It was a real pleasure working with you, always in a friendly environment.

Thank you to my supervisor and *mate* Bidou for giving me so much freedom during my PhD, without it, this cotutelle program would have never happened. Also, it was good having you in Australia last year, our discussions on my work with Owen on this occasion were very helpful. Thank you Jean-Marc for the long conversations we had on the phone, for your availability, even when it was summer holidays in France. Your advice for the writing of this thesis was always very helpful.

I would like to thank everyone who has helped me directly or indirectly in the content of my thesis. In particular, thank you to Darryl Veitch for allocating some of your time to talk about my research. Thanks also to David Rolls for the many discussions we had. I would also like to express my gratitude to Patrice Abry for providing me with Matlab code, which has been very helpful for the numerical results obtained in this thesis. Finally, thanks to Kostya Borovkov and Aihua Xia for being part of the committee for my confirmation talk at Melbourne University.

Then, I want to thank all my wonderful Australian friends who have helped me correct my English in this thesis. In particular, a big thank you to Tom Ballard, Ellie Button, Andrew Downes, Felix Ellet, Matthew Green, Paul Keeler, Heather Lonsdale, Rohan Malhotra, Anthony Mays, Jasper Montana, Michael Wheeler and Matt Zuparic for their suggestions. Cheers to my officemates, without them I wouldn't know that the capital city of Lesotho is Maseru and I still couldn't place Swaziland on a map. Thanks also to Andrew for reading part of my thesis while we spent hours and hours in Singapore waiting for passport controls to go to Malaysia. What a day. I also want to thank my friend Jasper for his understanding and his support throughout my thesis. Cheers to my great friend Prune for providing me a shelter and a mattress everytime I go back to Grenoble. Her place has become the most friendly squat in the whole city. I also want to give a big thank you to my parents, for making my exile at the other side of the world possible, but also for you support, and your wonderful help in all the horrible administrative tasks which makes this cotutelle possible. I promise I will go back to France ... one day.

Last, but not least, I want give a warm thank you to Daniel Baudois for transmitting to me his passion for signal processing. Without you as a lecturer, I would probably not have pursued my studies in this area.

Contents

List of Figures

List of Symbols

Chapter 1

Introduction

> "It is not necessary that you leave the house. Remain at your table and listen. Do not even listen, only wait. Do not even wait, be wholly still and alone. The world will present itself to you for its unmasking, it can do no other, in ecstasy it will writhe at your feet."
>
> Franz Kafka

Throughout the centuries, mathematicians have attached lots of credit to objects with nice properties, like spheres, lines, circles or differentiable curves and functions. Other mathematical objects were considered as "pathological", like irregular sets with no derivative almost everywhere. In a letter addressed to Emile Bernard in April 15th 1904, Paul Cézanne said: "Let me repeat what I told you when we were here: render nature with the cylinder, the sphere and the cone, arranged in perspective so that each side of an object or of a plane is directed towards a central point" [38]. For him, it was essential to learn how to paint "with reference to these simple shapes". Almost a century later, Mandelbrot claimed "clouds are not spheres, mountains are not cones, coastlines are not circle, and bark is not smooth, nor does lightning travel in straight line" [85]. The birth of the fractal geometry rests on Mandelbrot's observations of the world. His merit was to put these observations together with irregular objects scientists were considering as purely mathematical objects, to notice their shared properties and unify them to create what he called a fractal geometry. The term fractal, as defined by Mandelbrot, comes from the Latin *fractus* meaning broken and refers to very irregular sets. The rapid development of fractal geometry was soon recognized in many areas of science, as witnessed by the exponential growth in the number of papers on fractals during the past 20 years.

In this chapter, I recall the concept of a 'physical fractal' and strict self-similarity on sets through different famous examples. This will lead us to see how this concept can be adapted to the context of processes, which is the main focus of my research work. Self-similar processes whose graphs have a non integer dimension can be obtained through simple recursive procedures. Fixed points of Iterated Functions Systems (IFS) are one of them and are briefly presented later. We also present the class of random self-similar processes through different examples, including Brownian motion and fractional Brownian motions. Then we introduce multifractal processes

and review a few methods to estimate their properties. Finally, I present a general outline of my doctoral research work.

1.1 Fractals, multifractals and other oddities

1.1.1 A first glimpse into self-similarity

Traditional geometry cannot describe very accurately irregular objects of our world, for example a coastline, a mountain or a tree. Before the birth of fractal geometry, complicated geometrical shapes could only be represented by a map or by their images. When looking at a map of scale $1 : 100,000$, is it possible to measure the length of a coastline? What becomes this length if we now look at a more detailed map of scale $1 : 10,000$? As we zoom in, bays appear, increasing the total length of the coast. When seen at different scales, the coastline appears similar in some sense. Likewise for a mountain; a detail of the edge of a mountain looks like the whole mountain. One aim of mathematics is to provide approximate models to describe these objects from the real world. We propose to review famous examples which incorporate the property of *similarity* of their geometrical shape at different scales of magnification.

Let us start exploring this mathematical world with the famous Cantor ternary set (1875), a subset of the real line which contains no intervals but has as many points as an interval. This set, categorized as pathological and called 'monster' by the mathematician Charles Hermitte, can be obtained through a recursive procedure. Begin with the interval $C_1 = [0, 1]$ and remove its middle third. At the first stage, you obtain a set C_2 which is the union of two intervals: $C_2 = [0, 1/3] \cup [2/3, 1]$. Repeat the same procedure on the two smaller intervals. Keep going infinitely many times. The limit set, $C = \bigcap_n C_n$ is the Cantor ternary set, an uncountable set with zero Lebesgue measure. Parts of C look like the original set C, up to a scaling factor.

The journey continues with the Sierpinski gasket constructed by the Polish mathematician W. Sierpinski. The idea is similar: start with a single region and remove parts of its interior. Let the initial set be a triangle S_0 together with its interior. Then, connect the midpoints of the sides with line segments and remove the interior of the small middle triangular region. At the first stage, three smaller triangles replace the initial one. Call this set S_1. Repeat the same procedure to the three smaller triangles and continue infinitely many times. The limit set $S = \bigcap_n S_n$ is the Sierpinski triangle. An approximation of S is presented in the left upper corner of Figure 1.1, together with other limit sets obtained via similar recursive procedures. The sets are obtained with Matlab code that I have re-implented to produce them. The Sierpinski triangle possesses a similar property as the Cantor ternary set: each of the smaller triangles is an exact replica of S. Many other such examples exist, such as the Von Koch curve (a compact curve with infinite length), the Monger sponge, the Sierpinski carpet, to name but a few. In general, a figure or set which can be decomposed into parts which are exact replicas of the whole is called discrete strict self-similar.

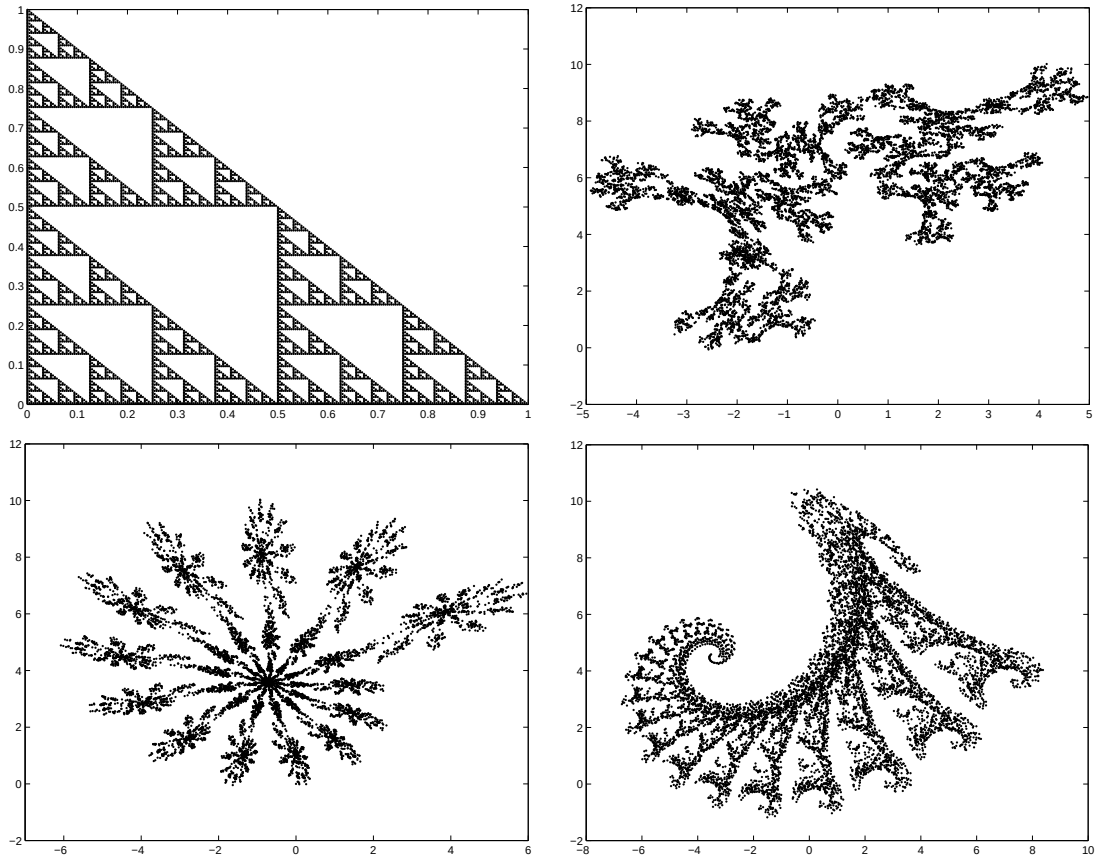
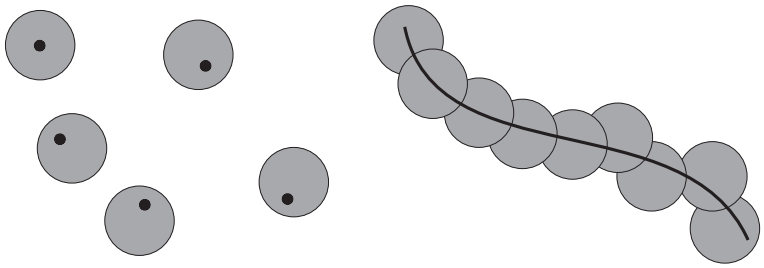


Figure 1.1: Self-similar sets.

Let us emphasize the major difference between fractals encountered in mathematics, for which self-similarity occurs at all scales of magnification, and self-similar objects found in nature, called 'physical fractals', for which this self-similarity holds only for a range of scales. When looking at a coastline, for example, one cannot zoom indefinitely but only up to the atomic level. Mathematical fractals have played an important role in synthetic imagery and graphics [18] and have touched all branches of science. Applications range from biology (plant trees, bronchial trees, blood circulation systems) to network traffic [2], from art (in decoration, architecture, [108]) to pure mathematics.

1.1.2 Various definitions of dimensions

The next questions that come to mind are how to measure the irregularity of a set or a curve and how 'big' is a fractal set? Are the fractals presented in Figure 1.1 similar in some way? Also, we have not yet given an accurate definition of a fractal, other than the fact that it is an irregular set. Is it possible to give a more precise definition? A possible answer to these questions is related to the concept of dimension, which gives a good quantitative measure of how much a set or curve fills the space. The definition of dimension is not unique and throughout the history of mathematics, many definitions have been given. For example, the Euclidean dimension D_E is the number of coordinates needed to address the object. The topological dimension D_T requires coverings of the object. A covering is by definition a collection of open sets in a topological space X whose union contains the object. We also define a refinement of a cover $\mathcal{C}$ of X as a new cover $\mathcal{C}'$ such that each set in $\mathcal{C}'$ is contained in some set in $\mathcal{C}$. Then, an object A has topological dimension D_T if every covering $\mathcal{C}$ of A possesses a refinement $\mathcal{C}'$ in which every point of A belongs to at most D_T+1 sets in $\mathcal{C}'$, where D_T is the smallest such integer. For example, as illustrated below, a set of points distributed in the plane have $D_T = 0$ and $D_E = 2$. A smooth curve in the plane has $D_T = 1$ and $D_E = 2$.



When dealing with irregular objects, other definitions of dimensions have to be used to compare them. These dimensions are referred to as fractional dimensions since they can take non integer values. The Hausdorff dimension $D_{\mathcal{H}}$ of a set is a central notion in fractal geometry. Fractional dimensions first appeared in 1919 with mathematician Felix Hausdorff but had to wait until the works of Mandelbrot to be related in a systematic way to fractals. Other fractional (or fractal) dimensions are the box counting and packing dimensions.

HAUSDORFF DIMENSION. The Hausdorff dimension of a set is defined from its Hausdorff measure. To motivate the definition of the Hausdorff measure, we address the following question: given a curve $c \subset \mathbb{R}^2$, how do we measure its length? We want some measure theory notion, independent of the properties of the curve, such as its differentiability. The idea is once again to cover c with balls B_j. Denote by $|B_j|$ the diameter of B_j. Then, an approximation of the length $L(c)$ of c is $\sum |B_j|$. However, the covering can have poorly placed balls or extra balls not needed to cover c. A better approximation of the length is then to take the infimum over all possible coverings:

$$L(c) \approx \inf \left\{ \sum |B_j| \mid c \subseteq \bigcup B_j \right\}.$$

A second problem arises. A ball can do too well because it is too large: the sum is therefore too small because the balls do not follow the contours of the curve. Therefore, we approximate the length of c as the sum of the diameter of a 'good' covering of c, in the limit where the diameter of the balls tends to 0. Formally, fix $\delta > 0$ and $A \subset \mathbb{R}^p$, then the previous discussion motivates us to define

$$\mathcal{H}^n_\delta(A) = \inf \left\{ \sum_{j=1}^{\infty} |B_j|^n \mid A \subseteq \bigcup_{j=1}^{\infty} B_j,\ |B_j| \leqslant \delta \right\}.$$

In the previous definition, we raise $|B_j|$ to the power of n by analogy with covering sets with balls in $\mathbb{R}^n$. Since the infimum is made over a smaller number of possible coverings as δ decreases, $\mathcal{H}^n_\delta(A)$ increases and it is legitimate to consider the limit as $\delta \to 0$:

$$\mathcal{H}^n(A) = \lim_{\delta \to 0^+} \mathcal{H}^n_\delta(A).$$

$\mathcal{H}^n(A)$ is the Hausdorff measure of A. In fact, one can show that as n increases $\mathcal{H}^n(A)$ jumps from ∞ to 0, and that there is at most one s with $0 < \mathcal{H}^s(A) < \infty$. It is always possible to define the biggest n for which $\mathcal{H}^n(A)$ is infinite, or equivalently the smallest n such that $\mathcal{H}^n(A) = 0$. This particular value of n is the Hausdorff dimension of A:

$$D_{\mathcal{H}}(A) = \sup \{n \mid \mathcal{H}^n(A) = \infty\} = \inf \{n \mid \mathcal{H}^n(A) = 0\}.$$

The corresponding Hausdorff measure of A is however not necessarily finite. In general, $D_{\mathcal{H}} \geqslant D_T$. In fact, one cannot give a precise definition of a fractal, but Falconer in [44] notices that most fractal sets K share the following properties:

1. K has details/irregularities at all scales.
2. K cannot be described using equations from classical geometry.
3. K is approximately, strictly or statistically self-similar.
4. The 'fractal' dimension of K is usually strictly greater than its topological dimension.
5. K can be constructed most of the time in a simple way, perhaps recursively.

BOX-COUNTING DIMENSION. Hausdorff dimension cannot be calculated directly in practice due to the infimum encountered in the definition. To this end, more

tractable definitions are needed. Box counting dimension dates back to the 1930's and is also known as Kolmogorov entropy, entropy dimension, or capacity dimension. Its widespread use is due mainly to its ease of calculation. The idea is to cover the object A with sets of diameter r. Call N_r the number of such sets needed to cover A. The box dimension is then

$$D_B(A) = \lim_{r \to 0} \frac{\log N_r}{-\log r}$$

if the limit converges (otherwise replace lim by lim inf or lim sup, respectively the lower and upper box counting dimensions). The box dimension is therefore the power law behaviour of the measurement of the object at scale r. The number of sets that can cover A is of order $r^{-D_B(A)}$. The previous definition remains the same if for N_r we consider the smallest number of cubes of diameter r that can cover A [44], hence the name box counting dimension. To obtain an estimate of $D_B(A)$, it suffices to plot $\log N_r$ versus $\log r$. The slope of a linear interpolation gives an estimation of $D_B(A)$.

For sufficiently smooth objects like a straight line, it is possible to have equality among the various definitions of dimension. However, in general this does not hold and $D_E \geqslant D_B \geqslant D_{\mathcal{H}} \geqslant D_T$. We now review a recursive procedure to create sets for which $D_E > D_{\mathcal{H}}$.

1.1.3 Iterated Function Systems (IFS)

IFS are a simple way to generate fractal objects. Hutchinson pioneered the theory of non random self-similar fractal sets and measures via contracting mapping methods in his early work in 1981 [60]. He also introduced the notion of scaling operators. The terminology Iterated Function Systems appeared later on with the works of Barnsley and Demko [15].

The idea of an IFS is to recursively apply a set of contractive operators on a given set. A contraction $\omega : \mathbb{R}^d \to \mathbb{R}^d$ is a mapping for which there exists a $c \in (0, 1)$ such that $\forall (x, y)$, $|\omega(x) - \omega(y)| \leqslant c|x - y|$. c is called the contraction factor of ω. An IFS consists of a collection of contractions $\{\omega_1, \ldots, \omega_M\}$ with $M \geqslant 2$, with contraction factor $c_1, \ldots, c_M$. Let $\mathcal{H}(\mathbb{R}^2)$ be the set of compact sets of $\mathbb{R}^2$ and define the operator W by:

$$W(A) = \bigcup_{i=1}^{M} \omega_i(A)$$

for all $A \in \mathcal{H}(\mathbb{R}^2)$. Then W is contractive with contraction factor $c = \max(c_i)$. One can prove the existence and uniqueness of an attractor A^* or fixed point of a given IFS under mild conditions. A^* satisfies

$$A^* = \bigcup_{i=1}^{M} \omega_i(A^*) = W(A^*). \tag{1.1}$$

The IFS is said to satisfy the *open set condition* if there exists a non empty open set $O \subseteq \mathbb{R}^n$ such that $\cup \omega_i(O) \subseteq O$, where $\omega_i(O)$ and $\omega_j(O)$ are disjoint if $i \neq j$. The

open set condition is strong enough to ensure good mathematical results (existence of a fixed point, properties of the fixed point, derivation of its Hausdorff dimension, etc.) but sufficiently weak to include a large number of examples. Under the open set condition, one can show that the Hausdorff dimension $s = D_{\mathcal{H}}(A)$ of the attractor A of the IFS satisfies [93]

$$\sum_{i=1}^{M} c_i^s = 1. \tag{1.2}$$

We have already encountered two attractors of IFS in this introduction. The Cantor ternary set C is the fixed point of an IFS with two maps $\omega_1(x) = x/3$ and $\omega_2(x) = x/3 + 2/3$ and satisfies $C = \omega_1(C) \cup \omega_2(C)$. From Equation (1.2), it follows that the Hausdorff dimension of C is $\log 2/\log 3$. The Sierpinski triangle is also the fixed point of an IFS, consisting of 3 similarities of ratio 1/2. Its Hausdorff dimension is $\log 3/\log 2$.

IFS received a great deal of interest in data compression. The target image is represented by the attractor of an IFS. The difficulty of the method is to obtain a set of contractive maps that approximate correctly the target image. When successful, the advantage is an enormous compression of information to encode the image.

Fractals encountered so far are all deterministic (IFS considered have non-random contraction mappings) and are strictly self-similar. It is possible to randomize the construction of fractals to break the strict self-similarity. The motivation to do so is that 'physical' fractals are statistically self-similar. In other words, when zooming in on a natural fractal, the detail has the same properties as the whole object without being exactly the same. Random fractals can also be easily obtained via random IFS when considering random maps ω_i. Results on fractal sets have been adapted to the study of random self-similar measures and functions. My research works were partly focused on IFS acting over the space of functions to produce self-similar random processes. The method used was based on previous results from Hutchinson and Rüschendorff [63] and generalized one of their results on the existence and uniqueness of the attractor of an IFS, when allowing more randomness in the model.

Sets encountered so far possess no characteristic space scale. To characterize them and give them a dimension, we had to look at the construction procedure and to understand how scales interact with each other. Motivated by this observation, we want a process exhibiting scale invariance not to have a characteristic time or scale, so that the whole signal or parts of it cannot be distinguished. One cannot use traditional techniques to study these processes. Instead, it is relevant to understand how properties of the process are related across scales. Dimension is one way of thinking about fractals, the other paradigm is scaling, better suited to signals. To illustrate this notion, we consider two famous processes as introductory examples: the Brownian motion and the fractional Brownian motion. Then, we discuss the notion of strict self-similarity and introduce models with a richer structure. In this thesis, a random function $X(t)$ will be called either process or signal. The first term is more widely used among mathematicians while the second term is more common in the signal processing community.

1.1.4 Brownian Motion and fBm

Brownian motion $B(t)$ was first observed by botanist Robert Brown in 1828 as he noticed the very irregular movement of pollen suspended in water [28]. The origin of the motion remained however unexplained and we had to wait until the works of Bachelier in stock price fluctuations in 1900 [11] and Einstein in 1905 [39] to obtain an explanation of the motion. Einstein predicted the motion of a sufficiently small particle caused by the random bombardment of the molecules of the liquid: the random number of collisions on the particle by molecules coming from different directions, with different strength, would cause an irregular motion of the particle. It turned out that the model gave a good description of the motion observed by Brown. A rigorous mathematical treatment of Brownian motion was given by Norbert Wiener in 1923. Applications of Brownian motion are in fact far beyond the study of microscopic particles floating in water. Its applications include modelling stock prices or thermal noise in electric circuits, but cover also random perturbations in many branches of science like physics, biology or economy. But Brownian motion has also played a major role in understanding fractals and self-similar processes. It can be constructed as a limit random walk in the plane or space. Limits of random walks on a Sierpinski gasket were also studied by Barlow and Perkins in 1988 [13]. They proposed to associate a branching process with the random walk (details about its construction are postponed to Chapter 3). This association was the primary motivation of the definition of Embedded Branching Processes (EBP) introduced by Jones in 2004 [69]. Part of my doctoral works consisted in extending EBP to a wider class of processes that we called Multifractal Embedded Branching Processes or MEBP.

Let us go back to the properties of $B(t)$. Provided $B(0) = 0$, it can be shown that the Brownian motion is the only process satisfying the following 3 properties
1. For all $\tau \geqslant 0$, the increment process $\Delta B_\tau(t) = B(t+\tau) - B(t)$ is Gaussian and stationary.
2. For $t_1 \leqslant t_2 \leqslant t_3 \leqslant t_4$, increments $B(t_4) - B(t_3)$ and $B(t_2) - B(t_1)$ are independent.
3. $B(t)$ is continuous.

$B(t)$ is the only Gaussian process with stationary and independent increments. $B(t)$ and $\Delta B_\tau(t)$ satisfy an interesting scaling law:

$$B(ct) \overset{d}{=} c^{1/2} B(t) \tag{1.3}$$

$$\Delta B_\tau(ct) \overset{d}{=} c^{1/2} \Delta B_\tau(t) \tag{1.4}$$

for all $c > 0$. Here $\overset{d}{=}$ denotes equality in distribution. Relation (1.3) suggests that the sample paths of Brownian motion cannot be distinguished from a rescaled version, by dilating the time axis by a factor c and the amplitude axis by a factor $c^{1/2}$: there is no reference scale of time. Furthermore, the Hausdorff and box dimensions of its graph are non integer and equal 1.5 almost surely. This observation builds a first bridge between the notions of scaling law and fractal dimensions.

For modelling purposes, Brownian motion suffers from its simplicity. It is unlikely to be able to model all random processes with a graph whose dimension is 1.5, or equivalently, whose scaling exponent is fixed to 1/2. To obtain more general models, at least one of the 3 properties given above need to be relaxed. We can drop for example either the independence of increments, leading to the fractional Brownian motion (fBm) or the finite variance of increments to obtain the class of α-stable Lévy processes. fBm $B_H(t)$ was introduced by Mandelbrot and Van Ness in 1968 [89] as a moving average of $dB(t)$, where past increments are convoluted by the kernel $(t-s)^{H-1/2}$, for $0 < H < 1$.

$$B_H(0) = 0$$

$$B_H(t) = \frac{1}{\Gamma(H+1/2)} \left[\int_{-\infty}^{0} \left(|t-s|^{H-1/2} - (-s)^{H-1/2} \right) dB(s) + \int_{0}^{t} |t-s|^{H-1/2} dB(s) \right].$$

In fact, fBm has the following properties

1. B_H is continuous and $B_H(0) = 0$.
2. $B_H(t+\tau) - B_H(t)$ is normally distributed with mean 0 and variance τ^{2H}.
3. B_H has stationary increments.

For $H = 1/2$, fBm reduces to Brownian motion. Under the assumption of finite variance, the covariance structure of fBm is fixed

$$\mathbb{E}B_H(t)B_H(s) = \frac{\mathbb{E}|B_H(1)|^2}{2}(|t|^{2H} + |s|^{2H} - |t-s|^{2H}).$$

A similar expression holds for increments of fBm (which are called fractional Gaussian noise (fGn)). The study of fGn can inform us about the behaviour of fBm. When $1/2 < H < 1$, fGn is positively correlated and the correlation function is not integrable (it decreases as the power law τ^{2H-2}). We are in presence of long-range dependence (LRD) or long memory [20]. These notions can be compared with more classical processes with exponential decay of their covariance function, like ARMA processes.

One can show that fBm with parameter H has Hausdorff and box dimensions $2-H$ almost surely. $B_H(t)$ and its increments $\Delta B_{H,\tau}(t) := B_H(t+\tau) - B_H(t)$ satisfy a similar scaling law to (1.3)

$$B_H(ct) \overset{d}{=} c^H B_H(t) \tag{1.5}$$

$$\Delta B_{H,\tau}(ct) \overset{d}{=} c^H \Delta B_{H,\tau}(t). \tag{1.6}$$

The behaviour and properties of fBm can be fully derived through its only parameter H. This simplicity is convenient for modelling but is also a drawback since fBm is too simple for many real world problems. It is unlikely that a whole range of scale invariant signals can be modeled by a class of processes with a single parameter. There exists other processes which generalise the Brownian motion. Stochastic differential equations are really the first extension to include a wider class of random processes.

1.1.5 Strict Self-Similarity

Relations (1.3) and (1.5) can be unified in the fundamental notion of self-similarity. A process X is said to be self-similar with index H if and only if a change of the time scale is equivalent to a change in the state space scale

$$\forall c > 0 \quad \{X(ct), t \in \mathbb{R}\} \stackrel{fdd}{=} \{c^H X(t), t \in \mathbb{R}\}. \tag{1.7}$$

Equality (1.7) holds in the sense of finite dimensional distributions ($\stackrel{fdd}{=}$), that is for any $d \geqslant 1$, $t_1, \ldots, t_d$,

$$(X(ct_1), \ldots, X(ct_d)) \stackrel{d}{=} (c^H X(t_1), \ldots, c^H X(t_d)).$$

A process satisfying (1.7) was originally called semi-stable by Lamperti in 1962 [77]. Mandelbrot used the term self-similar 20 years later [85]. It follows immediately from this definition that for all $t > 0$, the moments of X behave as a power law

$$\mathbb{E}|X(t)|^q = \mathbb{E}|X(1)|^q |t|^{qH} \tag{1.8}$$

for all q such that $\mathbb{E}|X(t)|^q$ is finite. Relation (1.8) shows that self-similar signals are non-stationary. There exists however a relation between stationary and self-similar processes, given by Lamperti's theorem [77].

1.1.6 H-SSSI processes

One usually restricts the class of self-similar processes to one of self-similar processes with stationary increments (SSSI processes) as they are more convenient to work with since their increments satisfy $Y_a(t) := X(t+a) - X(t) \stackrel{d}{=} X(a) - X(0)$. If one assumes finite variance, then the covariance structure of X is constrained to satisfy

$$\mathbb{E}X(t)X(s) = \frac{1}{2}[|t|^{2H} + |s|^{2H} - |t-s|^{2H}]\mathbb{E}|X(1)|^2.$$

For the covariance function to be definite non-negative, it follows that the range of possible values of H is $(0, 1]$. The unique H-SSSI Gaussian process is the fractional Brownian motion introduced above and is the most widely used to model phenomena possessing scale invariance properties. Also, the self-similarity is transmitted to Y_a:

$$\mathbb{E}|Y_a(t)|^q = \mathbb{E}|X(1)|^q |a|^{qH}. \tag{1.9}$$

We have already encountered this situation for fBm, where both the process and its (stationary) increments are self-similar. Note that in (1.8) and (1.9), the exponent is linear with q.

1.1.7 Hölder regularity and fractal processes

The concept of local Hölder regularity is closely related to the notion of self-similarity. Self-similar processes have parts (details) which are statistically similar to the whole. Zooming in on a detail of the process helps us learn about its the local fluctuations. Information about these local fluctuations can be made precise with the definition of the Hölder exponent of a process $X(t)$ at a specific time $t = t_0$. It compares $X(t_0)$ with a polynomial $P_{t_0}(t)$. The process X is said to belong to $C_{t_0}^{h(t_0)}$ if there is a polynomial P_{t_0} of degree at most equal to the integer part of $h(t_0)$ such that

$$|X(t) - P_{t_0}(t)| \leqslant K|t - t_0|^{h(t_0)}$$

in a neighborhood of t_0. The largest value H of $h(t_0)$ such that $X \in C_{t_0}^{h(t_0)}$ is the Hölder exponent of X at $t = t_0$. Alternatively, one says that the Hölder regularity of X at $t = t_0$ is H. Fractional Brownian motion $B_H(t)$ with self-similarity parameter H (see Equation 1.9) has a fixed local Hölder exponent equal to H for all t. It follows that almost surely, for all t, $|B_H(t+\delta) - B_H(t)|$ is bounded above by $K\delta^H$, for some finite constant K. We refer to processes with a constant Hölder exponent along their sample paths as monofractals. It is possible to consider a more general class of processes for which the Hölder exponent varies smoothly and deterministically with time. These processes are called multifractional. The major drawback is once again the lack of flexibility since the Hölder exponent is the same at a given time for all realisations of the process. In network traffic applications, the rich structure of the data does not allow multifractional processes as good models [1]. Instead, it is interesting to consider the case when $h(t)$ varies in an erratic way. The name multifractal was given to such processes, in opposition to monofractals with a single Hölder exponent. Due to the highly irregularity of $h(t)$, it is not realistic to describe the fluctuations of multifractal processes in terms of the evolution of h with time. Instead, we depict the process by means of its multifractal spectrum $D(h)$, a global description of its local fluctuations. $D(h)$ is defined as the Hausdorff dimension of the set of points with a given Hölder regularity h. For monofractal processes, $D(h)$ degenerates to a single point at $h = H$, $D(H) = 1$ and generally the convention is to set $D(h) = -\infty$ for $h \neq H$.

1.1.8 Multiplicative cascades

The oldest and best known multifractal processes are the multiplicative cascades introduced by Mandelbrot in the context of intermittent turbulence in 1974 [84]. We review the construction of the deterministic binomial measure μ, indicate some generalizations and give results about their local behaviour.

Deterministic example. Let the support of μ be the real interval $[0, 1)$. The idea is to allocate a mass or probability to each subinterval of $[0, 1)$ of the form

$$I_{n,k} = \left[\frac{k}{2^n}, \frac{k+1}{2^n}\right)$$

for $k = 0, \ldots, 2^n - 1$, $n = 0, 1, 2, \ldots$ Let $t \in [0, 1)$, the singleton $\{t\}$ can be expressed

as the intersection of all $I_{n,k}$ for which $t \in I_{n,k}$

$$\{t\} = \bigcap_{\substack{n \geqslant 0 \\ t \in I_{n,k}}} I_{n,k}.$$

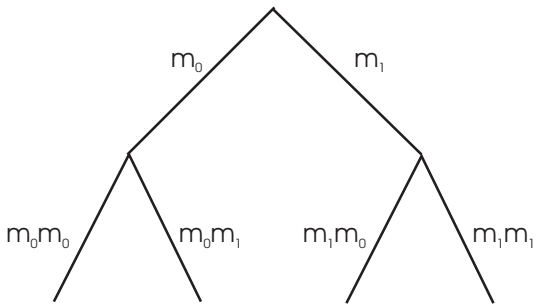
The binomial measure is constructed as the limit of a sequence of measures $\{\mu_n\}$. Let $\mu_0([0,1)) = 1$. Start by splitting $[0,1)$ into two subintervals of equal length $[0,1/2)$ and $[1/2,1)$. Allocate a mass m_0 to $[0,1/2)$ and m_1 to $[1/2,1)$, with $m_0 + m_1 = 1$,

$$\mu_1([0,1/2)) = m_0 \qquad \mu_1([1/2,1)) = m_1.$$

At stage $n = 2$, we divide each $[0,1/2)$ and $[1/2,1)$ into two subintervals of equal length to obtain intervals $[0,1/4)$, $[1/4,1/2)$, $[1/2,3/4)$ and $[3/4,1)$. $[0,1/4)$ receives a fraction m_0 of the mass of $[0,1/2)$ and $[1/4,1/2)$ a fraction m_1 of the mass of $[0,1/2)$. Thus

$$\mu_2([0,1/4)) = m_0^2 \qquad \mu_2([1/4,1/2)) = m_0 m_1$$
$$\mu_2([1/2,3/4)) = m_1 m_0 \qquad \mu_2([3/4,1)) = m_1^2.$$

The figure below illustrates the binary branching structure associated with this construction



At stage n, the initial mass $m_0 + m_1 = 1$ is distributed among 2^n dyadic intervals, which defines a measure μ_n, piecewise uniform. Consider the dyadic expansion of t

$$t = \sum_{i=1}^{n} \xi_i 2^{-i}$$

for $\xi_i \in \{0,1\}$. Let $N = \sum_{i=1}^{n} \xi_i$ be the number of ones among the first n binary digits of t, then, if $t \in I_{n,k}$,

$$\mu_n(I_{n,k}) = m_0^{n-N} m_1^N.$$

Since for all $m \geqslant n$, $\mu_m(I_{n,k}) = \mu_n(I_{n,k})$, we may define the binomial measure μ to be the limit of the sequence $\{\mu_n\}$ such that $\mu(I_{n,k}) = \mu_n(I_{n,k})$. Since $\mu([0,1)) = 1$ and μ is positive, the limit is well defined since (μ_n) is an increasing sequence, bounded above. The construction of μ can be extended to all half-open subintervals $[a,b)$ of $[0,1)$. By Carathéodory's extension theorem, we can uniquely extend μ to

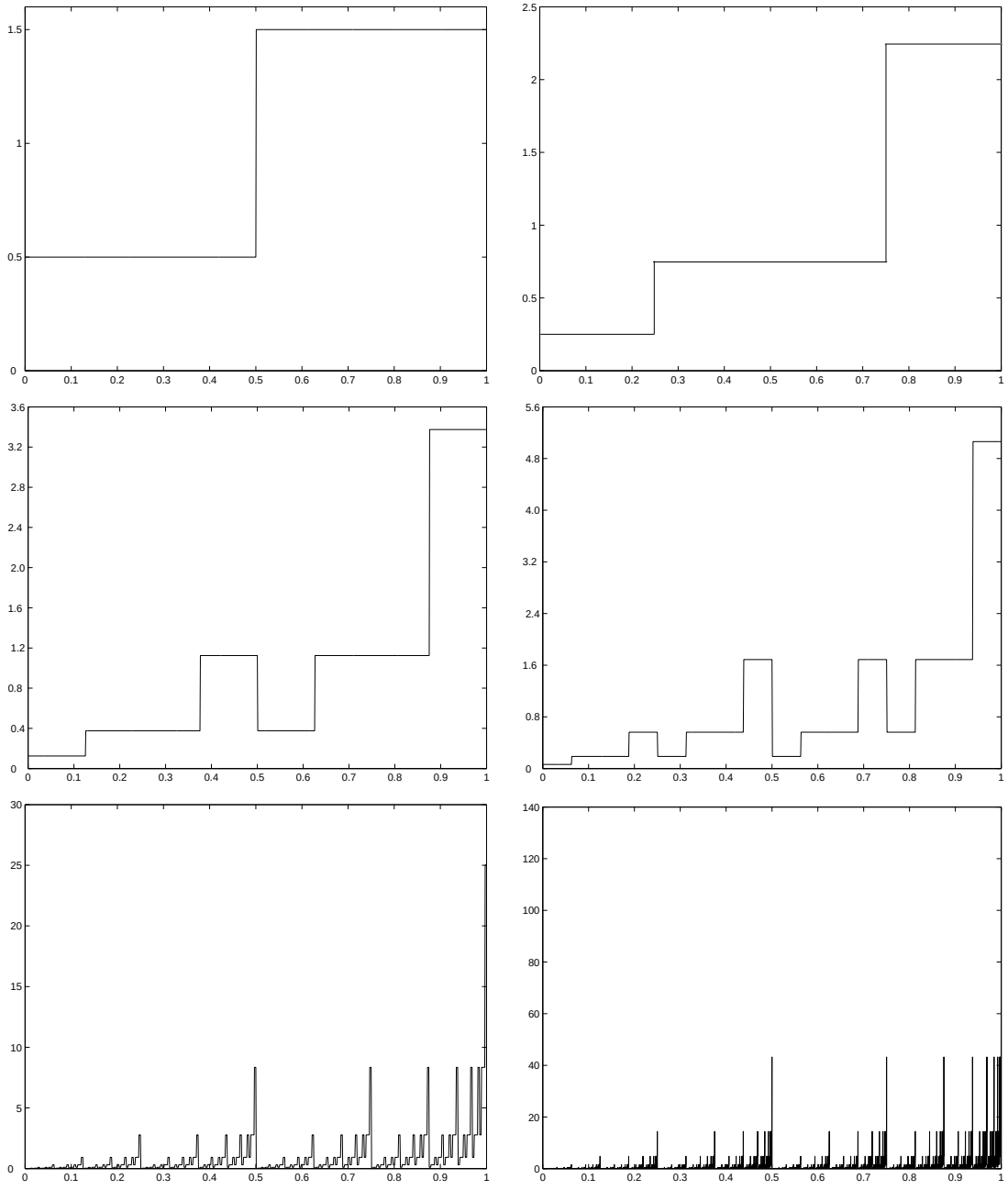


Figure 1.2: **Construction of the deterministic binomial measure ν.** From top to bottom, left to right, construction of the measure after 1, 2, 3, 4, 8 and 12 iterations, for $m_0 = 1/4$ and $m_1 = 3/4$. The measure is renormalized so that $\int d\mu = 1$.

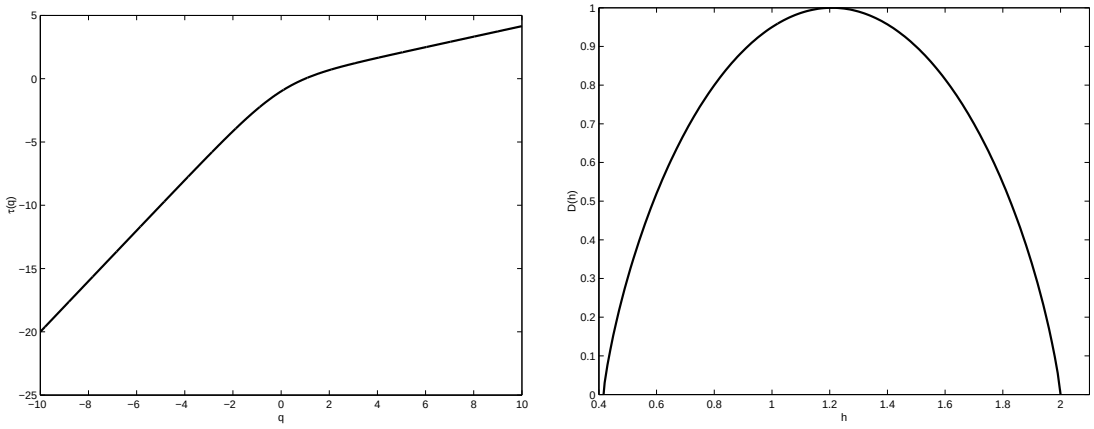


Figure 1.3: Binomial cascade: partition function $\tau(q)$ and Hausdorff spectrum $D(h)$ when $m_0 = 0.25$ and $m_1 = 0.75$.

the σ-algebra generated by the dyadic intervals of $[0, 1)$. The first few steps of the construction of μ with $m_0 = 1/4$ and $m_1 = 3/4$ are illustrated in Figure 1.2.

The binomial measure is by construction very irregular and possesses no density. We are generally interested in the process $X(t) = \int_0^t d\mu$. $X(t)$ is multifractal unless $m_0 = m_1 = 1/2$ and its multifractal spectrum (defined at the end of the previous section) is given by the Legendre-Fenchel (LF) transform of a so-called partition function $\tau(q)$

$$D(h) = \inf_q (qh - \tau(q)) \tag{1.10}$$

where $\tau(q)$ is given by

$$\tau(q) = -\log_2(m_0^q + m_1^q). \tag{1.11}$$

A review of the Legendre-Fenchel transform is given in Appendix A. As illustrated in Figure 1.3, the spectrum of the binomial measure with $m_0 = 0.25$ and $m_1 = 0.75$ is concave. In fact, the concave $\cap$-shape of the spectrum of X is typical of multiplicative cascades. All results and observations about the multifractal spectrum in this section can be found in [102].

The dyadic deterministic construction can be easily generalized by splitting $[0, 1)$ into $b > 2$ subintervals of equal length, each receiving a mass $m_0, \ldots, m_{b-1} > 0$, $\sum_k m_k = 1$. To this construction, we can associate a deterministic b-ary construction tree, whose branches are equipped with the non-random weights $m_0, \ldots, m_{b-1}$. The spectrum of the integral of the limit measure is then the LF transform of the partition function

$$\tau(q) = -\log_b(m_0^q + \ldots + m_{b-1}^q). \tag{1.12}$$

Random example. The procedure can be randomized by allocating a random mass to each subinterval at each iteration. Dyadic intervals $I_{n,k}$ therefore receive a random mass

$$\mu(I_{n,k}) = M_{n,k}^{(n)} \ldots M_{1,k}^{(1)}.$$

We usually make the following assumption about the random weights, for the random binary cascade:

- All multipliers $M_{1,k}^{(1)}, \ldots, M_{n,k}^{(n)}$ are independent and positive a.s.

- Conservation of the mass in the mean: $\mathbb{E}(M_{n,2k}^{(n+1)} + M_{n,2k+1}^{(n+1)}) = 1$
- $M_{n,k}^{(n)} \stackrel{d}{=} M_0$ if k is even and $M_{n,k}^{(n)} \stackrel{d}{=} M_1$ if k is odd.

where $\stackrel{d}{=}$ denotes equality in distribution. Products of multipliers can be very small and despite the conservation in the mean, the total mass distributed over the interval $[0, 1)$ may degenerate in some cases. However, under further conditions on the multipliers, one can show $\mathbb{E}\mu([0, 1)) = 1$. The Hausdorff spectrum of the integral of the random binomial measure is then the LF transform of

$$\tau(q) = -\log_2 \mathbb{E}(M_0^q + M_1^q). \tag{1.13}$$

The generalization of the construction of the random measure by splitting $[0, 1)$ into $b > 2$ subintervals of equal length yields the partition function $\tau(q) = -\log_2 \mathbb{E}(M_0^q + \ldots M_{b-1}^q)$. Many other random cascades can be contemplated. In Chapter 3, we derive an upper bound for the Hausdorff spectrum of a new process defined as the integral of a measure obtained from a cascade construction on a random tree. The novelty is the way we define an embedding from the boundary of the random tree to intervals of the real line, which differs from previously proposed random partitions, for example [100].

1.1.9 Multifractal formalism

For estimation, detection or classification purposes in signal processing, it is important to be able to estimate the spectrum $D(h)$ of a signal. It is a way to distinguish a monofractal process from a multifractal process for example. In practice, we are facing a double problem. Firstly, locating all points of the process with a given Hölder regularity is not feasible due to the finite precision of the data. Secondly, as discussed before, the difficulty in estimating the Hausdorff dimension of a given set due to the presence of the infimum in its definition. Alternative methods of estimations were sought, and gave birth to what is known as the *multifractal formalism*. By multifractal formalism we mean a formalism where we calculate the Legendre-Fenchel transform (see Appendix A) of a partition function, which generally provides an upper bound for the multifractal spectrum. When this upper bound is the multifractal spectrum, we will say that the multifractal formalism holds.

Multifractal formalism is associated with the study of the moments of multiresolution quantities $T_X(a, t)$, obtained from a comparison of the original process with a reference pattern $\psi(t)$ dilated and located at different positions [3, 7]

$$T_X(a, t) =< X, \psi_{a,t} >= |a|^{-1} \int X(u)\psi((u - t)/a)du \tag{1.14}$$

where $\psi_{a,t}(u) = |a|^{-1}\psi((u - t)/a)$. In previous sections, we considered increments of SSSI processes. It is easy to see that increments are the result of the comparison of the original process $X(t)$ with $\psi(t) = \delta(t + \tau_0) - \delta(t)$, where $\delta(t)$ is the Dirac distribution located at t. As we will see later, we can increase the regularity of $\psi(t)$

by considering techniques based on wavelets coefficients. The study of $\mathbb{E}|T_X(a,t)|^q$ is in practice replaced by time averages, under the assumption that the $\{T_X(a,t_k)\}_k$ form a stationary sequence, for some partition t_k, $k \in \mathbb{Z}$. Then, a 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

$$\frac{1}{n_a}\sum_{k=1}^{n_a}|T_X(a,t_k)|^q \simeq C_q|a|^{\zeta(q)}$$

where n_a is the number of $T_X(a,t_k)$ available at scale a. $\zeta(q)$ is called the partition function. Usually, this behaviour is valid only for a limited range of finite scales and a limited range of q. It is worth mentioning that the power law behaviour of time averages may differ from the power law behaviour (if it exists) of ensemble averages. To illustrate this counterintuitive fact, it has has been demonstrated that Compound Poisson Motions [19] have stationary increments which satisfy [32]

$$\mathbb{E}|T_X(a,t)|^q \simeq C_q|a|^{\lambda(q)}$$

in the limit as $a \to 0$. This expression holds for a finite range of $q \in (0, q_c^+)$ where q_c^+ is the largest q such that $\mathbb{E}|T_X(a,t)|^q$ is finite. Although it is tempting to believe that $\zeta(q)$ and $\lambda(q)$ agree on $(0, q_c^+)$, it is now acknowledged that equality holds only for a smaller range of q values [92, 98].

The choice of $T_X(a,t)$ plays a central role for the estimation of the partition function. Multiresolution quantities based on a wavelet decomposition of the process are the most powerful tool to date [1, 3, 7, 8, 66], since they allow the study of the signal at different scales and positions. We introduce $\xi(q)$, the LF transform of $D(h)$,

$$\xi(q) = 1 + \inf_h(qh - D(h)).$$

In the next section we recall relations between $\xi(q)$ and $\zeta(q)$, for two expressions of $T_X(a,t)$.

Wavelet based estimators. The discrete wavelet transform is a time/scale representation of a signal $X(t)$ using a multiresolution analysis, decomposing a signal into two parts: approximations and details. Approximations are obtained as the result of projections of the signal onto a low frequency function ϕ_0, called the scaling function [82, 91]. This operation realizes a low-pass filter and retains the slow variations of the signal. Details are obtained after comparison of the signal with a high frequency function called the mother wavelet ψ_0. This projection performs a high-pass filter of the signal and only keeps its fast variations. To reconstruct the signal from its projections on ψ_0 and ϕ_0, ψ_0 must satisfy the admissibility condition [50]

$$\int_{\mathbb{R}} \psi_0(t)dt = 0.$$

A wavelet is described by the number N of its vanishing moments

$$\int t^q\psi_0(t)dt = 0, \quad 0 \leqslant q < N.$$

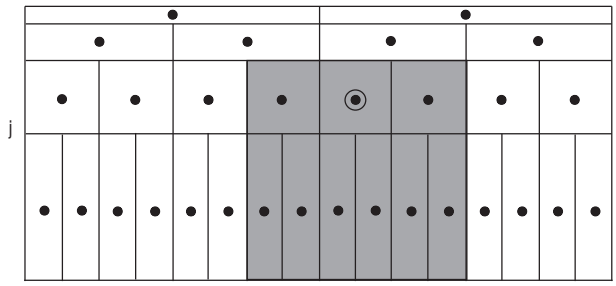


Figure 1.4: The Wavelet leader $L_X(j,k)$ in black circle is defined as the maximum of all detail coefficients $d_X(\cdot,\cdot)$ (black dots) over $3\lambda_{j,k}$ represented in a shaded area in the picture.

Let $\psi_{j,k}(t) = 2^{-j/2}\psi_0(2^{-j}t-k)$ be a copy of ψ_0, scaled by a factor 2^{-j} and shifted by k. We call the detail coefficient $d_X(j,k)$ the inner product of $X(t)$ with $\psi_{j,k}(t)$

$$d_X(j,k) =< X, \psi_{j,k} >= \int_{\mathbb{R}} X(t)\psi_{j,k}(t)dt. \tag{1.15}$$

The mother wavelet can be chosen so that $\{2^{-j/2}\psi_0(2^{-j}t-k), j \in \mathbb{Z}, k \in \mathbb{Z}\}$ forms an orthonormal basis of the space of square integrable signals $L^2(\mathbb{R})$. Let $X \in L^2(\mathbb{R})$, then X can be decomposed

$$X(t) = \sum_{j\in\mathbb{Z}}\sum_{k\in\mathbb{Z}} d_X(j,k)\psi_{j,k}(t).$$

Let us go back to the estimation of the partition function $\zeta(q)$. Noticing the similarity between Equations (1.14) and (1.15), we now have more powerful tools to compare the process with wavelets of higher regularity than $\psi(t) = \delta(t+\tau_0) - \delta(t)$, by increasing the number of vanishing moments. Estimators based on wavelet coefficients study the behaviour of

$$S_1(q,j) = \frac{1}{n_j}\sum_{k=1}^{n_j} |d_X(j,k)|^q$$

in the limit $2^j \to 0$. We define

$$\zeta_1(q) = \liminf_{j\to-\infty}\big(\frac{\log_2 S_1(q,j)}{j}\big). \tag{1.16}$$

It was noted in [3] that $\zeta_1(q) = \xi(q)$ for all positive q, giving only an upper bound of the increasing part of the multifractal spectrum (see Appendix A). However, when the process possesses oscillating singularities of the form $|t-t_0|^h \sin(|t-t_0|^{-\beta})$ with $h, \beta > 0$, the wavelet coefficient based estimator fails. In practice, the difficulty of estimating the partition function for negative q comes from numerical issues. Since wavelet coefficients can be very small, raising them to a negative power increases the uncertainty and leads to estimation errors. In 2006, Jaffard, Lashermes and Abry proposed a better estimator based on the so called wavelet leaders [66]. Let $\lambda_{j,k} =$

$[k2^j, (k+1)2^j)$ be the k-th dyadic interval at scale 2^j and $3\lambda_{j,k} = \lambda_{j,k-1} \cup \lambda_{j,k} \cup \lambda_{j,k+1}$ be the neighborhood around $\lambda_{j,k}$. The wavelet leaders $L_X(j,k)$ are then defined by

$$L_X(j,k) = \sup_{\lambda' \subset 3\lambda_{j,k}} |d_X(\cdot,\cdot)|$$

where the supremum is taken on $d_X(\cdot,\cdot)$ in the neighborhood $3\lambda_{j,k}$ over all finer scales $2^{j'} < 2^j$. This is illustrated in Figure 1.4. The estimation of the partition function using wavelet leaders requires a slightly stronger condition on the process than its continuity:

DEFINITION 1. *A process X is said Hölder uniform if there exists $\epsilon > 0$ such that*

$$\exists C > 0 \text{ such that } \forall t, s \in \mathbb{R},\ |X(t) - X(s)| \leqslant C|t-s|^\epsilon.$$

The $\zeta(q)$ estimator is based on the computation of

$$S_2(q,j) = \frac{1}{n_j} \sum_{k=1}^{n_j} |L_X(j,k)|^q.$$

It is known that

$$\zeta_2(q) = \liminf_{j \to -\infty} \left(\frac{\log_2 S_2(q,j)}{j}\right) \tag{1.17}$$

agrees with $\xi(q)$ for all q [3] for Hölder uniform processes, whether the process possesses oscillating singularities or not. Thus, inverting the LF transform we obtain an upper bound of the Hausdorff spectrum

$$D(h) \leqslant \inf_{q \neq 0} (1 + qh - \zeta_2(q)). \tag{1.18}$$

For concave $D(h)$, the LF transform is involutive (see Appendix A) and therefore equality in (1.18) holds. We have seen earlier that this is the case for multiplicative cascades. The major advantage of the wavelet leader based multifractal formalism is the ability to estimate the partition function for all values of q. Another important property is that estimation of $\zeta(q)$ for positive q is independent of the wavelet basis chosen. For negative q, a similar result holds if the wavelet belongs to the Schwartz class. However, as noted by the authors, estimations using Daubechies wavelets (which are not in the Schwartz class) performed well in [66], indicating that this assumption could be weakened. For all the above reasons, in simulation trials, we have chosen to estimate the spectrum of MEBP processes using the wavelet leader technique presented in this introduction.

Wavelet leaders appeared more than ten years after an initial robust estimator of the partition function was proposed in 1993, called the Wavelet Transform Modulus Maxima (WTMM). Since it is sufficient to derive the position and values of the maximum of the wavelet transform to characterize the singular behaviour of functions [83], Arneodo, Bacry and Muzy used this technique to derive the spectrum of singularities of a signal [12] and applied it in the context of turbulence [7]. A comparative study in [66] indicated equivalent performances of the two techniques.

1.2 Research work

In this thesis, I propose two new models to generate fractal processes, whose construction relies on a branching process. A branching process is by definition a system of particles which live for a random time and can give birth to offspring up to the moment of their death. Conditioned on when and where they are born, offsprings are independent of their parent and siblings. We are particularly interested in the Galton-Watson process, the oldest and simplest branching process. We describe it as follows. Start with a single ancestor. Suppose it lives exactly one unit of time and that it gives birth to a random number of children when it dies. Let p be the distribution of this random variable. Each offspring from the first generation behaves exactly the same way as the initial particle and independently of the others. They live one unit of time and give birth to a random number of offspring at the moment of death, according to the distribution p. And so on. The process can be described mathematically using a discrete time index, giving the size of the population Z_n at time $n = 0, 1, 2, \ldots$ The random variables Z_n possess very interesting and well known properties (such as the Markov property) and provide intuition for more complicated processes.

This model was first studied by Bienaymé in 1845, where he shows verbally in a communication that the theorem on extinction of families is known to him [21]. His contribution is however absent from the branching process literature [57]. It was then reintroduced by Galton and Watson in 1874 when they were studying the problem of extinction of surnames of noble English families. Galton noticed that in many cases, surnames which were once very common totally disappeared after a few generations. He addressed the problem by asking how many generations would elapse before a name would disappear, given that a man has $0, 1, 2, \ldots$ sons with probabilities $p_0, p_1, p_2, \ldots$ His friend the Reverend H.W. Watson solved the problem using the iteration of generating functions [47]. Generalizations of the original Galton-Watson process can be found for example in [54].

Applications of branching processes go beyond the study of population demography, and have been applied to model cell growth in many areas of biology [117]. Other applications include polymerase chain reactions and gene amplification, to cite but a few [67, 74]. Galton-Watson processes have also been used to produce random fractal sets in the theory of Iterated Function Systems [43, 90]. The novelty of my doctoral work is to introduce new models for generating fractal signals using branching processes of the Galton-Watson type. The first model is a generalization of the construction of Iterated Function Systems acting over the space of signals. The second model concerns signals whose so-called crossing tree is a Galton-Watson process, which we call Multifractal Embedded Branching Processes (MEBP) processes. We briefly introduce them here.

1.2.1 Galton-Watson Iterated Function Systems

Deterministic fractal sets satisfy relation (1.1). They can be constructed via a deterministic recursive procedure. Starting from an initial set, apply M contractive maps to it. Repeat the procedure *ad infinitum*. The M-ary tree associated with

this construction is a deterministic M-ary tree. This procedure can be randomized in many different ways. We can apply at each iteration a fixed number of random maps, but we can also consider a random number of random maps at each step. The theory of random fractal sets were studied by Falconer [43], Graf [48], Mauldin and Williams [90]. They derived the exact Hausdorff dimension for random sets [49] which reduces to the result of Moran in the deterministic setting [93]. In [61] and [62], Hutchinson and Rüschendorff introduced new probability metrics for random measures and obtained stronger results.

In Chapter 2, we propose a construction of a random IFS based on the works of Hutchinson and Rüschendorff. We consider a random number of maps at each iteration of the algorithm, each map being random. The construction tree is then a Galton-Watson branching process. We study conditions of existence and uniqueness of a fixed point of the IFS. It is shown in [14] that the fractal attractor of a deterministic IFS continuously depends on the parameters of the IFS. We extend this result and show, in a special case, that the moments of the fixed point continuously depend on the distribution of the number of maps used at each iteration of the algorithm.

1.2.2 Multifractal Embedded Branching Processes

In the second part of this thesis, we propose a new class of multifractal processes, called Multifractal Embedded Branching Processes (MEBP) processes, which can be efficiently simulated on-line (Chapter 3). MEBP are defined using the crossing tree, an ad-hoc space-time description of the process, and are such that the crossing tree is a Galton-Watson branching process. The crossing tree of a given realisation of a signal is obtained in the same way as Barlow and Perkins [13] associated a branching process with a diffusion on the Sierpinski gasket. For any suitable branching process there is a family of discrete-scale invariant processes—identical up to a continuous time change—for which it is the crossing tree. We identify one of these as the Canonical Embedded Branching Process (CEBP), and then construct MEBP from it using a multifractal time change. To allow on-line simulation of the process, the time change is constructed from a multiplicative cascade on the crossing tree. Time-changed self-similar signals, in particular time-changed Brownian motion, are popular models in finance [86, 87].

Brownian motion can be constructed as a CEBP, so MEBP include a class of time changed Brownian motion, suggesting their suitability for modelling purposes. We also propose to imitate an fBm with an MEBP. Proofs of the existence and continuity of MEBP are given, together with an efficient on-line algorithm for simulating them (a Matlab implementation is freely available from the web page of Jones). Also, using an approach of Riedi, an upper bound on the multifractal spectrum of the time change is derived (Chapter 4). Estimation of the spectrum using the wavelet leaders is also carried out, supporting the theoretical results. Further results about the Hausdorff spectrum of the time change defined on the boundary of the crossing tree are also obtained.

Chapter 2

Galton-Watson Iterated Function Systems

The terminology Iterated Function Systems (IFS), introduced by Barnsley and Demko [15], refers to a finite set of contractive mappings which completely specify a fractal set. The study of IFS was pioneered by Hutchinson in 1981 [60] and Barnsley and coworkers. Hutchinson proved the existence of a unique fractal set and measure of an IFS using a contraction mapping principle, whereas Barnsley proved the same result using a probabilistic set up. In the first section of this chapter, we review the basic construction of fractal sets and measures using IFS. We mainly focus on the method proposed by Hutchinson since it will be relevant for the proposed extension in a later section. In addition, we briefly review alternatives to the construction of IFS, like Recursive Iterated Function Systems (RIFS), and explain how to add randomness to the model. In the second section we consider the space of p integrable functions and define a set of contractive mappings acting over this space, using the approach developed in [63]. These IFS rely on an M-ary underlying construction tree, if M is the total number of contractions used. We propose to extend the model by allowing a random construction tree, with random contractions. The existence of a unique fixed point is shown following the approach of Hutchinson and Rüschendorff. Properties of the fixed point are described in section 2.5.

2.1 Self-similar sets and measures

In this section we review the definition and construction of self-similar sets. But first, we recall fundamental mathematical notions such as complete metric spaces, the fixed point theorem and the Hausdorff metric.

2.1.1 Contractive operators in metric spaces

Let $\mathbb{X}$ be a space and define a metric $d : \mathbb{X} \times \mathbb{X} \to \mathbb{R}$ on this space. $(\mathbb{X}, d)$ is then called a metric space. The concept of metric leads to the notion of convergence. A sequence of points $(x_n)_{n\in\mathbb{N}}$ of $\mathbb{X}$ is said to converge to an element of $\mathbb{X}$ when the distance determined by the metric d between the two elements can be made arbitrarily small by sending n to infinity. We are mainly concerned with sequences

known as Cauchy sequences, which satisfy

$$\forall \epsilon > 0 \;\; \exists r \in \mathbb{N} \;\; \forall (p,q) \in \mathbb{N}^2 \;\; p \geqslant r \;\; q \geqslant r \;\; \Rightarrow d(x_p, x_q) \leqslant \epsilon$$

where $(x_n)_{n\in\mathbb{N}}$ is a sequence of $\mathbb{X}$. The definition of a Cauchy sequence states that the distance between 2 elements of the sequence can be made arbitrarily small if we consider p and q large enough. Note that a Cauchy sequence does not necessarily converge. A metric space for which every Cauchy sequence converges is called complete.

DEFINITION 2. *If $f : \mathbb{X} \to \mathbb{X}$, we define the Lipschitz constant of f by*

$$Lip(f) = \sup_{x \neq y} \frac{d(f(x), f(y))}{d(x,y)}.$$

f is Lipschitz if $Lip(f) < \infty$ and f is contractive if $Lip(f) < 1$

The following theorem plays a central role in the theory of Iterated Function Systems. It is known as the Banach fixed point theorem.

THEOREM 1. *Let $(\mathbb{X}, d)$ be a complete metric space and f a contractive map. Then f possess a fixed point in $\mathbb{X}$. Moreover, this fixed point is unique.*

2.1.2 Complete space of compact sets

Let $K \subset \mathbb{R}^2$. Suppose

- K is bounded, that is there exists $r > 0$ such that for all $x \in K$, $d(x, 0) \leqslant r$.
- K is closed. A set K is closed if its complement K^c is open, that is if for all $x \in K^c$, there exists $r > 0$ such that $B(x, r) \subset K^c$, where $B(a, r)$ is the open ball of center a and radius r.

K is compact if and only if for every sequence $(x_n)_{n\in\mathbb{N}}$, it is possible to extract a subsequence which converges in K. It is a standard result that if the underlying space is of finite dimension, then K is compact if and only if K is closed and bounded. The set of compact subsets of $\mathbb{R}^2$ is usually denoted $\mathcal{H}(\mathbb{R}^2)$.

Let $x \in \mathbb{R}^2$ and $S \in \mathcal{H}(\mathbb{R}^2)$. Define the distance between x and $K \in \mathcal{H}(\mathbb{R}^2)$ by

$$d(x, K) = \inf\{d(x, y) \mid y \in K\}$$

and the distance between K and S by $d(K, S) = \inf\{d(x, S) \mid x \in K\}$. Then, define the Hausdorff metric on compact subsets of $\mathbb{R}^2$ by

$$d_{\mathcal{H}}(K, S) = \sup\{d(K, S), d(S, K)\}.$$

It follows that $(\mathcal{H}(\mathbb{R}^2), d_{\mathcal{H}})$ is a complete metric space (see for example [14]). This result motivates the study of sequences of compact sets defined via contraction mappings and the definition of IFS and self-similar sets.

2.1.3 Construction of self-similar sets

Consider a finite set $\{\omega_1, \ldots, \omega_M\}$ of contraction maps $\mathbb{R}^2 \to \mathbb{R}^2$. If $K \in \mathcal{H}(\mathbb{R}^2)$, define the scaling law W by

$$W(K) = \bigcup_{i=1}^{M} \omega_i(K).$$

Denote by s_i the Lipschitz constant of ω_i. W is contractive with Lipschitz constant $s = \max\{s_1, \ldots, s_M\}$. Using the terminology of Hutchinson and Rüschendorff, K^* satisfies the scaling law W if $K^* = W(K^*)$ [63]. Such a set is by construction discrete self-similar, since it can be decomposed into a union of scaled identical M copies of itself. They proved the existence of a unique self-similar set K^* using the contraction mapping theorem [60]. The Hausdorff dimension r of K^* is usually fractional and satisfies

$$\sum_{i=1}^{M} s_i^r = 1.$$

Let $W^p(K) = W(W^{p-1}(K))$. Starting from an arbitrary initial bounded set $K_0 \neq \emptyset$, Hutchinson proved that $W^p(K_0) \to K^*$ in the Hausdorff metric. K^* is therefore known as the fixed point or the attractor of the IFS. This gives an algorithm to generate approximations of K^*. Select a starting point x_0 and define $K_0 = \{x_0\}$. Let $K_1 = W(K_0)$, $K_2 = W(K_1)$, and so on. As $n \to \infty$, $d_{\mathcal{H}}(K_n, K^*) \to 0$. So for n large enough, we can obtain a good approximation of the attractor.

Pictures of K^* can also be obtained through a random and faster procedure, called the 'Random Iteration' or 'Chaos game' [14, 15]. Consider an initial point x_0 and apply a contractive mapping ω_{i_1} chosen uniformly among the M possible contractions. $x_1 = \omega_{i_1}(x_0)$. Select again another transformation, independently from the previous one, and apply it to x_1. Repeat the procedure many times to obtain a sequence of points. It is a famous result that the orbit $\{x_n\}$ is dense in K^*. This random algorithm can be slightly modified by not picking ω_k with uniform probability but with probability p_k, with $\sum_k p_k = 1$. The orbit still converges to the attractor of the IFS, however some regions of the fixed point are visited more than others, depending on the values of p_k. In fact the random algorithm generates a picture of a measure μ, as suggested by the relation [42]

$$\mu(B) = \lim_{n\to\infty} \frac{1}{n+1} \sum_{k=0}^{n} \chi_B(x_k)$$

where χ_B is the characteristic function of B. This equality makes explicit the relation between a measure, which support is the attractor of the IFS and the relative visitation frequency of a set B. This notion is made precise using the notion of self-similar measure, and it was proved in [60] that the measure μ previously obtained is the unique self-similar measure μ of total mass one, such that

$$\mu = \sum_{i=1}^{M} p_i \mu \circ \omega_i^{-1}.$$

The support of μ is the attractor of the IFS without probabilities. The chaos game was extended by Barnsley and coworkers [16] in 1989 by allowing dependence between the choice of mappings, where the probability of selecting a map at iteration n depends on the selection made at step $n-1$. They called the procedure Recurrent Iterated Function Systems (RIFS) and proved the existence of a unique compact set called attractor of the IFS. The major difference lies in the fact that the attractor of an RIFS need not exhibit self-similarity.

The previous constructions have been randomized in various ways to obtain statistically self-similar sets. The theory was extensively investigated by Falconer [43], Mauldin and Williams [90] and Graf [48]. Later, Arbeiter [6] and Olsen [97] studied the theory of random self-similar measures.

IFS considered so far are acting over the space of sets and measures. It is possible to adapt these definitions to the study of self-similar signals and obtain fast algorithms to simulate them.

2.2 Iterated Function Systems on functions

In this section we present the model and introduce the working spaces. The random IFS model presented is referred to as a Galton-Watson IFS, referring to the random structure of its underlying construction tree.

2.2.1 Deterministic IFS

Let $L_p(\mathbb{X})$ be the space of p-integrable signals $\mathbb{X} \to \mathbb{R}$ where $\mathbb{X}$ is a compact subset of the real line. $||\cdot||_p$ is the usual norm defined on $L_p(\mathbb{X})$: $||f||_p = (\int |f(x)|^p dx)^{1/p}$, leading to the natural metric d_p defined by $d_p(f,g) = ||f-g||_p$ where f and g are in L_p. (L_p, d_p) is then a complete metric space (Riesz-Fisher theorem, [96]). It is common to consider without loss of generality the case $\mathbb{X} = [0,1]$.

Similarly, an IFS on functions consists of recursively applying a contractive operator T. Starting with an initial function f_0, we denote by $T^n f_0$ the n-th iterate of T acting on f_0. For a class of operators T, the IFS converges to a function f^*

$$T^n f_0 \to f^* \text{ as } n \to +\infty \tag{2.1}$$

in $L_p(\mathbb{X})$. f^* is the unique function satisfying $f = Tf$ and similarly, we say that f^* satisfies the scaling law T. That is, f^* is the fixed point or attractor of the IFS associated with T. It is generally assumed that T can be decomposed into a set of M nonlinear operators $\phi_i : \mathbb{R} \times \mathbb{X} \to \mathbb{R}$ for $1 \leqslant i \leqslant M$. Each ϕ_i deforms the original signal and maps it to a subinterval $\mathbb{X}_i = \varrho_i(\mathbb{X})$ of $\mathbb{X}$. Specifically,

$$(Tf)(x) = \sum_{i=1}^{M} \phi_i[f(\varrho_i^{-1}(x)), \varrho_i^{-1}(x))]\mathbf{1}_{\varrho_i(\mathbb{X})}(x) \tag{2.2}$$

where $\{\varrho_i(\mathbb{X})\}_{i=1}^M$ partitions $\mathbb{X}$. $\mathbf{1}_{\varrho_i(\mathbb{X})}$ is the indicator function of the interval $\varrho_i(\mathbb{X})$. The procedure is illustrated in Figure 2.1. In (2.2), ϕ_i are functions of two variables. The second variable is optional however (we introduce it to obtain more general

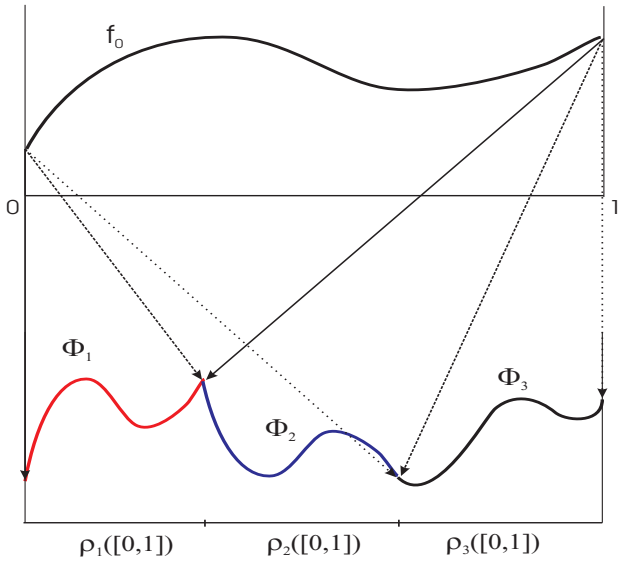


Figure 2.1: Operator T acting on f_0. Here, $M = 3$.

results) and one can define the operator T with $\phi_i : \mathbb{R} \to \mathbb{R}$. The underlying construction tree is an M-ary deterministic tree. Conditions of convergence of the IFS are derived explicitly in [63] for non-linear functions $\phi_i : \mathbb{R} \to \mathbb{R}$. The result can be easily generalized to functions $\phi_i : \mathbb{R} \times \mathbb{X} \to \mathbb{R}$, as above.

THEOREM 2. *Suppose that ϱ_i are strict contractions with contraction factors $r_i < 1$ for $i = 1, \ldots, M$, and that ϕ_i are Lipschitz in their first variable with Lipschitz constants s_i. If for some p, $\lambda_p = \sum_{i=1}^{M} r_i s_i^p < 1$ and $\sum_{i=1}^{M} r_i \int |\phi_i(0,x)|^p dx < \infty$, then T has a unique fixed point in $L_p(\mathbb{X})$.*

This is a specific case of Theorem 4, so the proof is not given here. The conditions for convergence are quite weak. The second condition only requires that ϕ_i must be p integrable with respect to their second variable.

Figure 2.2 presents attractors of two different IFS, one continuous and one discontinuous. Conditions for continuity are derived in section 2.5.1 in a more general setting.

The deterministic model acting on functions is not flexible enough to model natural signals. This is mainly due to its deterministic self-similarity as observed in Figure 2.2. One way to break this pattern is to add randomness to the construction. Section 2.2.3 defines random IFS with random operators and a random construction tree.

2.2.2 $\mathbb{L}_p$ spaces

Before giving the definition of a Galton-Watson IFS we need to specify the space where the fixed point lies. Let $(\Sigma, \mathcal{F}, P)$ be a probability space. We endow L_p with its σ-algebra $\mathcal{L}_p$ ([107], définition 25.2). Then, a p-integrable random process is a

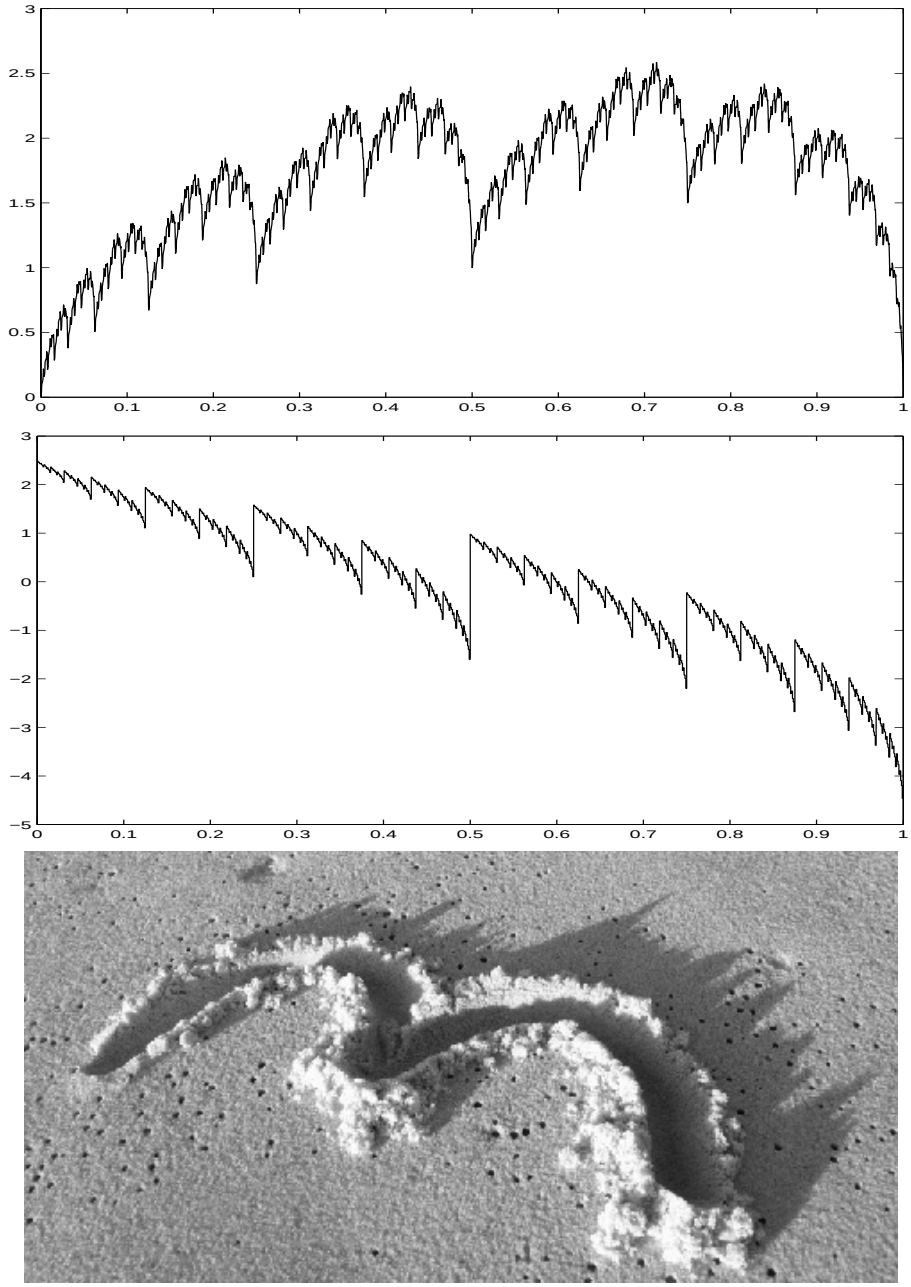


Figure 2.2: Top signal: continuous attractor of the IFS defined with the maps $\phi_1(u, v) = s_1 u + v^3$ and $\phi_2(u, v) = s_2 u + (1 - v^2)$, where $s_1 = s_2 = 0.75$. The middle discontinuous signal is also obtained as the fixed point of an IFS, whose parameters are $\phi_1(u, v) = s_1 u + 1$ and $\phi_2(u, v) = s_2 u - 1$, where $s_1 = 0.6$ and $s_2 = 0.8$. $\mathbb{X} = [0, 1]$ in both cases. The bottom figure is a natural fractal obtained at dusk, on a beach in Queensland.

random variable $f : \Sigma \to L_p(\mathbb{X})$, provided f is measurable. Define

$$\mathbb{L}_p = \{f : \Sigma \to L_p(\mathbb{X}), f \text{ measurable} \mid \mathbb{E}\Big[||f||_p^p\Big] < +\infty\}$$

where $\mathbb{E}$ denotes expectation under P. We denote by f_σ a realisation of the random process $f \in \mathbb{L}_p$, where $\sigma \in \Sigma$. $f(x) : \Sigma \to \mathbb{R}$ is the random variable obtained by evaluating f at x. The goal is to define a metric d_p^* over $\mathbb{L}_p$ such that $(\mathbb{L}_p, d_p^*)$ is a complete metric space. Let

$$||f||_p^* = \mathbb{E}^{\frac{1}{p}}\big[||f||_p^p\big] \tag{2.3}$$

for all random p-integrable functions $f \in \mathbb{L}_p$.

LEMMA 1. *$||\cdot||_p^*$ is a norm on $\mathbb{L}_p$, $p \geqslant 1$.*

Proof. The lemma is obvious for $p = 1$. We consider $p \neq 1$ in the following. The key is to derive the triangle inequality for $||\cdot||_p^*$, using the Hölder and Minkowski inequalities.

Integrals are defined in this proof with respect to the Lebesgue measure. First note that if a and b are non negative reals and p and q are such that $\frac{1}{p} + \frac{1}{q} = 1$ and $1 < p, q < \infty$, then $ab \leqslant \frac{a^p}{p} + \frac{b^q}{q}$. This inequality can be derived using the concavity of log. This gives $\mathbb{E}\int |\bar{f}\bar{g}| \leqslant \mathbb{E}[\int \frac{1}{p}|\bar{f}|^p + \int \frac{1}{q}|\bar{g}|^q]$ where we define $\bar{f} = \frac{f}{||f||_p^*}$ and $\bar{g} = \frac{g}{||g||_q^*}$, from which

$$||fg||_1^* = \mathbb{E}\int |fg| \leqslant ||f||_p^* ||g||_q^* \tag{2.4}$$

follows. This is the equivalent of the Hölder inequality for random p and q integrable functions.

Applying the triangle inequality to $|f + g|$, $||f + g||_p^{*p}$ is smaller than $\mathbb{E}\int |f + g|^{p-1}(|f| + |g|) = \big|||f + g|^{p-1}|f|\big|_1^* + \big|||f + g|^{p-1}|g|\big|_1^*$. Thus, using the previous Hölder's inequality:

$$||f + g||_p^{*p} \leqslant \big|||f + g|^{p-1}\big|_q^* \big[||f||_p^* + ||g||_p^*\big].$$

Since $pq = p + q$,

$$\big|||f + g|^{p-1}\big|_q^* = \mathbb{E}^{\frac{1}{q}}\Big(\int |f + g|^p\Big) = ||f + g||_p^{*p-1}. \tag{2.5}$$

Hence:

$$||f + g||_p^{*p} \leqslant ||f + g||_p^{*p-1}\big[||f||_p^* + ||g||_p^*\big].$$

When $||f + g||_p^* = 0$, the inequality is trivial. When it is not, we can divide each side of the inequality by $||f + g||_p^{*p-1}$ which concludes the proof of lemma. □

Lemma 1 leads us to define the metric d_p^* as follows: $d_p^*(f, g) = ||f - g||_p^*$. It is straightforward to adapt the proof of the Riesz-Fisher theorem [96] to show that $(\mathbb{L}_p, d_p^*)$ is a Banach space.

2.2.3 Galton-Watson IFS

The operator T acting over the space $\mathbb{L}_p$ is now defined as follows:

$$(Tf)(x) = \sum_{j=1}^{Z} \phi_j[f^{(j)}(\varrho_j^{-1}(x)), \varrho_j^{-1}(x))]\mathbf{1}_{\varrho_j(\mathbb{X})}(x) \tag{2.6}$$

where $(Z, \phi_1, \varrho_1, \ldots, \phi_Z, \varrho_Z)$ is random and $f^{(j)}$ are i.i.d. copies of $f \in \mathbb{L}_p$. The ϱ_j are affine maps and randomly partition $\mathbb{X}$ into Z subintervals. The contraction factor of ϱ_j is the random variable r_j, such that $0 < r_j < 1$ almost surely. ϕ_j are functions of two variables, Lipschitz in their first variable, with random Lipschitz factor s_j. Z is distributed according to a probability vector $\mathbf{q} = (q_1, q_2, \ldots)$. The underlying construction tree has therefore a random number of offspring at each node, as illustrated in Figure 2.3. Assuming that, in this construction, the random variable Z is independent and identically distributed from one node to another, the construction tree is a Galton-Watson branching process [95], hence the name of the IFS. A good survey on Galton-Watson processes can be found in [10]. The natural question is to know whether T possess a fixed point f^*, and if this fixed point is unique. By fixed point, we mean a function f^* which satisfies $f^* = Tf^*$ in $\mathbb{L}_p$. Also, consider an arbitrary random "seed" f_0 and let $T^n f_0 = T(T^{n-1} f_0)$ be the n-th iterate of f_0. Under the condition of existence of a fixed point, we want to find conditions such that $d_p^*(T^n f_0, f^*) \to 0$ as $n \to \infty$, providing a simple algorithm to produce an approximation of the fixed point. It is worth noting that the number of maps used to produce the n-th iterate of f_0 varies with n and is a random variable ν, defined above. In [35], Daoudi proposed to produce fractal functions based on a generalization of IFS (GIFS), allowing the number of maps used at each iteration to vary. The first fundamental difference between his model and ours is that the number of contractions used in [35] is not random whereas it is in the present construction. Moreover, Daoudi works in $\mathcal{H}(\mathbb{R}^2)$ and not in $\mathbb{L}_p$.

Hutchinson and Rüschendorff [63] have shown the existence and uniqueness of a fixed point in $\mathbb{L}_p$ with random maps and constant ν. Also, they have introduced new probability spaces to prove existence of a unique fixed point for random sets and measures, previously obtained by Falconer and Mauldin & Williams. We adopt their method to prove that T possess a unique fixed point f^* under some conditions.

2.3 Space of extended trees

An ad-hoc structure of Σ is needed in order to build i.i.d. copies of the signal f. We show how to do this in the present section using extended Galton-Watson trees.

The construction of the probability space of extended Galton-Watson trees relies on two famous theorems in measure theory: the Ionescu-Tulcea theorem and the Daniell-Kolmogorov extension theorem. We use the first theorem to build a probability space of the first n generations of extended trees for any finite integer n, then extend the construction to infinite trees using the Daniell-Kolmogorov extension theorem. An element of that space therefore consists of a realisation of a Galton-Watson tree whose branches are equipped with realisations of the IFS operators.

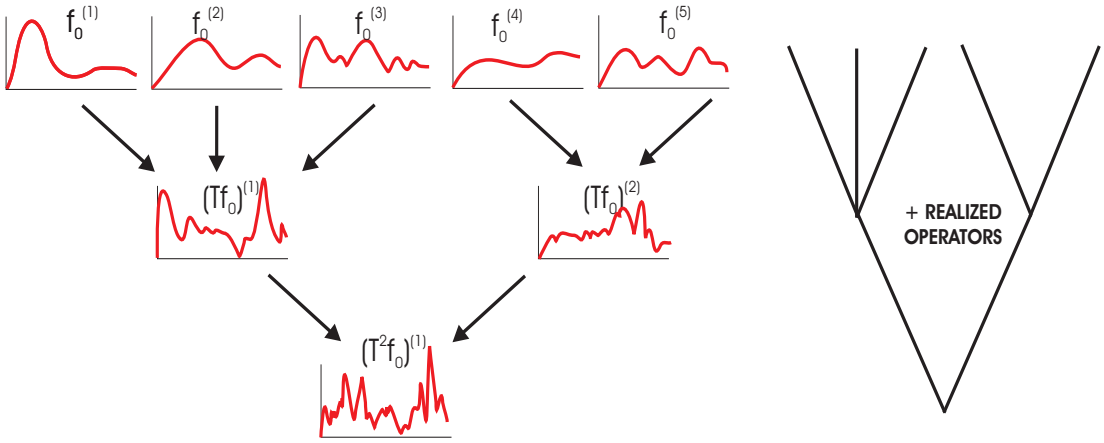


Figure 2.3: The first two generations of a Galton-Watson IFS. $f_0^{(j)}$ are i.i.d. realizations of an initial process f_0. Likewise, $(Tf_0)^{(j)}$ are i.i.d. realisations of the process Tf_0. The associated tree is a realisation of two generations of a Galton-Watson tree. $(T^2f_0)^{1)}$ is not uniquely determined by its construction tree: there is a loss of information. Branches of the tree need to be endowed with realisations of the random mappings of the IFS.

Ionescu-Tulcea [111]. The result of Ionescu-Tulcea relies on the concept of probability kernels. Let $(A_1, \mathcal{A}_1)$ and $(A_2, \mathcal{A}_2)$ be two measurable spaces. A probability kernel is a function $\kappa_2 : A_1 \times \mathcal{A}_2 \to [0,1]$ such that for all $E \in \mathcal{A}_2$, $a \mapsto \kappa_2(a, E)$ is a measurable function on A_1 and such that for all $a \in A_1$, $E \mapsto \kappa_2(a, E)$ is a probability measure on $(A_2, \mathcal{A}_2)$. We interpret κ_2 as a probability distribution on $(A_2, \mathcal{A}_2)$ conditioned on state $a \in A_1$ and write it either $\kappa_2(E|a)$ or $\kappa_2(a, E)$. Let $\kappa_3 : A_1 \times A_2 \times \mathcal{A}_3 \to [0,1]$ be a probability measure on $(A_3, \mathcal{A}_3)$ given we were in state (a_1, a_2) for $a_i \in A_i$, $i = 1, 2$ in the previous step. Then the kernel $\kappa_2 \otimes \kappa_3$ defined as

$$(\kappa_2 \otimes \kappa_3)(a_1, E) = \int \int \mathbf{1}_E(b, c)\kappa_2(a_1, db)\kappa_3(a_1, b, dc) \tag{2.7}$$

measures Borel subsets E of $\mathcal{A}_2 \times \mathcal{A}_3$ from an initial state $a_1 \in A_1$. Ionescu-Tulcea let us chain correctly n measurable spaces $(A_i, \mathcal{A}_i)$, $i = 1, \ldots, n$ by defining a joint probability on the product space $\prod_{i=1}^{n} A_i$ from n probability kernels κ_i. The result of Ionescu-Tulcea is then the following [111].

THEOREM 3. *Let κ_1 be a probability measure on $(A_1, \mathcal{A}_1)$ and for all $n \geqslant 2$, $\kappa_n : \left(\prod_{i=1}^{n-1} A_k\right) \times \mathcal{A}_n \to [0,1]$ a probability kernel. Then there exists a unique probability measure on $\prod_{i=1}^{n} A_k$ given by $\bigotimes_{i=1}^{n} \kappa_i$, a generalization of Equation (2.7).*

Daniell-Kolmogorov [31]. The Daniell-Kolmogorov extension theorem extends a measure defined on a sequence of finite product spaces to a measure on an infinite product space. Let $A_1, A_2, \ldots$ be a sequence of measurable spaces and μ_n a measure on the product space $A_1 \times \ldots \times A_n$. We say that the sequence of probability measures

μ_n forms a projective family if $\mu_{n+1}(\cdot \times A_{n+1}) = \mu_n$ for all $n \in \mathbb{N}$. Daniell-Kolmogorov states that if μ_n forms a projective family, then there exists a measure μ on $\prod_{i=1}^{\infty} A_i$ such that μ_n is equal to the projection of μ onto $\prod_{i=1}^{n} A_i$.

Space of extended trees. Let $(\Delta, \mathcal{D}, P)$ be the probability space of elements of the form

$$\delta = (Z, \phi_1, \varrho_1, \ldots, \phi_Z, \varrho_Z).$$

An element of this space carries information about the node to which it is attached: it contains the number of children of the node (random variable ν) and the operators attached to each of its branches. The probability measure $\kappa_1 = P$ lets us build the sample space for first generation of the tree, denoted by $K^1 = \Delta$. We define K^j the sample space of the j-th generation of the tree by

$$K^j := \{\{\delta(i)\} \mid i = 1, \ldots, Z_j \quad \delta(i) \in \Delta \quad Z_j \in \{1, 2, 3, \ldots\}\}.$$

The σ-algebra associated with K^j is

$$\mathcal{D}_j = \sigma\Big(\bigcup_{k \geqslant 1} \mathcal{D}^k\Big) = \{d_1 \cup d_2 \cup \ldots | d_i \in \mathcal{D}^i\} \tag{2.8}$$

where $\mathcal{D}^k = \mathcal{D} \times \ldots \times \mathcal{D}$ k times. In (2.8), note that the right hand side does not depend on j. This comes from the definition of K^j which is the same for all $j \geqslant 2$. The σ-algebra attached to each K^j is therefore the same. Also, we need to consider the smallest σ-algebra spanned by the union of $\mathcal{D}^k$ since the union of σ-algebras is not in general a σ-algebra.

The construction of κ_2 supposes we know the first generation and in particular its size Z_1. For $d = d_1 \cup d_2 \cup \ldots \in \mathcal{D}_2$, with $d_j = E_1^j \times \ldots \times E_j^j \in \mathcal{D}^j$, we define

$$\kappa_2(d|Z_1) = \prod_{i=1}^{Z_1} P(E_i^{Z_1}). \tag{2.9}$$

Sets d_i for $i \neq Z_1$ therefore receive a zero measure. This is illustrated in Figure 2.4. By taking the product of $P(E_i^{Z_1})$ we ensure independence from one node of the tree to the next.

The procedure for constructing κ_2 is repeated n times to build a probability measure on the first n generations $\prod_{i=1}^{n} K^i$, from Ionescu-Tulcea. Then, Daniell-Kolmogorov let us extend the measure to infinite trees since by construction $\bigotimes_{i=1}^{n} \kappa_i$ forms a projective family. Let K be the infinite product space, $\mathcal{K}$ its σ-algebra and κ the probability distribution over this space.

Definition 3. *$(K, \mathcal{K}, \kappa)$ is the probability space of extended Galton-Watson trees.*

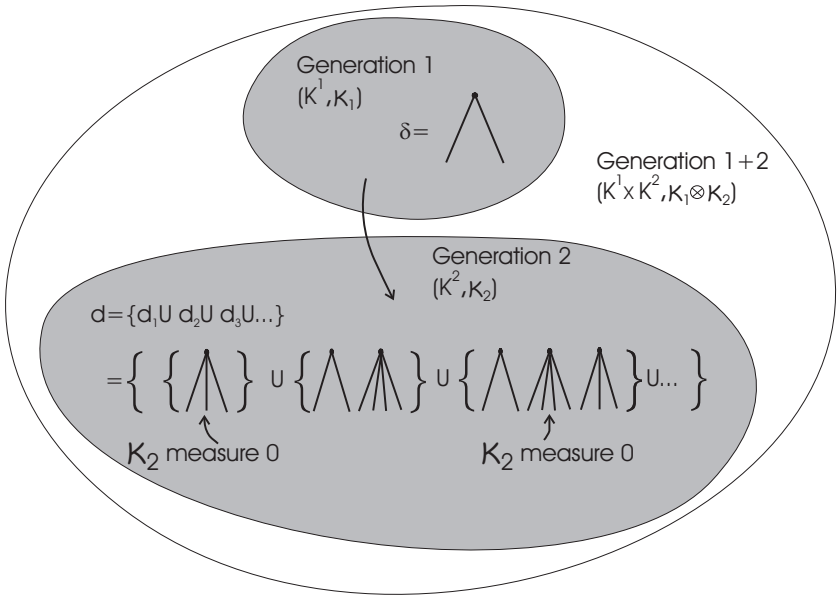


Figure 2.4: Spaces K^1, K^2 and $K^1 \times K^2$ with their respective probability measures κ_1, κ_2 and $\kappa_1 \otimes \kappa_2$. $\delta \in K^1$ has 2 children. Conditionally on δ, only $d_2 \in K^2$ represented here has non-zero measure as it is the only element composed of 2 families. κ_2 assigns measure to each family in d_2 independently. To keep the figure simple, operators attached to the branches of the tree are not represented.

By construction, extended trees are Galton-Watson trees whose branches are marked with random operators. We use classical notation to label nodes and branches of the tree: let $\emptyset$ be the root of the tree and $Z_\emptyset$ be the number of branches rooted at $\emptyset$. Then each node coming from the root is denoted by i, for $i = 1, \ldots, Z_\emptyset$. The second generation of the tree is denoted ij for $1 \leqslant j \leqslant Z_i$. More generally, a node $\mathbf{i}$ is an element of $U = \bigcup_{n \geqslant 0} \mathbb{N}^{*n}$, where $\mathbb{N}^* := \{1, 2, \ldots\}$, and a branch is a couple of nodes $(\mathbf{i}, \mathbf{i}j)$ where $\mathbf{i} \in U$ and j is a strictly positive integer. Lastly, we consider $k_\mathbf{i}$ the subtree of $k \in K$ rooted at $\mathbf{i}$: $k_\mathbf{i} = \{\mathbf{j} \mid \mathbf{j} \in U \text{ and } \mathbf{ij} \in k\}$. By construction, the random variables k_i, $i = 1 \ldots Z_\emptyset$, are independent and identically distributed (Equation (2.9)).

To be consistent with the fact that the fixed point lies at the root of its construction tree, we write $Z_\emptyset$ for Z in (2.6) for the remainder of this chapter.

2.4 Existence and Uniqueness of a fixed point

This section makes precise the conditions under which the Galton-Watson IFS defined in Equation (5.7) possesses a unique fixed point.

THEOREM 4. *Let $(K, \mathcal{K}, \kappa)$ be the space of extended trees and define $\mathbb{L}_p$ using $(\Sigma, \mathcal{F}, P) = (K, \mathcal{K}, \kappa)$. If $\mathbb{E} \sum_{j=1}^{Z_\emptyset} r_j \int |\phi_j(0,x)|^p dx < +\infty$ for some $1 < p < +\infty$, where r_j is the contractive factor of ϱ_j with $0 < r_j < 1$ almost surely, each $\phi_j(.,.)$ is a.s. Lipschitz*

in its first variable, with Lipschitz constant s_j *and and* $\lambda_p = \mathbb{E}\sum_{j=1}^{Z_\emptyset} r_j s_j^p < 1$ *, where* $\mathbb{E}$ *denotes the expectation under* κ*, there exists a unique function* f^* *which satisfies* $f^* = Tf^*$ *in* $\mathbb{L}_p$*. Moreover, for all* $f_0 \in \mathbb{L}_p(\mathbb{X})$*,*

$$d_p^*(T^n f_0, f^*) \leqslant \frac{\lambda_p^{n/p}}{1-\lambda_p^{1/p}} d_p^*(f_0, Tf_0) \tag{2.10}$$

which tends to 0 *as* $n \to +\infty$*.*

Proof. The proof is in two steps. We first need to check that $\mathbb{L}_p$ is closed under T. Next, we have to show that T is contractive in the complete metric space $(\mathbb{L}_p, d_p^*)$. The Banach fixed point theorem will ensure the existence and uniqueness of a limit function in $\mathbb{L}_p$.

Let $f \in \mathbb{L}_p$. We make explicit the construction of i.i.d. copies of $f \in \mathbb{L}_p$. Using notations of section 2.2.2, write f_k for the realisation of f at point $k \in K$, then define $f_k^{(j)}$ by $f_k^{(j)} := f_{k_j}$. Since the random variables k_j are i.i.d., so are the functions $f_k^{(j)}$.

First step. Let $f \in \mathbb{L}_p$. We want to show that $Tf \in \mathbb{L}_p$, or equivalently $\mathbb{E}\int_{\mathbb{X}} |(Tf)(x)|^p dx < +\infty$. To this end, first notice that in the expression (2.6) of Tf, the indicator function partitions $\mathbb{X}$ into disjoint subintervals, so that the absolute value of the sum equals the sum of absolute values. Thus

$$\mathbb{E}\int_{\mathbb{X}} |(Tf)(x)|^p dx = \mathbb{E}\sum_{j=1}^{Z_\emptyset} \int_{\varrho_j(\mathbb{X})} |\phi_j[f^{(j)}(\varrho_j^{-1}(x)), \varrho_j^{-1}(x))]|^p dx.$$

Since the ϱ_j are affine with contraction factor $0 < r_j < 1$, its inverse is also affine with almost everywhere existing Jacobian, and we can perform the change of variable $y = \varrho_j^{-1}(x)$. We bound the Jacobian of the transformation by r_j, the contraction factor of ϱ_j. $\mathbb{E}\int_{\mathbb{X}} |(Tf)(x)|^p dx$ is therefore bounded above by:

$$\mathbb{E}\sum_{j=1}^{Z_\emptyset} r_j \mathbb{E}[\int_{\mathbb{X}} |\phi_j[f^{(j)}(y), y]|^p dy \big| \phi_j]. \tag{2.11}$$

In (2.11), we have also used the law of total probability

$$\mathbb{E}(\cdot) = \mathbb{E}[\mathbb{E}(\cdot|\{Z_\emptyset, \{\phi_j, \varrho_j\}\})]$$

where the second expectation is conditioned on the IFS parameters. Terms depending on $Z_\emptyset$ and ϱ_j can be put outside the second expectation, leaving us with a term which only depends on ϕ_j, hence (2.11). Note that the term in the inner expectation does not depend any more on the ϱ_i's and is just $d_p^{*p}(\phi_j[f^{(j)}, Id], 0)$ after conditioning on the IFS parameters. Id stands for the identity function and 0 is the zero function. Using the triangle inequality, and the fact that for any reals x and y we have $|x+y|^p \leqslant 2^p(|x|^p + |y|^p)$ it follows that $\mathbb{E}\int_{\mathbb{X}} |(Tf)(x)|^p dx$ is bounded by:

$$2^p \mathbb{E}\sum_{j=1}^{Z_\emptyset} r_j d_p^{*p}(\phi_j[f_\kappa^{(j)}, Id], \phi_j[0, Id]) + 2^p \mathbb{E}\sum_{j=1}^{Z_\emptyset} r_j d_p^{*p}(\phi_j[0, Id], 0). \tag{2.12}$$

Using the Lipschitz property of the ϕ_j, the first term of (2.12) is smaller than

$$2^p \mathbb{E} \sum_{j=1}^{Z_\emptyset} r_j s_j^p d_p^{*p}(f_\kappa^{(j)}, 0)$$

which is bounded since $f \in \mathbb{L}_p$. The second term of (2.12) is proportional to $\mathbb{E} \sum_{j=1}^{Z_\emptyset} r_j \int |\phi_j(0,x)|^p dx$ and is finite by assumption. $\mathbb{E} \int_{\mathbb{X}} |(Tf)(x)|^p dx < +\infty$ follows.

Second step. We now prove the contractive property of T under the conditions of Theorem 4. Take f and g in $\mathbb{L}_p$ and consider $d_p^{*p}(Tf, Tg)$. As in step 1, we expand expressions of Tf and Tg and swap the sum and absolute value, we then use the law of total probability and perform the change of variable $y = \varrho_j^{-1}(x)$, whose Jacobian is bounded by r_j. We obtain

$$\begin{aligned} d_p^{*p}(Tf, Tg) &= \mathbb{E} \int \Big| (Tf)(x) - (Tg)(x) \Big|^p dx \\ &\leqslant \mathbb{E} \sum_{j=1}^{Z_\emptyset} r_j \mathbb{E}^* \Big[\int_{\mathbb{X}} \Big| \phi_j[f^{(j)}(y)), y] - \phi_j[g^{(j)}(y), y] \Big|^p dy \Big] \end{aligned}$$

where $\mathbb{E}^* = \mathbb{E}[\cdot | \{Z_\emptyset, \{\phi_i, \varrho_i\}\}]$. Lastly, we use the Lipschitz property of the non linear random maps ϕ_j to conclude that

$$d_p^{*p}(Tf, Tg) \leqslant \lambda_p d_p^{*p}(f, g)$$

where the definition of λ_p is given in the theorem statement. Under the assumption $\lambda_p < 1$, the contractive property follows and from the Banach fixed point theorem there exists a unique function f^*, attractor of the Galton-Watson IFS. Moreover,

$$d_p^{*p}(T^n f_0, f^*) \leqslant \lambda_p d_p^{*p}(T^{n-1} f_0, f^*)$$

which leads to

$$d_p^{*}(T^n f_0, f^*) \leqslant \lambda_p^{n/p} d_p^{*}(f_0, f^*).$$

Now using the triangle inequality:

$$d_p^{*}(f_0, f^*) \leqslant d_p^{*}(f_0, Tf_0) + \lambda_p^{1/p} d_p^{*}(f_0, f^*)$$

so that

$$d_p^{*}(T^n f_0, f^*) \leqslant \frac{\lambda_p^{n/p}}{1 - \lambda_p^{1/p}} d_p^{*}(f_0, Tf_0)$$

which concludes the proof of the theorem. □

To illustrate, we present in Figure 2.5 a snapshot of the fixed point of a certain IFS and its mean. The IFS parameters are detailed in the Figure caption.

The theorem not only states that starting from an initial function the IFS converges in $\mathbb{L}_p$ to a unique fixed point under the metric d_p^* but also that the convergence is exponential. It follows that the convergence of $T^n f_0$ towards f^* is almost sure. To show this, let $\epsilon > 0$, then

$$P(d_p^p(T^n f_0, f^*) > \epsilon) \leqslant \frac{\mathbb{E} d_p^p(T^n f_0, f^*)}{\epsilon} \leqslant C\lambda_p^n$$

where

$$C = \frac{d_p^{*p}(f_0, Tf_0)}{\epsilon(1-\lambda_p^{1/p})^p}.$$

It follows that

$$\sum_{n \geqslant 1} P(d_p^p(T^n f, f^*) > \epsilon) < \infty$$

and from Borel-Cantelli lemma we have P-almost sure convergence.

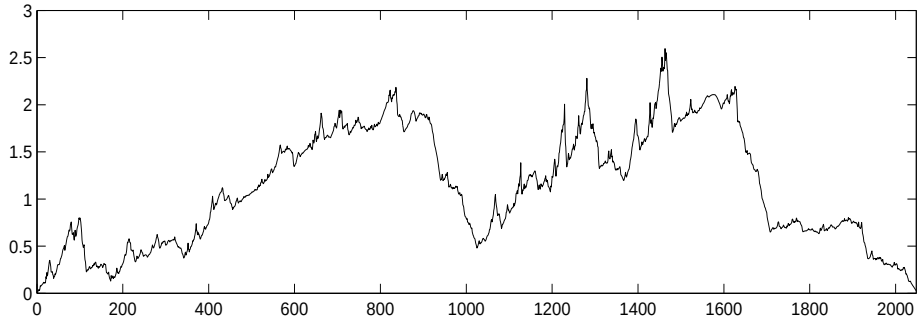
f^* is the unique fixed point for which $f^* = Tf^*$ in $\mathbb{L}_p$ but there may be some other $f^0 \neq f^*$ such that the law of f^0 equals the law of Tf^0. The following result can be proven in the same way as Hutchinson and Rüschendorff [63].

Corollary 1. *The distribution of f^* is the unique distribution which satisfies $f^* \stackrel{d}{=} Tf^*$, where $\stackrel{d}{=}$ denotes equality in distribution.*

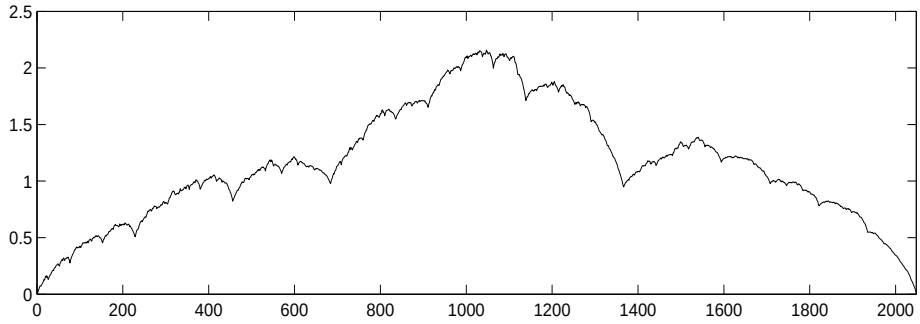
The idea is to define a new space of probability distributions of elements of $\mathbb{L}_p$ and a new metric over this space which, leads to a complete metric space. Then one can prove that the operator T seen at the distribution level is contractive in this space and therefore admits a unique fixed point.

2.5 Properties of the fixed point

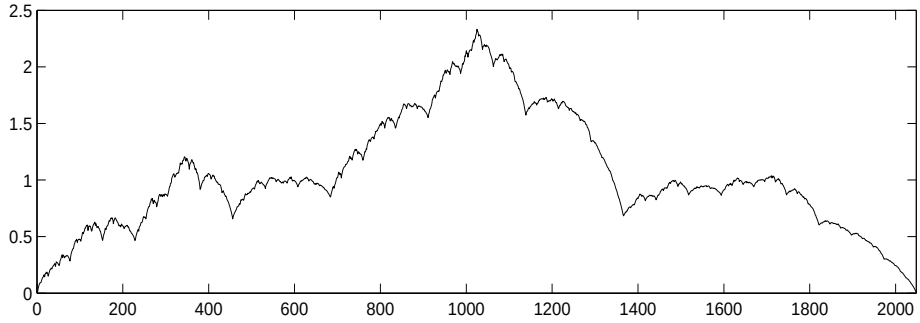
We now consider two properties of the fixed point. First, we derive conditions under which paths are a.s. continuous. Then, we look at the moments of the fixed point and show that under certain assumptions, moments of the attractor continuously depend on the probability vector $\mathbf{q}$. This fact is suggested by observing the Figure 2.5 where a small change in $\mathbf{q}$ induces 'small' variations in the mean of the fixed point. Indeed, the mean of the fixed point in the middle figure is obtained for $\mathbf{q}_1 = (0.2, 0.3, 0.5)$, and the bottom figure shows the mean of another Galton-Watson IFS with the same parameters, but with $\mathbf{q}_2 = (0.2, 0.2, 0.6)$, so that the distance (as defined below) between $\mathbf{q}_1$ and $\mathbf{q}_2$ is small. Then one can notice that the two means look quite similar, so that changing $\mathbf{q}_1$ by a small amount does not modify too much the shape of the mean of the fixed point. This observation will be made precise in Section 2.5.2.



(a)



(b)



(c)

Figure 2.5: A realization of the fixed point (a) and its mean (b). ϕ_j are decomposed as follows $\phi_j(x,t) = s_j x + X\zeta_j(t)$ for $j = 1, \ldots, Z_\emptyset$ where X is normally distributed with mean 1 and variance 0.25. When $Z_\emptyset = 1$, $s_1 = 0.6$ and $\zeta_1(t) = t(1-t)$. For $Z_\emptyset = 2$ we define $s_1 = 0.6$, $s_2 = 0.7$, $\zeta_1(t) = t^3$, $\zeta_2(t) = 1 - t^2$ and for $Z_\emptyset = 3$ we have $s_1 = 0.6$, $s_2 = 0.7$, $s_3 = 0.3$, $\zeta_1(t) = t^4$, $\zeta_2(t) = (t+1)(1 - 0.75t^3)$ and $\zeta_3(t) = 0.5(1-t^2)$. $Z_\emptyset$ takes the values 1, 2 or 3 with probabilities 0.2, 0.3 and 0.5 for the first 2 figures. The bottom figure is the mean obtained with the probabilities 0.2, 0.2 and 0.6.

2.5.1 Continuity of the sample paths

The results for Galton-Watson IFS are a straight-forward generalization of continuity results in the deterministic setting.

PROPOSITION 1. $\mathbb{X} = [a, b]$. *Let α be the unique random fixed point of $\phi_1(., a)$ and β the unique random fixed point of $\phi_{Z_\emptyset}(., b)$: $\phi_1(\alpha, a) = \alpha$ and $\phi_{Z_\emptyset}(\beta, b) = \beta$. Assume that α and β are the same for all possible realisations of ϕ_1 and $\phi_{Z_\emptyset}$. If $\phi_i(\beta, b) = \phi_{i+1}(\alpha, a)$ a.s. for all $i \in \{1, \ldots, Z_\emptyset - 1\}$ and all the operators considered are continuous, then f^* has continuous paths and $f^*(a) = \alpha$ and $f^*(b) = \beta$ (a.s.).*

Proof. We first note that $f^*(a)$ and $f^*(b)$ are respectively fixed points of ϕ_1 and $\phi_{Z_\emptyset}$:

$$\begin{aligned} f^*(a) &= \phi_1[f^*(\varrho_1^{-1}(a)), \varrho_1^{-1}(a)] = \phi_1[f^*(a), a] \\ f^*(b) &= \phi_{Z_\emptyset}[f^*(\varrho_{Z_\emptyset}^{-1}(b)), \varrho_{Z_\emptyset}^{-1}(b)] = \phi_{Z_\emptyset}[f^*(b), b]. \end{aligned}$$

Those equalities have to remain true whatever $Z_\emptyset$ is, which is realized under the assumption of Proposition 1.

Let $\varrho_i[a, b] = [a_{i-1}, a_i]$ for $i \in \{1, \ldots, Z_\emptyset\}$ and $a_0 = a$, $a_{Z_\emptyset} = b$ almost surely. We only have to prove the continuity at the random points a_i of the interval $[a, b]$ since we consider continuous operators and d_p^* is complete on the set of continuous functions [63]. Therefore, if the n-th iterate of T is continuous, the limit process also belongs to the space of continuous functions.

$f^*(a_i)$ can be expressed in two different ways as the point a_i is at the intersection of $\varrho_i[a, b]$ with $\varrho_{i+1}[a, b]$:

$$f^*(a_i) = \phi_i[f^*(\varrho_i^{-1}(a_i)), \varrho_i^{-1}(a_i)] = \phi_i[f^*(b), b] = \phi_i[\beta, b]. \tag{2.13}$$

We can show in a similar way that $f^*(a_i) = \phi_{i+1}[\alpha, a]$. Under the condition of the proposition the continuity of f^* at points a_i follows. □

With this model, it is possible to obtain continuous paths or random processes everywhere discontinuous by adjusting the IFS parameters. Allowing only one discontinuity by not joining two operators $\phi_{Z_\emptyset,i}$ and $\phi_{Z_\emptyset,i+1}$ at the random point a_i will result in an everywhere discontinuous fixed point. A snapshot of a continuous fixed point is represented in Figure 2.5.

2.5.2 Continuous dependency w.r.t. q

The continuity of the moments of the fixed point with respect to $\mathbf{q}$ is suggested in Figure 2.5. This observation is related to the one made by Barnsley in [14] where the attractor of a deterministic IFS is continuously varying with respect to the IFS parameters, leading to applications in image synthesis. We prove the result here for the model presented in [37], with deterministic maps and a random tree.

Consider the set of deterministic maps

$$\{\{\phi_{k,1}, \ldots, \phi_{k,k}, \varrho_{k,1}, \ldots, \varrho_{k,k}\}\}_{k=1,2,\ldots}.$$

Given $Z_\emptyset = j$, then apply $\{\phi_{j,1}, \ldots, \phi_{j,j}, \varrho_{j,1}, \ldots, \varrho_{j,j}\}$. $\phi_{k,j}$ and $\varrho_{k,j}$ may differ for different values of k, $j = 1, \ldots, k$. The operator T becomes:

$$(Tf)(x) = \sum_{j=1}^{Z_\emptyset} \phi_{Z_\emptyset,j}[f^{(j)}(\varrho_{Z_\emptyset,j}^{-1}(x)), \varrho_{Z_\emptyset,j}^{-1}(x))]\mathbf{1}_{\varrho_{Z_\emptyset,j}(\mathbb{X})}(x). \tag{2.14}$$

Lipschitz factor of $\phi_{Z_\emptyset,j}$ is $s_{Z_\emptyset,j}$ and the contraction factor of $\varrho_{Z_\emptyset,j}$ is denoted by $r_{Z_\emptyset,j}$. Since the operators attached to the branches of the tree are the same for a given number of offsprings, to one realisation of the tree is associated one and only one realisation of the fixed point.

THEOREM 5. *Suppose conditions of Theorem 4 hold. Let $f^* \in \mathbb{L}_p$ be the fixed point of a Galton-Watson IFS, bounded number of offspring and deterministic maps of the form $\phi(u, v) = su + \zeta(v)$, where $0 \leqslant s < 1$ and ζ is a nonlinear function. Suppose that $\lambda_r = \mathbb{E}\sum_{j=1}^{Z_\emptyset} r_{Z_\emptyset,j} s_{Z_\emptyset,j}^r < 1$ for $r = 1, \ldots, p$. Then the r-th moment of f^* continuously varies with respect to the probability generating vector $\mathbf{q}$, for $r = 1, \ldots, p$.*

Proof. We prove the theorem by induction on r, for $r = 1, \ldots, p$. The first step of the proof shows that the continuity property holds for the mean of the fixed point. In the second step, we generalize it to any higher order integer moment. Let P_v be the space of probability vectors

$$P_v = \{\mathbf{p} = (p_i, i \in \mathbb{N}^*) \mid \sum_i p_i = 1\}$$

where $\mathbb{N}^* := \{1, 2, \ldots\}$. This space is endowed with the metric $l(\mathbf{p}, \mathbf{q}) = \sum_i |p_i - q_i|$.

First step. By definition of $\mathbb{L}_p$, $\mathbb{E}_\mathbf{p} f_\mathbf{p}^* \in L_1$, if we denote by $f_\mathbf{p}^*$ the fixed point of the IFS with probability vector $\mathbf{p}$ and by $\mathbb{E}_\mathbf{p}$ the expectation under $\kappa_\mathbf{p}$, the probability measure defined on K with probability generating vector $\mathbf{p}$. We adopt this notation in this section to emphasize the dependence on $\mathbf{p}$. Note that by changing the probability vector, we change the measure $\kappa_\mathbf{p}$ on the space of extended Galton Watson trees. Therefore, if we now call $f_\mathbf{q}^*$ the fixed point of the same IFS with probability generating vector $\mathbf{q}$, the expectation with respect to this new measure is different from $\mathbb{E}_\mathbf{p}$ and we denote it by $\mathbb{E}_\mathbf{q}$ (expectation under the new measure $\kappa_\mathbf{q}$). The continuity of the mean of the fixed point with respect to the generating vector follows if we show the continuity of the map $\psi : P_v \to L_1$ which associates with each probability vector the mean of the fixed point of the Galton-Watson IFS. Let $\mathbf{p} \in P_v$, we want to show that for all $\epsilon > 0$, there exists $\eta > 0$ such that

$$\forall \mathbf{q} \in P_v \;\; l(\mathbf{p}, \mathbf{q}) \leqslant \eta \;\Rightarrow\; d_1(\mathbb{E}_\mathbf{p} f_\mathbf{p}^*, \mathbb{E}_\mathbf{q} f_\mathbf{q}^*) \leqslant \epsilon. \tag{2.15}$$

Let $\epsilon > 0$ and $\mathbf{p} \in P_v$. We first use the fact that f^* and Tf^* have the same distribution, therefore the same mean:

$d_1(\mathbb{E}_\mathbf{p} f_\mathbf{p}^*, \mathbb{E}_\mathbf{q} f_\mathbf{q}^*) =$

$$\int |\mathbb{E}_{\mathbf{p}} \sum_{j=1}^{Z^1_\emptyset} \phi_{Z^1_\emptyset,j}[f^*_{\mathbf{p}} \circ \varrho^{-1}_{Z^1_\emptyset,j}, \varrho^{-1}_{Z^1_\emptyset,j}]\mathbf{1}_{\varrho_{Z^1_\emptyset,j}(\mathbb{X})} - \mathbb{E}_{\mathbf{q}} \sum_{j=1}^{Z^2_\emptyset} \phi_{Z^2_\emptyset,j}[f^*_{\mathbf{q}} \circ \varrho^{-1}_{Z^2_\emptyset,j}, \varrho^{-1}_{Z^2_\emptyset,j}]\mathbf{1}_{\varrho_{Z^2_\emptyset,j}(\mathbb{X})}|$$

where $\kappa_{\mathbf{p}}(Z^1_\emptyset = k) = p_k$ and $\kappa_{\mathbf{q}}(Z^2_\emptyset = k) = q_k$. We omit the variable x in the integrand to keep the notation clear. By conditioning with respect to $Z^1_\emptyset$ and $Z^2_\emptyset$, the right hand side becomes:

$$\int |\sum_{i\geqslant 1} p_i \sum_{j=1}^{i} \mathbb{E}_{\mathbf{p}} \phi_{i,j}[f^*_{\mathbf{p}} \circ \varrho^{-1}_{i,j}, \varrho^{-1}_{i,j}]\mathbf{1}_{\varrho_{i,j}(\mathbb{X})} - \sum_{i\geqslant 1} q_i \sum_{j=1}^{i} \mathbb{E}_{\mathbf{q}} \phi_{i,j}[f^*_{\mathbf{q}} \circ \varrho^{-1}_{i,j}, \varrho^{-1}_{i,j}]\mathbf{1}_{\varrho_{i,j}(\mathbb{X})}|.$$

The sums can be taken outside the integral. By setting $y = \varrho^{-1}_{i,j}(x)$ and bounding the Jacobian by $r_{i,j}$, where $r_{i,j}$ is deterministic here as we consider non random maps, $d_1(\mathbb{E}_{\mathbf{p}} f^*_{\mathbf{p}}, \mathbb{E}_{\mathbf{q}} f^*_{\mathbf{q}})$ is less than

$$\sum_{i,j} r_{i,j} \int |p_i \mathbb{E}_{\mathbf{p}} \phi_{i,j}[f^*_{\mathbf{p}}(y), y] - q_i \mathbb{E}_{\mathbf{q}} \phi_{i,j}[f^*_{\mathbf{q}}(y), y]| dy.$$

By assumption $\phi_{i,j}(u, v) = s_{i,j}u + \zeta_{i,j}(v)$. The Lipschitz factor of $\phi_{i,j}$ is $s_{i,j}$ in this case. Using the triangle inequality of $|\cdot|$ it follows that $d_1(\mathbb{E}_{\mathbf{p}} f^*_{\mathbf{p}}, \mathbb{E}_{\mathbf{q}} f^*_{\mathbf{q}})$ is bounded by

$$\sum_{i,j} r_{i,j} \Big[\int s_{i,j} |p_i \mathbb{E}_{\mathbf{p}} f^*_{\mathbf{p}} - q_i \mathbb{E}_{\mathbf{q}} f^*_{\mathbf{q}}| + |p_i - q_i||\zeta_{i,j}(y)| dy \Big].$$

The term $|p_i \mathbb{E}_{\mathbf{p}} f^*_{\mathbf{p}} - q_i \mathbb{E}_{\mathbf{q}} f^*_{\mathbf{q}}|$ can be further bounded above by $|\mathbb{E}_{\mathbf{p}} f^*_{\mathbf{p}}||p_i - q_i| + q_i|\mathbb{E}_{\mathbf{p}} f^*_{\mathbf{p}} - \mathbb{E}_{\mathbf{q}} f^*_{\mathbf{q}}|$ by adding and subtracting $q_i \mathbb{E}_{\mathbf{p}} f^*_{\mathbf{p}}$ and using the triangle inequality. Suppose $\mathbf{p}$ and $\mathbf{q}$ are chosen such that $l(\mathbf{p}, \mathbf{q}) \leqslant \eta$. We have

$$d_1(\mathbb{E}_{\mathbf{p}} f^*_{\mathbf{p}}, \mathbb{E}_{\mathbf{q}} f^*_{\mathbf{q}}) \leqslant \eta \textstyle\sum_{i,j} r_{i,j} s_{i,j} \int |\mathbb{E}_{\mathbf{p}} f^*_{\mathbf{p}}| + \\ \textstyle\sum_{i,j} q_i r_{i,j} s_{i,j} \int |\mathbb{E}_{\mathbf{p}} f^*_{\mathbf{p}} - \mathbb{E}_{\mathbf{q}} f^*_{\mathbf{q}}| + \eta \sum_{i,j} r_{i,j} \int |\zeta_{i,j}(y)| dy.$$

In the first term of the right hand side, $\int |\mathbb{E}_{\mathbf{p}} f^*_{\mathbf{p}}| < M < \infty$ since $\mathbb{E}_{\mathbf{p}} f^*_{\mathbf{p}} \in L_1$. We have a bounded number of maps so $\sum_{i,j} r_{i,j} s_{i,j}$ is also bounded. In the second term, $\int |\mathbb{E}_{\mathbf{p}} f^*_{\mathbf{p}} - \mathbb{E}_{\mathbf{q}} f^*_{\mathbf{q}}|$ is the distance between $\mathbb{E}_{\mathbf{p}} f^*_{\mathbf{p}}$ and $\mathbb{E}_{\mathbf{q}} f^*_{\mathbf{q}}$ and $\sum_{i,j} q_i r_{i,j} s_{i,j}$ is exactly the contraction factor $\lambda_1(\mathbf{q}) < 1$ of the map T. Since the map $\mathbf{p} \mapsto \sum p_i r_{i,j} s_{i,j}$ is linear, it follows that the distance between $\lambda_1(\mathbf{q})$ and $\lambda_1(\mathbf{p})$ is small for $\mathbf{p}$ sufficiently close to $\mathbf{q}$. Therefore $\lambda_1(\mathbf{q}) \leqslant 1 - \varepsilon_p$ for some small $\varepsilon_p > 0$. Finally, the third term is bounded by assumption. It follows that $d_1(\mathbb{E}_{\mathbf{p}} f^*_{\mathbf{p}}, \mathbb{E}_{\mathbf{q}} f^*_{\mathbf{q}})$ is smaller than

$$\frac{\eta}{\varepsilon_p} \Big[M \sum_{i,j} r_{i,j} s_{i,j} + \sum_{i,j} r_{i,j} \int |\zeta_{i,j}(y)| dy \Big] := \frac{\eta}{\varepsilon_p} \gamma.$$

Set $\eta = \frac{\varepsilon_p}{\gamma} \epsilon$, it follows that $d_1(\mathbb{E}_{\mathbf{p}} f^*_{\mathbf{p}}, \mathbb{E}_{\mathbf{q}} f^*_{\mathbf{q}}) \leqslant \epsilon$.

Second step. We now show that the previous result holds for the r-th moment of the fixed point, as long as $r \leqslant p$. To fix ideas, let us start with the second order

moment. First note that f^{*2} is distributed like $(Tf^*)^2$. Then, proceeding as before we bound $d_1(\mathbb{E}_{\mathbf{p}} f^{*2}_{\mathbf{p}}, \mathbb{E}_{\mathbf{q}} f^{*2}_{\mathbf{q}})$ by

$$\sum_{i,j} r_{i,j} \int |p_i \mathbb{E}_{\mathbf{p}} \phi^2_{i,j}[f^*_{\mathbf{p}}(y), y] - q_i \mathbb{E}_{\mathbf{q}} \phi^2_{i,j}[f^*_{\mathbf{q}}(y), y]| dy.$$

By developing the square, the following terms appear: $s^2_{i,j} f^{*2}$, $\zeta^2_{i,j}(y)$ and $2 s_{i,j} \zeta_{i,j}(y)$. It follows that

$d_1(\mathbb{E}_{\mathbf{p}} f^{*2}_{\mathbf{p}}, \mathbb{E}_{\mathbf{q}} f^{*2}_{\mathbf{q}}) \leqslant$

$$\textstyle\sum_{i,j} r_{i,j} \Big[\int s^2_{i,j} |p_i \mathbb{E}_{\mathbf{p}} f^{*2}_{\mathbf{p}} - q_i \mathbb{E}_{\mathbf{q}} f^{*2}_{\mathbf{q}}| + 2 s_{i,j} \zeta_{i,j}(y) |p_i \mathbb{E}_{\mathbf{p}} f^*_{\mathbf{p}} - q_i \mathbb{E}_{\mathbf{q}} f^*_{\mathbf{q}}| + \zeta^2_{i,j}(y) dy \Big].$$

First note that for all $r \leqslant p$, $f \in \mathbb{L}_p(\mathbb{X}) \Rightarrow f \in \mathbb{L}_r(\mathbb{X})$. Indeed, notice that $|f(t)|^r = |f(t)|^r \cdot 1$ and apply the Hölder inequality (2.4):

$$\mathbb{E} \int |f(t)|^r dt \leqslant \Big(\mathbb{E}^{1/s} |f(t)|^{rs} dt \Big) \Big(\mathbb{E}^{1/u} \int_{\mathbb{X}} dt \Big) \tag{2.16}$$

where s and u are chosen such that $\frac{1}{s} + \frac{1}{u} = 1$ and $1 < s, u < \infty$. Put $s = p/r$. The right hand side becomes

$$\mathbb{E}^{r/p} \int |f(t)|^p dt.$$

The inequality

$$||f||^*_r \leqslant ||f||^*_p$$

follows and if $f \in \mathbb{L}_p$, then clearly $f \in \mathbb{L}_r$. Note that this result only holds when we consider functions defined on compact intervals. No similar result holds for functions of the real line as the right hand side of (2.16) is infinite.

Back to the proof, it follows that f^* is in $\mathbb{L}_2$ and the last term of the right hand side is bounded above by assumption. The second term of the right hand side is also smaller than η times some constant, as noted previously in step 1 of the proof. In the first term, we proceed as in step 1 and it follows that the right hand side is less than some constant times η. The continuity follows.

When dealing with the r-th order moment, the term $\phi^r_{i,j}[f^*(y), y]$ appears in the bounding factor. Expanding the expression, the terms $\int |p_i \mathbb{E}_{\mathbf{p}} f^{*j}_{\mathbf{p}} - q_i \mathbb{E}_{\mathbf{q}} f^{*j}_{\mathbf{q}}|$ for $1 \leqslant j \leqslant r-1$ show up, all less than some constant times η by the induction hypothesis. The conclusion follows by noting that the triangle inequality can be applied as before to the term $\int |p_i \mathbb{E}_{\mathbf{p}} f^{*r}_{\mathbf{p}} - q_i \mathbb{E}_{\mathbf{q}} f^{*r}_{\mathbf{q}}|$. □

2.5.3 Test for multifractality

So far, we have not investigated the fractal properties of the fixed point of Galton-Watson IFS. In this section, we give empirical results showing the multifractal behaviour of these random signals by estimating their partition function $\zeta_2(q)$ using wavelet leaders. We only perform tests for multifractality, no attempt is made to support those simulations theoretically.

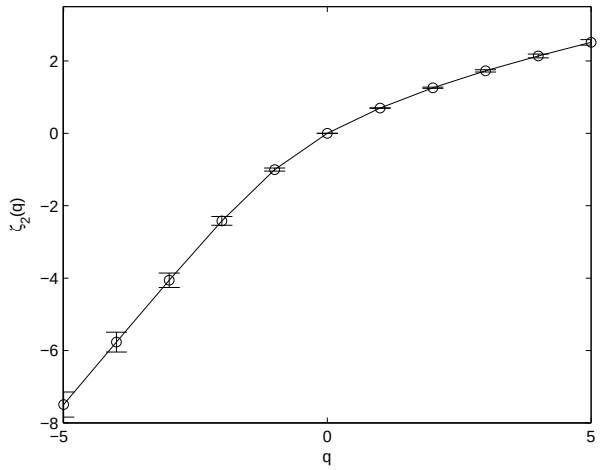


Figure 2.6: Estimation of the partition function $\zeta_2(q)$ of the fixed point of a Galton-Watson IFS using wavelet leaders. The parameters of the IFS are given in the figure caption 2.5, with $Z_\emptyset$ taking values 1, 2 or 3 with probabilities 0.2, 0.3 and 0.5 respectively. The partition function is obtained by averaging 100 estimations. 95% confidence intervals are also plotted.

Consider a Galton-Watson IFS whose characteristics are the same as the one presented in the Figure caption 2.5, with $Z_\emptyset$ taking values 1, 2 or 3 with probabilities 0.2, 0.3 and 0.5 respectively. Figure 2.6 presents the estimated function $\zeta_2(q)$, obtained after averaging over 50 estimations using the wavelet leader technique. The length of each signal is 2^{12}. We have used Daubechies wavelets with 2 vanishing moments and the wavelet scale of analysis ranges from $j_1 = 3$ to $j_2 = 9$. $\zeta_2(q)$ clearly appears non-linear, characteristic of multifractal processes.

To support the previous observation, we perform a multifractality test proposed by Wendt and Abry, using a non parametric bootstrap hypothesis test using wavelet coefficients and leaders [114]. We have used the Matlab toolbox developed by Wendt in his PhD [113]. Consider the polynomial expansion of the partition function $\zeta_2(q) = c_1 q + c_2 q^2/2 + c_3 q^3/3 + \ldots$. When the multifractal formalism holds and when $\zeta_2(q)$ is linear with q (corresponding to $c_p = 0$ for all $p \geqslant 2$), the process is monofractal (see the end of Section 4.1 for more details). Any departure from a linear behaviour is characteristic of a multifractal process. Wendt and Abry have designed a test to decide whether $c_p = 0$ or not. The case $p = 2$ permits to conclude between a mono and multifractal stochastic process.

First, c_p is estimated by linear regression of the log-cumulants of $\log d_X(j,k)$ vs $\log 2^j$ over the range of scales 2^{j_1} to 2^{j_2}. $d_X(j,k)$ is the wavelet coefficient at scale j and position k. $d_X(j,k)$ can be replaced by the wavelet leader $L_X(j,k)$. Then, test $c_p = c_{p,0}$ (null hypothesis) vs the two sided alternative $c_p \neq c_{p,0}$. The basic test is then

$$T = c_p - c_{p,0}.$$

The distribution of the statistic under the null hypothesis is unknown in general and is estimated using non parametric bootstrap techniques. From the empirical

distribution, one can design an acceptance region $\mathcal{T}_{1-\alpha} = [t_{\alpha/2}, t_{1-\alpha/2}]$ where t_α is the α quantile of the null distribution. α is the error rate in rejecting the null hypothesis. During the simulations, $c_{p,0}$ is replaced by its estimated value $\hat{c}_p$. The empirical null distribution is obtained by considering R bootstrap resamples of the wavelet coefficients or leaders which are used to obtain R resamples of the log cumulants $\hat{c}_p^{*,(r)}$, $r = 1 \dots R$. Wendt and Abry proposed six tests in total but we only focus on one using percentiles of the null distribution. We choose arbitrarily a significance level of $\alpha = 0.1$ in our simulations and perform the estimation 100 times to test the null hypothesis for $p = 2$ with the wavelet leaders. We could consider other significance levels. The averaged rejection rate $\hat{\alpha} = 0.99$ obtained in our simulations is close to 1, confirming the multifractal behaviour of the fixed point studied here (the rejection rate would be as high when considering for example $\alpha = 0.05$). This study shows the existence of a class of multifractal fixed points. In a future project, it would be interesting to obtain theoretical results on the spectrum of fixed point of Galton-Watson IFS in terms of its parameters. The study could follow the lines of Jaffard [65], who derived the spectrum of a class of deterministic IFS signals, working in the wavelet domain.

Chapter 3

A multifractal embedded branching process

In this section we present a new class of multifractal processes, constructed using an Embedded Branching Process (EBP), which admits an infinite dimensional Markov representation. Simulation of processes which exhibit long-range dependance is problematic in practice because it is hard in general to simulate $X(n+1)$ given $X(1), \ldots, X(n)$, due to their highly correlated structure. Using a dynamic restriction of the state space of the Markov representation of the process, we are able to simulate $X(n+1)$ given $X(n)$ in $O(\log n)$ operations. Our class of processes includes Brownian motion subjected to a continuous-multifractal time-change, which is constructed from a multiplicative cascade applied to the embedded branching process. The study of the multifractal properties of the time change are postponed to Chapter 4. We present in Section 3.1 the construction of the Canonical EBP process (CEBP) and extend the model in Part 3.2. The simulation algorithm is detailed in section 3.3.2. Finally, in Section 3.4, we imitate an fBm with an MEBP.

3.1 EBP and the crossing tree

3.1.1 Construction of the Canonical Embedded Branching Process (CEBP).

The idea behind embedded branching processes takes its origins to the study of diffusion on fractal sets [13]. As noted in the introduction of this thesis, it is possible to associate a branching process to a random walk on a Sierpinski gasket.

Consider the graph S of a Sierpinski triangle with vertices $(0,0)$, $(1,0)$ and $(0,1)$ whose construction was detailed in Section 1.1.1, considering the initial set to be a triangle with vertices $(0,0)$, $(1,0)$ and $(0,1)$.Define $S_0 = S \cup (S-1)$ where $S - 1 = \{(x-1, y) \mid (x,y) \in S\}$. Start the random walk on the graph S_0 from one vertex, say $(0,0)$, and go to $(0,1)$, $(1,0)$, $(-1,0)$ or $(-1,1)$, each direction with equal probability. This gives a first crossing of size 1, represented as the root node on the crossing tree, in red in Figure 3.1. Then decompose this first crossing of size 1 into a random number of subcrossings of size 1/2. For example, starting at $(0,0)$, we can go to $(1/2,0)$, $(0,1/2)$, $(-1/2,0)$ or $(-1/2,1/2)$ with equal probability.

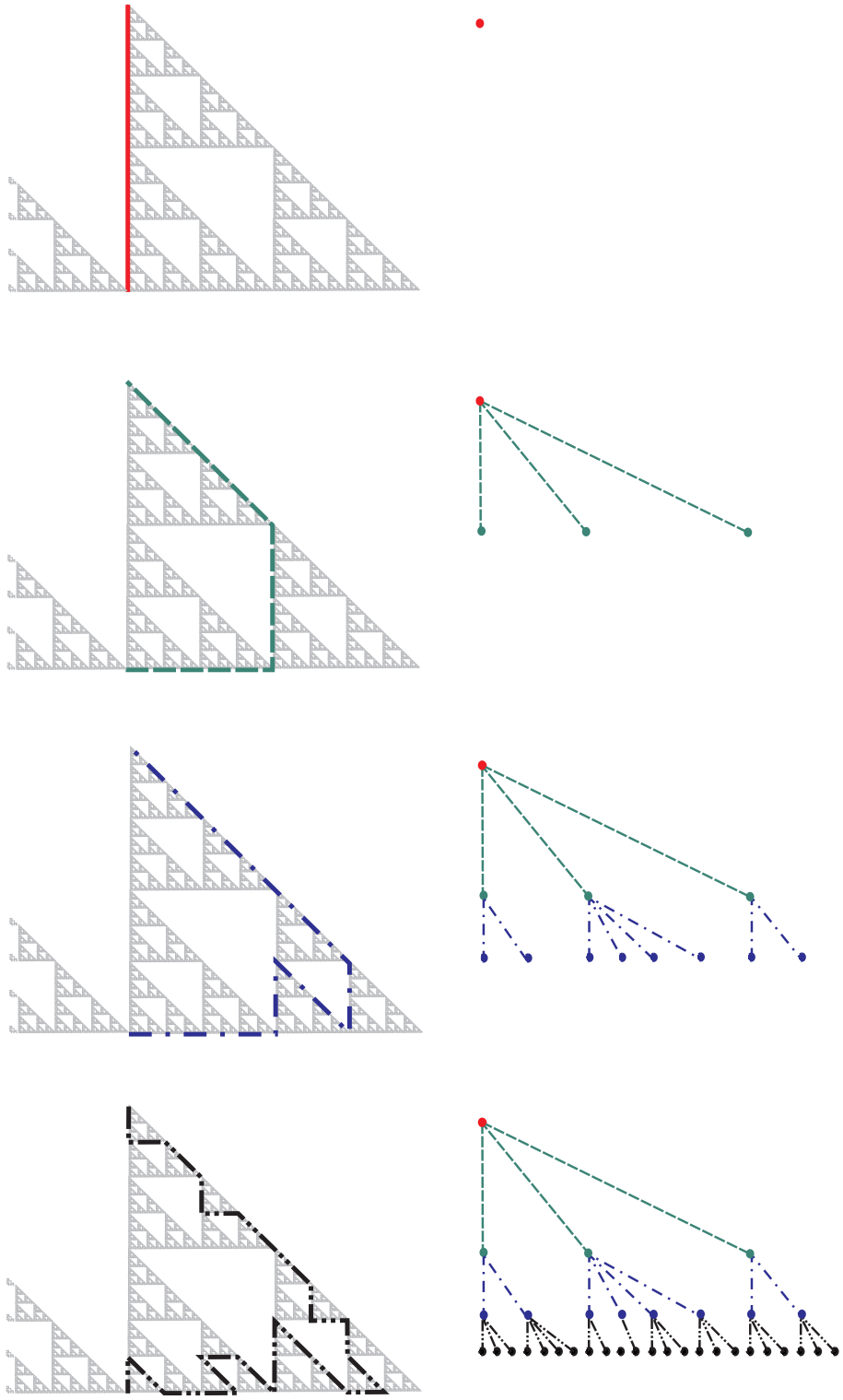


Figure 3.1: Crossing tree of a random walk on a Sierpinski triangle

Figure 3.1 decomposes the first crossing in red into three green subcrossings of size 1/2. Associate vertices of S_0 visited by the random walk with nodes of a tree, and subcrossings with the branches of a tree to obtain the crossing tree of the random walk. In our example, the random walk is given by $\{(0,0),(1/2,0),(1/2,1/2),(0,1)\}$ and the crossing tree after one step consists of a root node (crossing of size 1) and its three children (three subcrossings of size 1/2). Repeat the procedure. CEBP processes are defined in a similar way with their crossing times. Their construction is now given.

Let $X : \mathbb{R}^+ \to \mathbb{R}$ be a continuous process, with $X(0) = 0$. For $n \in \mathbb{Z}$ we define level-n crossing times T_k^n by putting $T_0^n = 0$ and

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

Definition 4. *The k-th level-n (equivalently scale 2^n) crossing C_k^n is the bit of sample path from T_{k-1}^n to T_k^n plus the extra information of the time and place the crossing starts:*

$$C_k^n := \{(t, X(t)) \mid T_{k-1}^n \leqslant t < T_k^n\}$$

This is illustrated in Figure 3.2, where the level 3, 4 and 5 crossings of a given sample path are shown. When passing from a coarse scale to a finer one, we decompose each level n crossing into a sequence of level $n-1$ crossings. To define the crossing tree, we associate nodes with crossings, and the children of a node are its subcrossings. This is illustrated in Figure 3.2.

The crossing tree is an efficient way of representing a signal, and can also be used for inference. In [71] the crossing tree is used to test for self-similarity and to obtain an asymptotically consistent estimator of the Hurst index of a self-similar process with stationary increments, and in [72] it is used to test for stationarity.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be the probability space of extended Galton-Watson trees. In this chapter and the following one, by extended Galton-Watson tree we mean a Galton-Watson tree whose branches are endowed with random variables, which may depend on the size of the generation. In fact, we will use later on results on multitype Galton-Watson processes. Its construction is similar to the probability space $(K, \mathcal{K}, \kappa)$ constructed in section 2.3 of the previous chapter.

Let the random variable Z_k^n be the number of level-$(n-1)$ subcrossings that make up the k-th level-n crossing. The orientation of a crossing can be up or down. Subcrossings consist of excursions (up-down and down-up pairs) followed by a direct crossing (down-down or up-up pairs), whose direction depends on the parent crossing: if the parent crossing is up, then the subcrossings end up-up, if the parent is down, they end down-down.

Definition 5. *A continuous process X is called an Embedded Branching Process (EBP) process if the random variables Z_k^n are independent and identically distributed.*

Given an EBP process, let $p(x) = \mathbb{P}(Z_k^n = x)$ be the offspring distribution, noting that Z_k^n takes values in $2\mathbb{N}$ since subcrossings come by pairs.

Definition 6. *p is said to be regular if $p(2) < 1$ and $\sum_x x \log(x) p(x) < \infty$.*

Figure 3.2: A process and levels $n = 3$, 4 and 5 of its crossing tree. In the top frame we have joined the points T_k^n at each level, and in the bottom frame we have identified the k-th level-n crossing with the point $(2^n, T_{k-1}^n)$ and linked each crossing to its subcrossings.

Let α_k^n be a vector of size Z_k^n whose components are the types of the Z_k^n subcrossings of C_k^n. Clearly, each of the $(Z_k^n - 2)$ first entries of α_k^n come by pair, each pair being up-down or down-up. The last two components are either the pair up-up or down-down. Let $p_{c|z}(\cdot) = \mathbb{P}(\alpha_k^n = \cdot \mid Z_k^n = z)$ be the orientation distribution.

We adopt the following notation for indexing the crossing tree. These notations were partially introduced in the previous chapter, but we recall them here for convenience. Let $\emptyset$ be the root of the tree, representing a single level-0 crossing. The first generation of children are labeled by i, $1 \leqslant i \leqslant Z_\emptyset$, where $Z_\emptyset$ is the number of children of $\emptyset$. The second generation are then labeled ij, $1 \leqslant j \leqslant Z_i$, and so on. More generally, a node is an element of $U = \cup_{n \geqslant 0} \mathbb{N}^{*n}$ and a branch is a couple $(\mathbf{u}, \mathbf{u}j)$ where $\mathbf{u} \in U$ and $j \in \mathbb{N}$. The length of a node $\mathbf{i} = i_1 \dots i_n$ is $|\mathbf{i}| = n$. If $|\mathbf{i}| > n$, $\mathbf{i}|_n$ is the curtailment of $\mathbf{i}$ after n terms. Conventionally $|\emptyset| = 0$ and $\mathbf{i}|_0 = \emptyset$. A tree Υ is a set of nodes, that is a subset of U, such that

- $\emptyset \in \Upsilon$
- If a node $\mathbf{i}$ belongs to the tree then every ancestor node $\mathbf{i}|_k$, $k \leqslant |\mathbf{i}|$, belongs to the tree
- If $\mathbf{u} \in \Upsilon$, then $\mathbf{u}j \in \Upsilon$ for $j = 1, \dots, Z_\mathbf{u}$ and $\mathbf{u}j \notin \Upsilon$ for $j > Z_\mathbf{u}$, where $Z_\mathbf{u}$ is the number of children of $\mathbf{u}$.

Let $F(s) = \sum_{j=0}^{\infty} p(j) s^j$ denote the probability generating function defined for complex s such that $|s| \leqslant 1$. Let Υ_n^{GW} be the n-th generation of the tree, that is the set of nodes of length n and Z_n^{GW} its cardinal and $\mu = \sum_x xp(x)$ be the mean family size. Then $\mu^{-n} Z_n^{GW}$ is a non-negative martingale and thus converges almost surely to some limit $W_\emptyset$ [10]. Consider $\Lambda(s) = \mathbb{E}(e^{-sW_\emptyset})$, the Laplace transform of $W_\emptyset$ defined for complex s such that $\mathrm{Re}(s) \geqslant 0$. Λ satisfies the Poincaré functional equation [109]

$$\Lambda(\mu s) = F(\Lambda(s)). \tag{3.1}$$

Let $\Upsilon_\mathbf{i} = \{\mathbf{j} \in \Upsilon \mid |\mathbf{j}| \geqslant |\mathbf{i}| \text{ and } \mathbf{j}|_{|\mathbf{i}|} = \mathbf{i}\}$. The boundary of the tree is given by $\partial\Upsilon = \{\mathbf{i} \in \mathbb{N}^\mathbb{N} \mid \forall n \geqslant 0,\ \mathbf{i}|_n \in \Upsilon\}$. Let $\psi(\mathbf{i})$ be the position of node $\mathbf{i}$ within generation $|\mathbf{i}|$

$$\begin{aligned} \psi : \Upsilon &\to \mathbb{N}^+ \\ \mathbf{i} &\mapsto \psi(\mathbf{i}) = \#\{\mathbf{j} \preccurlyeq \mathbf{i}\} \end{aligned}$$

where the operator $\preccurlyeq$ is defined by

$$\begin{aligned} \mathbf{j} \preccurlyeq \mathbf{i} \iff\ & |\mathbf{j}| = |\mathbf{i}| = n \text{ and the exists } k,\ 1 \leqslant k \leqslant n \text{ such that} \\ & \mathbf{j}|_{k-1} = \mathbf{i}|_{k-1} \text{ and } j_k < i_k \text{ with non-strict inequality if } k = n. \end{aligned}$$

When convenient we will write $Z_\mathbf{i}$ and $T_\mathbf{i}$ for $Z_{\psi(\mathbf{i})}^{|\mathbf{i}|}$ and $T_{\psi(\mathbf{i})}^{|\mathbf{i}|}$, and so on.

THEOREM 6. *For any regular distribution p on $2\mathbb{N}$ there exists a unique continuous EBP process X defined on $[0, T_1^0]$ such that*

- *Orientations follow a distribution $p_{c|z}$ which can depend on the number of children of the parent crossing.*
- *The level n crossing duration is distributed like $\mu^{-n}W_\emptyset$ where the Laplace transform of $W_\emptyset$ satisfies (3.1).*

Then we call X the Canonical EBP (CEBP) process with offspring distribution p and orientation distribution $p_{c|z}$. Let $\mu = \sum_x xp(x)$ be the mean family size. Let $H = \log 2/\log \mu$, then for all $a \in \{\mu^n, n \in \mathbb{Z}\}$,

$$X(t) \stackrel{fdd}{=} a^{-H}X(at) \text{ for finite dimensional distributions.} \tag{3.2}$$

Proof. This result can be found as Theorem 1 in [69], for a particular orientation distribution. The proof given there contains a mistake, which is corrected here.

X is obtained as the limit as $n \to +\infty$, of a sequence of random walks X^{-n} with steps of size 2^{-n} and duration μ^{-n}. If we add weights $1/\mu$ to each branch of the crossing tree of X^{-n}, then the product of the weights along a line of descent is μ^{-n}, which is the duration of any crossing of X^{-n}. This construction will be generalized in the next section, where we allow random weights on the crossing tree and therefore random crossing durations.

Put $X^0(0) = 0$ and $X^0(1) = 1$, so that the coarsest scale is $n = 0$. It is sufficient to construct a crossing from 0 to 1. We explain in Section 3.1.2, Corollary 2, how to extend the support of X to any compact interval of the form $[0, T]$. Given X^{-n} we construct $X^{-(n+1)}$ by replacing the k-th step of X^{-n} by a sequence of Z_k^{-n} steps of size $2^{-(n+1)}$, where the Z_k^{-n} are i.i.d. and $\mathbb{P}(Z_k^{-n} = x) = p(x)$. The new steps can be split into $(Z_k^{-n} - 2)/2$ excursions (an up-down or down-up pair) followed by a direct crossing (two ups or two downs, depending on the parent crossing). We allow a general orientation distribution of the $(Z_k^{-n} - 2)/2$ excursions, which can depend on Z_k^{-n}, that we denote $p_{c|z}$. For diffusion processes, we expect excursion to be up-down or down-up with equal probability since it is shown in [70] that diffusions look like continuous local martingales at small scales.

We extend X^{-n} by linear interpolation, from $\mu^{-n}\mathbb{Z}^+ \to 2^{-n}\mathbb{Z}$ to $\mathbb{R}^+ \to \mathbb{R}$. Interpolated X^{-n} has continuous sample paths and we show that it converges uniformly on any finite interval, from which the continuity of the limit process will follow. Let

$$T^{-n} = \inf\{t \mid X^{-n}(t) = 1\}$$

and set $X^{-n}(t) = 1$ for all $t > T^{-n}$. For $0 \leqslant m \leqslant n$, let $T_0^{-m,-n} = 0$ and

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

If $X^{-n}(T_k^{-m,-n}) = 1$ then set $T_{k+1}^{-m,-n} = \infty$. The duration of the k-th level $-m$ crossing of X^{-n} is $D_k^{-m,-n} = T_k^{-m,-n} - T_{k-1}^{-m,-n}$. Since the Z_k^{-n} are i.i.d. $\{\mu^n D_k^{-m,-n}\}_{n=m}^{\infty}$ is a Galton-Watson branching process. Since p is regular, there exists continuous strictly positive random variables D_k^{-m} such that $\mathbb{E}D_k^{-m} = 1$ and [73]

$$D_k^{-m,-n} \to D_k^{-m} \text{ a.s. as } n \to \infty.$$

Let $T_k^{-m} = \sum_{j=1}^k D_j^{-m} = \lim_{n\to\infty} T_k^{-m,-n}$. Clearly, $T_1^0 = \inf\{t \mid X(t) = 1\}$.

Take any $\epsilon > 0$, $\delta > 0$ and $T > 0$. To establish the a.s. convergence of the processes X^{-n} uniformly on compact intervals, we show that we can find a u so that for all $r, s \geqslant u$ and $t \in [0, T]$

$$|X^{-r}(t) - X^{-s}(t)| \leqslant \delta \text{ with probability } 1 - \epsilon. \tag{3.3}$$

Given $t \in [0, T]$, let $k = k(n, t)$ be such that

$$T_{k-1}^{-n} \leqslant t < T_k^{-n}.$$

For any $r, s \geqslant n$, the triangle inequality yields

$$|X^{-r}(t) - X^{-s}(t)| \leqslant |X^{-r}(t) - X^{-r}(T_k^{-n,-r})| + |X^{-s}(T_k^{-n,-s}) - X^{-s}(t)|, \tag{3.4}$$

since $X^{-r}(T_k^{-n,-r}) = X^{-s}(T_k^{-n,-s}) = X^{-n}(k\mu^{-n})$.

For any $n \leq u$ let $j = j(n, u)$ be the smallest j such that $T_j^{-n,-u} > T$. As $u \to +\infty$, $j(n, u) \to j(n)$ a.s., so for any n w can choose ϵ_0 such that

$$\mathbb{P}(\min_{i \leqslant j(n)} D_i^{-n} \geqslant \epsilon_0) \geq 1 - \epsilon,$$

and u such that for all $q \geq u$,

$$\mathbb{P}(\max_{i \leqslant j(n)} |T_i^{-n,-q} - T_i^{-n}| < \epsilon_0) \geq 1 - \epsilon,$$

which yields

$$\mathbb{P}(\max_{i \leqslant j(n)} |T_i^{-n,-q} - T_i^{-n}| < \min_{i \leqslant j(n)} D_i^{-n}) \geq 1 - \epsilon.$$

Thus, given n, we can find u such that for all $q \geq u$, with probability at least $1 - \epsilon$,

$$T_{k-2}^{-n,-q} < t < T_{k+1}^{-n,-q}.$$

Now, since $X^{-q}(T_{k-2}^{-n,-q}) = X^{-n}((k-2)\mu^{-n})$, $X^{-q}(T_{k+1}^{-n,-q}) = X^{-n}((k+1)\mu^{-n})$, and in three steps X^{-n} can move at most distance $3 \cdot 2^{-n}$, we have

$$|X^{-q}(t) - X^{-q}(T_k^{-n,-q})| \leqslant 3 \cdot 2^{-n}.$$

Then choosing n large enough such that $6 \cdot 2^{-n} \leqslant \delta$, we see that (3.3) follows from (3.4). Sending δ and ϵ to 0 shows that X^{-n} converges to some continuous limit process X uniformly on all closed intervals $[0, T]$, with probability 1.

The process X constructed here is unique given the joint distribution of the Z_k^{-n}, the crossing orientations and the crossing times. Solutions of (3.1) such that $\mathbb{E}D_1^0 = 1$ are unique provided $\mu > 1$ and $\mathbb{E}Z_1^0 \log Z_1^0 < \infty$, which follows from the regularity condition on p (Theorem 8.2 in [54]). The process X for which the crossing time distribution satisfy the Poincaré functional equation (3.1) is called the CEBP process. Note that in the construction of the CEBP, we allow a general expression for the joint distribution of the Z_k^{-n} and the orientations, but we specify the crossing durations.

By construction, simultaneously scaling the state space by 2^k and time space by μ^k does not change the distribution of $X(t)$, so for all $t \in [0, T_1^0]$,

$$X(t) \overset{fdd}{=} 2^{-k} X(\mu^k t) = (\mu^k)^{-\log 2/\log \mu} X(\mu^k t) \tag{3.5}$$

where $\overset{fdd}{=}$ is for finite dimensional distributions. □

Recall that Z_n^{GW} is the size of generation n. If p is regular then from Kesten & Stigum [73] we have that $W_\emptyset > 0$ a.s. and $\mathbb{E}W_\emptyset = 1$, where $W_\emptyset$ is the almost sure limit of martingale $\mu^{-n}Z_n^{GW}$. Therefore, for the CEBP process, we associate $W_\emptyset$ with the root node of the tree as it gives the duration of the crossing $\emptyset$, hence the notation $W_\emptyset$. Similarly, one can attach to node $\mathbf{i}$ in generation n the random variable $W_\mathbf{i}$, which is the normed-limit of the tree rooted at $\mathbf{i}$. This is illustrated Figure 3.3. The duration of crossing $\mathbf{i}$ is then $\mu^{-n}W_\mathbf{i}$.

We assume regularity of p throughout the remainder of the chapter. $H = \log 2/\log \mu$ is known as the Hurst index and the process X is said to be discrete scale-invariant. From [52], Brownian motion is an example of a CEBP process, corresponding to the $2\,\text{geometric}_1(1/2)$ offspring distribution: the Z_k^n are i.i.d. for all n and k, with $\mathbb{P}(Z_k^n = 2i) = 2^{-i}$, $i = 1, 2, \ldots$ (see Theorem 9). We present in Figure 3.4 three realisations of CEBP processes with geometric offspring distribution and parameters $\theta = 0.3, 0.5, 0.7$.

3.1.2 Extension of the support of the CEBP

In this section we explain how to scale up the support of the CEBP process to any compact interval of $\mathbb{R}^+$. To do so, we consider a sequence $\{X^{(n)}\}$ of CEBP processes and investigate two types of convergence to a limit process X: convergence for finite dimensional distributions and weak convergence (also called convergence in distribution).

Let $X^{(0)}$ be a CEBP process constructed as before, with duration $T_1^0 = \inf\{t \mid X^{(0)}(t) = \pm 1\}$, using the notation of the proof of Theorem 6. We define $X^{(0)}(t) = \pm 1$ for all $t \geqslant T_1^0$. Let $\mathcal{C}_T$ be the space of continuous functions with compact support $[0, T]$. We define a sequence $\{X^{(n)}\}$ of CEBP processes by

$$X^{(n)}(t) = 2^n X^{(0)}(\mu^{-n}t).$$

Proposition 2. *$X^{(n)}$ converges in finite dimensional distribution to a limit process X in $\mathcal{C}_T$. For all $k \geqslant 1$, and $0 \leqslant t_1 \ldots \leqslant t_k \leqslant T$,*

$$(X^{(n)}(t_1) \ldots X^{(n)}(t_k)) \overset{fdd}{\to} (X(t_1) \ldots X(t_k)).$$

Proof. Let $k \geqslant 1$ and $0 \leqslant t_1 \leqslant \ldots \leqslant t_k \leqslant T$. Let $n \leqslant m$. $X^{(n)}$ and $X^{(m)}$ have respective durations T_1^n and T_1^m. By construction, their finite dimensional distributions agree on $[0, T_1^n]$. It remains to show that

$$\mathbb{P}(\{T_1^n < T\} \text{ for infinitely many } n) = \mathbb{P}(\{T_1^n < T\} \text{ i.o.}) = 0.$$

If the total duration of $X^{(0)}$ is $W_\emptyset$, then $T_1^n = \mu^n W_\emptyset$. By the Borel-Cantelli lemma, it is sufficient to show that

$$\sum_{n=1}^{\infty} \mathbb{P}(\{T_1^n < T\}) = \sum_{n=1}^{\infty} \mathbb{P}(\{W_\emptyset < \mu^{-n}T\}) < \infty.$$

From Biggins and Bingham (Theorem 3 in [24]), the left tail of $W_\emptyset$ decays exponentially when the mean family size is larger than 2. More precisely, as $n \to \infty$,

$$-\log \mathbb{P}(W_\emptyset < \mu^{-n}T) = \alpha\gamma^n L(\mu^{-n}T) + o(\mu^{-n}T)$$

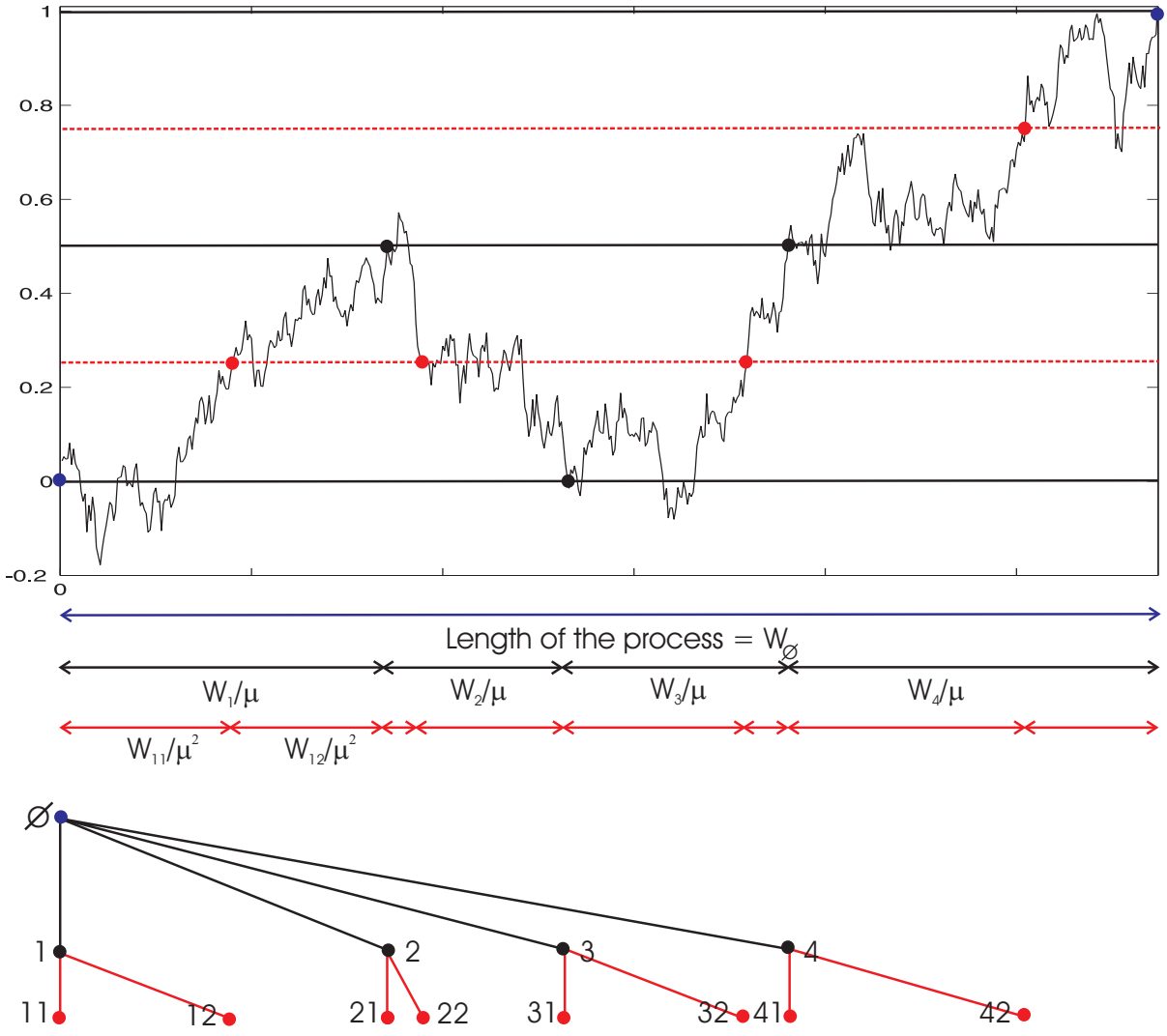


Figure 3.3: The top signal is a CEBP from 0 to 1. We consider its crossings of size 0, -1 and -2, together with the crossing durations. The crossing duration of size 0 is represented in blue, and its duration is $W_\emptyset$. This crossing corresponds to the root node $\emptyset$ of the crossing tree. Therefore, we attach the random variable $W_\emptyset$ to $\emptyset$. Similarly, crossings of size -1, in black, have durations $\mu^{-1}W_i$, $i = 1, \ldots, 4$ where W_i is the normed-limit of the tree rooted at i. For this reason, we attact the random variable $\mu^{-1}W_i$ to node i, and so on.

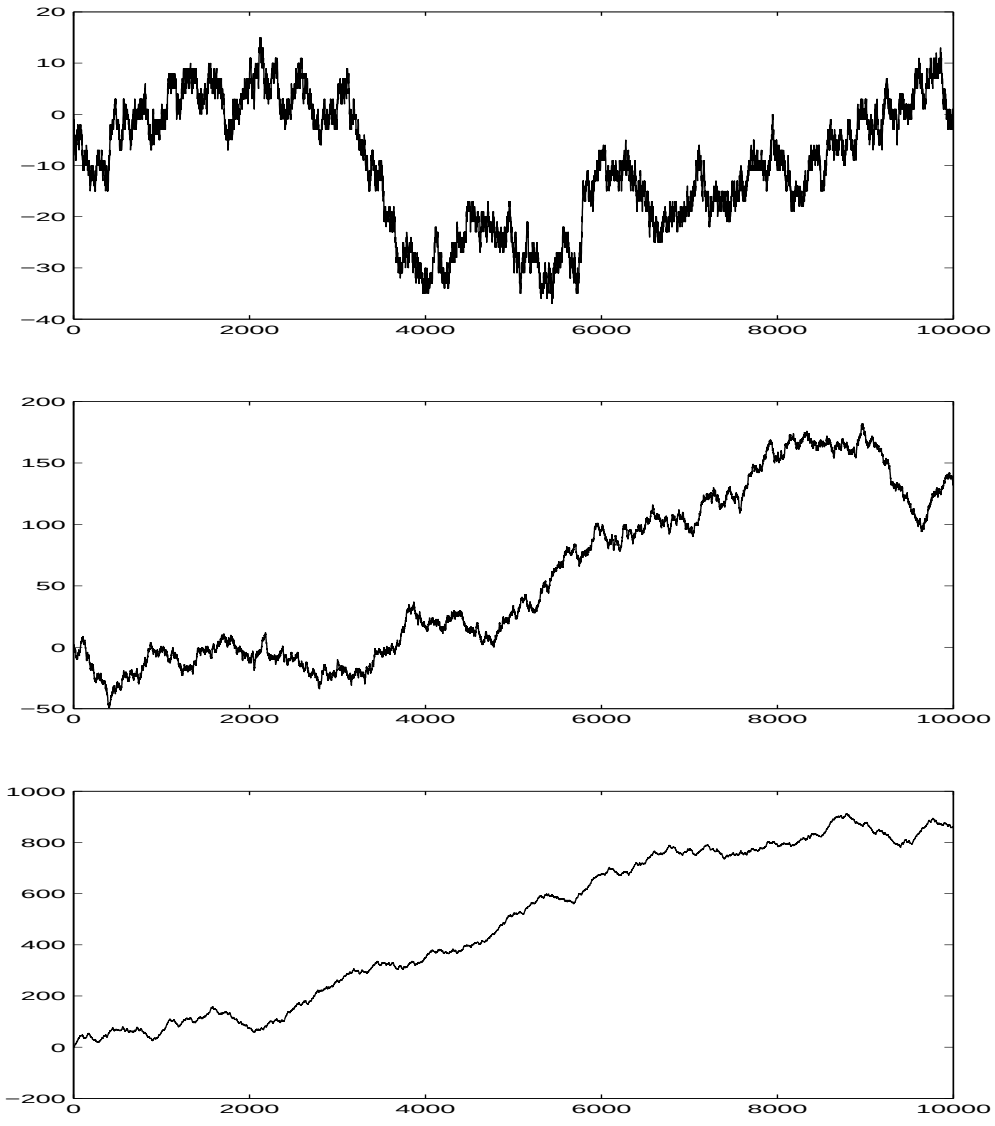


Figure 3.4: From top to bottom. Realisations of CEBP processes with geometric offspring distribution with parameter 0.3, 0.5 and 0.7.

where $\alpha > 0$, $\gamma > 1$ and L is a real strictly positive function and multiplicatively periodic. The result follows for CEBP processes since their mean family size is larger than 2. □

Convergence in finite dimensional distribution is not enough to establish the weak convergence of the sequence. The sequence $X^{(n)}$ converges in distribution to a limit process if and only if it converges in finite dimensional distribution and if $X^{(n)}$ is *tight* (Corollary 11.6.2 in [115]). We give a criteria for a sequence $\{W_n\}$ of random elements of $\mathcal{C}_T$ to be tight:

THEOREM 7. (Theorem 11.6.3 in [115]) *A sequence $\{W_n \mid n \geqslant 1\}$ of random elements of $\mathcal{C}_T$ is tight if and only if, for every $\epsilon > 0$, there exists a constant c such that*

$$\mathbb{P}(|W_n(0) > c|) < \epsilon \text{ for all } n \geqslant 1$$

and, for every $\epsilon > 0$ and $\eta > 0$, there exists $\delta > 0$ and n_0 such that

$$\mathbb{P}(v(W_n, \delta) \geqslant \epsilon) \leqslant \eta \text{ for all } n \geqslant n_0$$

where

$$v(W_n, \delta) := \sup \{|W_n(t_1) - W_n(t_2)| \mid 0 \leqslant t_1 \leqslant t_2 \leqslant T \text{ and } |t_1 - t_2| \leqslant \delta\}.$$

LEMMA 2. *$X^{(n)}$ is tight.*

Proof. The first condition is trivial since $X^{(n)}(0) = 0$ for all n. Let $\epsilon > 0$. Choose n^* large enough such that $2^{-n^*} \leqslant \epsilon$. Let $a > 0$ and $(t_1, t_2) \in [0, T]$. If

$$|t_1 - t_2| \leqslant a \leqslant \min_{\mathbf{i} \in \Upsilon_{n^*+1}} \mu^{-(n^*+1)} W_{\mathbf{i}}$$

then $[t_1, t_2]$ intersects at most two intervals of generation n^* and $v(X^{(n^*)}, a) \leqslant 2^{-n^*} \leqslant \epsilon$. It follows that the event $\{v(X^{(n^*)}, a) \geqslant \epsilon\}$ is false whenever $\{\min_{\mathbf{i} \in \Upsilon_{n^*+1}} \mu^{-(n^*+1)} W_{\mathbf{i}} \leqslant a\}$ is false and

$$\mathbb{P}(v(X^{(n^*)}, a) \geqslant \epsilon) \leqslant \mathbb{P}(\min_{\mathbf{i} \in \Upsilon_{n^*+1}} \mu^{-(n^*+1)} W_{\mathbf{i}} \leqslant a).$$

Let $\eta > 0$. Take a small enough such that

$$\mathbb{P}(\min_{\mathbf{i} \in \Upsilon_{n^*+1}} \mu^{-(n^*+1)} W_{\mathbf{i}} \leqslant a) \leqslant \eta$$

and set $\delta = a$. □

COROLLARY 2. *$X^{(n)}$ converges in distribution to a limit process X in $\mathcal{C}_T$.*

3.2 From CEBP to MEBP

In this section we construct Multifractal Embedded Branching Processes as time changed CEBP processes.

The crossing tree of a CEBP process X gives the number of subcrossings of each crossing. If we add a weight of $1/\mu$ to each branch of the tree, then truncating the tree at level n, the product of the weights down any line of descent is μ^{-n}, which is the duration of any single crossing by X^{-n}. We generalize this by allowing the weights to be random, then defining the duration of a crossing to be the product of the random weights down the line of descent of the crossing.

Let X be a CEBP (level 0 crossing) and consider its crossing tree. We assign weight $\rho_j(\mathbf{i})$ to the branch $(\mathbf{i}, \mathbf{i}j)$. $\rho_1(\mathbf{i}), \ldots, \rho_{Z_\mathbf{i}}(\mathbf{i})$ may be dependent and depend on $Z_\mathbf{i}$, but must be independent of other nodes. The weight attributed to node $\mathbf{i}$ is then

$$\rho_\mathbf{i} = \prod_{k=1}^{|\mathbf{i}|} \rho_{i_k}(\mathbf{i}|_{k-1}).$$

That is, $\rho_\mathbf{i}$ is the product of all weights on the line of descent from the root down to node $\mathbf{i}$. We use the weights to define a measure, ν, on the boundary of the crossing tree. The measure ν on $\partial\Upsilon$ is then mapped to a measure ζ on $\mathbb{R}$, with which we define a chronometer $\mathcal{M}$ (a non-decreasing process) by $\mathcal{M}(t) = \zeta([0,t))$. The MEBP is then given by $Y = X \circ \mathcal{M}^{-1}$, where X is the CEBP. The crossing trees of X and Y have the same structure, but have different crossing durations. In Figure 3.10 we plot a realisation of an MEBP and its associated CEBP.

3.2.1 The measure ν

For the remainder of the chapter we will make the following assumptions about the weights $\rho_1, \ldots, \rho_{Z_\emptyset}$.

Assumption 1.

$$\rho_i > 0, \quad \mathbb{E}\sum_{i=1}^{Z_\emptyset} \rho_i(\emptyset) = 1, \quad 0 > \mathbb{E}\sum_{i=1}^{Z_\emptyset} \rho_i(\emptyset)\log\rho_i(\emptyset) > -\infty$$
$$\textit{and } \mathbb{E}\sum_{i=1}^{Z_\emptyset} \rho_i(\emptyset)\log\sum_{i=1}^{Z_\emptyset} \rho_i(\emptyset) < \infty.$$

Define $\mathcal{W}^n = \sum_{\mathbf{i}\in\Upsilon_n} \rho_\mathbf{i}$. If we take constant weights equal to $1/\mu$, then $\mathcal{W}^n = \mu^{-n} Z_n^{GW}$, which is the martingale defined in Section 3.1. In fact, $\mathcal{W}^n$ is always a non-negative martingale, and thus converges almost surely to a random variable $\mathcal{W}$. For regular p and under Assumption 1, Biggins & Kyprianou [25] show that the limit $\mathcal{W}$ is strictly positive and $\mathbb{E}\mathcal{W} = 1$. The distribution of $\mathcal{W}$ is unknown in general.

Write $\mathcal{W}_\emptyset$ for $\mathcal{W}$ and attach it to the root of the tree. In a similar fashion, we can attach to node $\mathbf{i}$ a random variable $\mathcal{W}_\mathbf{i}$ independent within a generation, defined and distributed like $\mathcal{W}$, using $\Upsilon_\mathbf{i}$

$$\mathcal{W}_\mathbf{i} = \lim_{n\to\infty} \sum_{\mathbf{j}\in\Upsilon_\mathbf{i}, |\mathbf{j}|=n} \rho_\mathbf{j}/\rho_\mathbf{i}.$$

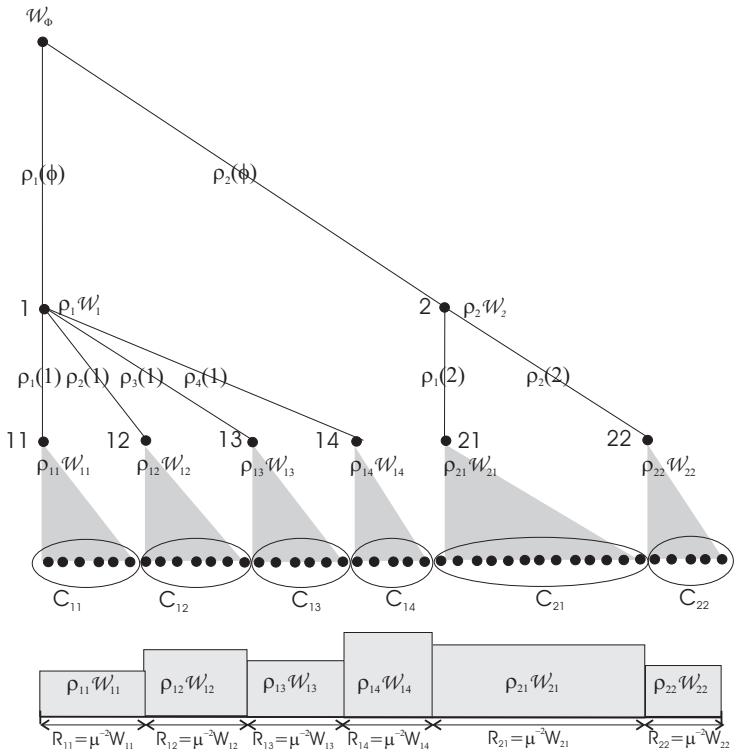


Figure 3.5: Measure of cylinder sets and their mapping onto intervals of the real line.

$\mathcal{W}_{\mathbf{i}}$ can be seen as the mass of $\Upsilon_{\mathbf{i}}$ and $\rho_{\mathbf{i}}$ as a weighting of $\Upsilon_{\mathbf{i}}$.

From the construction of $\mathcal{W}_{\mathbf{i}}$ it follows that almost surely, for all nodes $\mathbf{i}$,

$$\mathcal{W}_{\mathbf{i}} = \sum_{j=1}^{Z_{\mathbf{i}}} \rho_j(\mathbf{i})\mathcal{W}_{\mathbf{i}j}. \tag{3.6}$$

For $\mathbf{i} \in \Upsilon_n^{GW}$, let $C_{\mathbf{i}}$ be the cylinder set defined by $C_{\mathbf{i}} = \{\mathbf{j} \in \partial\Upsilon \mid \mathbf{j}|_n = \mathbf{i}\}$. In other words, $C_{\mathbf{i}}$ contains all the nodes on the boundary of the tree which have $\mathbf{i}$ as an ancestor. We define $\nu(C_{\mathbf{i}}) = \rho_{\mathbf{i}}\mathcal{W}_{\mathbf{i}}$. From (3.6), ν is additive, as one would expect for a measure. The term $\rho_{\mathbf{i}}$ is a weight corresponding to the 'past' of $\mathbf{i}$, from the root node to $\mathbf{i}$ and the term $\mathcal{W}_{\mathbf{i}}$ can be seen as the total weight of what is hanging below $\mathbf{i}$, its 'future', from $\mathbf{i}$ to all its descendants on the boundary of the tree. By Carathéodory's Extension Theorem, we can uniquely extend ν to the σ-algebra generated by these cylinder sets. The construction of ν is illustrated in Figure 3.5 where we have represented the cylinder sets corresponding to the second generation of the tree.

3.2.2 Existence and continuity of the MEBP

The measure ζ is a mapping of ν from $\partial\Upsilon$ to $[0, W_\emptyset] \subset \mathbb{R}$. By analogy with m-ary cascades, we call ζ a Galton-Watson cascade measure on $[0, W_\emptyset]$. Let T_k^{-n} denote the k-th level-$(-n)$ passage time of the CEBP process X and put $R_{\mathbf{i}} = [T^{-|\mathbf{i}|}_{\psi(\mathbf{i})-1},\ T^{-|\mathbf{i}|}_{\psi(\mathbf{i})})$.

Note that n goes up for the tree (generations $1, 2, 3 \ldots$) but down for the crossings (hitting times $T_k^{-1}, T_k^{-2}, T_k^{-3}, \ldots$). Then $|R_{\mathbf{i}}| = \mu^{-|\mathbf{i}|} W_{\mathbf{i}}$ and we define

$$\zeta(R_{\mathbf{i}}) := \nu(C_{\mathbf{i}}) = \rho_{\mathbf{i}} \mathcal{W}_{\mathbf{i}}.$$

Intervals $R_{\mathbf{i}}$ are illustrated in Figure 3.5. We can now give a ζ-measure of intervals of the form $[0, T_k^{-n})$ and use them to define a time change $\mathcal{M}$ of the original process:

$$\mathcal{M}(T_k^{-n}) := \zeta([0\ T_k^{-n})) = \sum_{\mathbf{i} \in \Upsilon_n^{GW} : \psi(\mathbf{i}) \leqslant k} \rho_{\mathbf{i}} \mathcal{W}_{\mathbf{i}}.$$

The MEBP process Y is defined through its hitting times $\mathcal{T}_k^{-n}$ that we define by $Y(\mathcal{T}_k^{-n}) = Y(\mathcal{M}(T_k^{-n})) = X(T_k^{-n})$, which yields

$$Y = X \circ \mathcal{M}^{-1}.$$

Thus the duration of crossing $\mathbf{i}$ for the time-changed process Y is $\rho_{\mathbf{i}} \mathcal{W}_{\mathbf{i}}$.

We need to make this definition more precise. So far, the MEBP process is defined only at its hitting times. We want to define $Y(t)$ for all $t \in [0, \mathcal{W}_\emptyset]$. In other words, we need to make sure that the chronometer $\mathcal{M}(t) = \zeta([0,\ t))$ is defined for all t. It is sufficient to construct Y for $t \in [0, \mathcal{W}_\emptyset]$. We explain later how to scale up the support of Y to $[0, \infty)$ without changing its statistical properties. For any point $t \in [0, W_\emptyset]$ we can find $\mathbf{i} \in \partial\Upsilon$ such that $t \in R_{\mathbf{i}|_n}$ for all $n \geq 0$. If $|R_{\mathbf{i}|_n}| \to 0$ for all $\mathbf{i} \in \partial\Upsilon$, then $\{t\} = \bigcap_n R_{\mathbf{i}}$ and we can define

$$\mathcal{M}(t) = \zeta([0,\ t)) = \lim_{n \to \infty} \zeta([0,\ T^{-n}_{\psi(\mathbf{i}|_n)})).$$

Thus to establish the existence of the MEBP it is sufficient to show that as $n \to \infty$, $\max_{\mathbf{i} \in \Upsilon_n^{GW}} |R_{\mathbf{i}}| \to 0$. To this end we need the following lemma due to Pakes [99], which gives a limit result on the right tail of the distribution of an extreme order statistic defined on a Galton-Watson tree.

LEMMA 3. (Theorem 4.1 in [99]) *Suppose that we are given a Galton-Watson tree with average family size $\mu > 1$ and regular offspring distribution, and equip the j-th node of generation n with a random variable $\{X_{j,n}\}$. Suppose independence and identical distribution F from one node to another. Denote by $X_{(1),n}$ the largest random variable at generation n and suppose that F is in the domain of attraction of an extremal density function $G(x) = e^{-\gamma(x)}$. That is, there exist constants $a(n)$ and $b(n) > 0$ such that*

$$\lim_{n \to +\infty} n[1 - F(a(n) + x b(n))] = \gamma(x). \tag{3.7}$$

Let $A_n = a(\mu^n)$, $B_n = b(\mu^n)$ and $W_\emptyset$ the limit martingale defined in Section 3.1. Then

$$\lim_{n \to +\infty} P(X_{(1),n} \leqslant A_n + x B_n) = \mathbb{E}[e^{-W_\emptyset \gamma(x)}]. \tag{3.8}$$

We use the lemma to prove the following result.

ASSUMPTION 2. *There exists $\epsilon > 0$ such that $\mathbb{E}Z_\emptyset^{1+\epsilon} < \infty$.*

PROPOSITION 3. *Let p a regular distribution and suppose Assumption 2 holds. Then,*

$$\lim_{n\to+\infty} \max_{\mathbf{i}\in\Upsilon_n^{GW}} \mu^{-n} W_{\mathbf{i}} = 0 \quad a.s.$$

Proof. Since for each $\mathbf{i} \in \partial\Upsilon$, $\mu^{-n} W_{\mathbf{i}|_n}$ is a nested decreasing sequence, we have that $\max_{\mathbf{i}\in\Upsilon_n^{GW}} \mu^{-n} W_{\mathbf{i}|_n}$ is a decreasing sequence. Thus, it is sufficient to prove that

$$\lim_{n\to+\infty} \mathbb{P}(\max_{\mathbf{i}\in\Upsilon_n^{GW}} \mu^{-n} W_{\mathbf{i}} = 0) = 1.$$

Let $\epsilon > 0$. It is a famous result that $\mathbb{E}Z_\emptyset^{1+\epsilon}$ and $\mathbb{E}W_\emptyset^{1+\epsilon}$ converge or diverge together [27]. Therefore, under Assumption 2, $\mathbb{E}W_\emptyset^{1+\epsilon} < \infty$. By Chebyshev's inequality, for $x > 0$,

$$\mathbb{P}(W_{\mathbf{i}} > x) \leqslant x^{-(1+\epsilon)} \mathbb{E}(W_{\mathbf{i}}^{1+\epsilon}) = \alpha x^{-(1+\epsilon)}$$

for some $\alpha < \infty$. That is, the tail of $W_{\mathbf{i}}$ decays as $x^{-(1+\epsilon)}$. Thus we can find i.i.d. random variables $V_{\mathbf{i}}$ such that

$$\mathbb{P}(W_{\mathbf{i}} > x) \leqslant \mathbb{P}(V_{\mathbf{i}} > x) \sim c x^{-(1+\epsilon)} \text{ as } x \to +\infty$$

for some constant c. $V_{\mathbf{i}}$ is in the domain of attraction of the extremal law $e^{-x^{-(1+\epsilon)}}$. That is

$$\lim_{n\to+\infty} n\mathbb{P}(V_{\mathbf{i}} > a(n) + x b(n)) = x^{-(1+\epsilon)}$$

with $a(n) = 0$ and $b(n) = (nc)^{1/(1+\epsilon)}$. Lemma 3 lets us conclude that:

$$\lim_{n\to+\infty} \mathbb{P}(\max_{\mathbf{i}\in\Upsilon_n^{GW}} V_{\mathbf{i}} > x c^{1/(1+\epsilon)} \mu^{n/(1+\epsilon)}) = 1 - \mathbb{E}(e^{-W_\emptyset x^{-(1+\epsilon)}}).$$

Take $m \leqslant \epsilon n$, then

$$\begin{aligned} \lim_{n\to+\infty} \mathbb{P}(\max_{\mathbf{i}\in\Upsilon_n^{GW}} \mu^{-n} W_{\mathbf{i}} > x) &\leqslant \lim_{n\to+\infty} \mathbb{P}(\max_{\mathbf{i}\in\Upsilon_n^{GW}} \mu^{-(\frac{n}{1+\epsilon}+\frac{m}{1+\epsilon})} V_{\mathbf{i}} > x) \\ &\leqslant 1 - \mathbb{E}[\exp(-W_\emptyset c x^{-(1+\epsilon)} \mu^{-m})] \\ &\longrightarrow 0 \text{ as } m \to \infty. \end{aligned}$$

This concludes the proof. □

Since MEBP processes are now defined as $Y(t) = X \circ \mathcal{M}^{-1}(t)$, the continuity of Y follows from the continuity of X if $\mathcal{M}^{-1}$ is continuous, or equivalently if $\mathcal{M}$ has no flat spots. In other words, we have to make sure no interval of positive length has zero ζ-measure. That is, for every $\mathbf{i} \in \partial\Upsilon$ we must have $\rho_{\mathbf{i}|_n} \mathcal{W}_{\mathbf{i}|_n} > 0$ for all $n \geq 0$. However, for regular p and under Assumption 1, this follows immediately from Biggins & Kyprianou, who proved the existence of the almost sure limit $\mathcal{W}_{\mathbf{i}|_n}$ [25] with strictly positive distribution [26].

We summarize conditions for existence and continuity of Y in the theorem below.

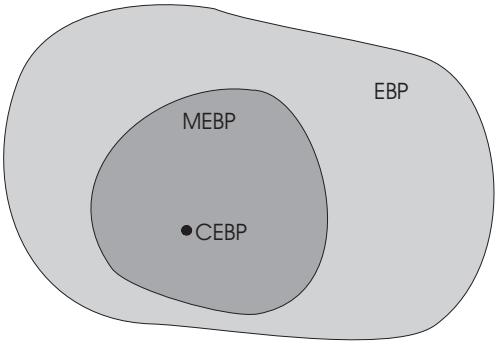


Figure 3.6: Relationships between the spaces of EBP, CEBP and MEBP processes.

THEOREM 8. *Let p be a regular offspring distribution, $p_{c|z}$ the orientation distribution and X the corresponding CEBP. If Assumptions 1 and 2 hold then we can construct a chronometer $\mathcal{M}$ such that $\mathcal{M}^{-1}$ is a.s. continuous, and so define*

$$Y(t) = X \circ \mathcal{M}^{-1}(t) \text{ for all } t \in [0, \mathcal{W}_\emptyset].$$

Moreover, Y is continuous a.s.

Figure 3.6 illustrates the relationships between the spaces of EBP, CEBP and MEBP processes.

It is not straightforward to extend the support of an MEBP to any compact interval of the real line and we cannot use a similar procedure as for CEBP processes. First, it is not known whether MEBP possess a discrete scale invariance like CEBP, which was used to establish the convergence in finite dimensional distribution. We also need results on the left tail of the distribution of $\mathcal{W}$. There are results in the case of finite Ω in [53]. Extension of the support of MEBP remains an open problem.

3.2.3 Continuity of $\mathcal{M}$

If $\mathcal{M}$ is continuous then it is easier to relate the properties of X and Y. It turns out that continuity of $\mathcal{M}$ is a much stronger restriction than continuity of $\mathcal{M}^{-1}$. In the remainder we will sometimes use the phrase 'a.s. on trees' for $\mathbb{P}$-a.s.

$\mathcal{M}$ is continuous if ζ has no atoms. In other words, if for all $\mathbf{i} \in \partial\Upsilon$, $\lim_{n\to+\infty} \rho_{\mathbf{i}|_n} \mathcal{W}_{\mathbf{i}|_n} = 0$. We have

PROPOSITION 4. *If p is regular and Assumption 1 holds and $\mathcal{W}_\emptyset$ admits moments of all (positive) orders, then*

$$\lim_{n\to+\infty} \max_{\mathbf{i}\in\partial\Upsilon} \rho_{\mathbf{i}|_n} \mathcal{W}_{\mathbf{i}|_n} = 0 \quad \text{a.s. on trees}$$

Proof. The proof uses an embedded branching random walk (BRW) [23]. A BRW is described by a triple (Q, M, χ), whose elements respectively describe the reproduction, movement and importance of each individual. We label individuals using elements of $U = \cup_{n\geqslant 0}\mathbb{N}^n$ (as for the crossing tree). With each individual $\mathbf{i}$ we associate an independent version of (Q, M, χ), denoted $(Q_\mathbf{i}, M_\mathbf{i}, \chi_\mathbf{i})$. Q is a point process

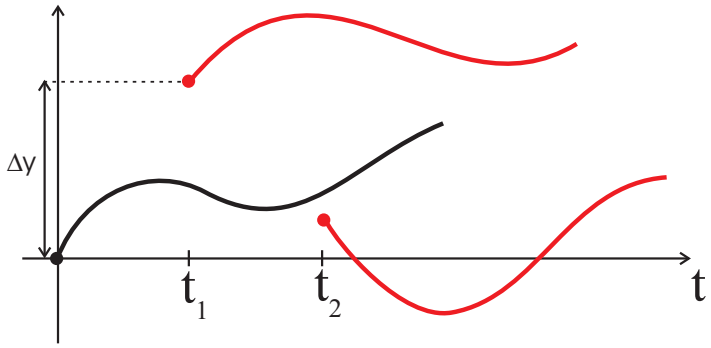


Figure 3.7: **Branching Random Walk** (Q, M, χ)**.** We start with one initial particle (black) which gives birth to two offspring (orange) at times t_1 and t_2 during its lifetime. The displacement of the first child from its parent's birth position is Δy. The pair $(t_1, \Delta y) \in \mathbb{R}\times\mathbb{R}^+$ gives the first point of the point process Q. The movement of black and orange particles are described by the stochastic process M. Finally, many choices are allowed for $\chi(t)$, which gives the importance of each individual. One can consider for example a unit weight if the particle is alive at time t and no weight if the particule is not yet born at time t, or dead.

on $\mathbb{R} \times \mathbb{R}^+$, its first coordinate gives a child's displacement from its parent's birth position and the second gives the parent's age at that child's birth. M is a stochastic process which describes an individual's movement during its lifetime. $\chi : \mathbb{R}^+ \to \mathbb{R}$ is the so-called characteristic function. The triple (Q, M, χ) is illustrated in Figure 3.7.

We construct a BRW as follows. The point process $Q_{\mathbf{i}}$ generates $Z_{\mathbf{i}}$ points at time 1, with displacements $\log \rho_{\mathbf{i}}(1), \ldots, \log \rho_{\mathbf{i}}(Z_{\mathbf{i}})$. Thus if $\sigma_{\mathbf{i}}$ is the birth position of individual $\mathbf{i}$, then $\sigma_{\mathbf{i}} = \log \rho_{\mathbf{i}}$. The birth time of individual $\mathbf{i}$ is just $|\mathbf{i}|$. Individual movement consists of a single jump at birth of size $\log \mathcal{V}_{\mathbf{i}}$, where the $\mathcal{V}_{\mathbf{i}}$ are i.i.d. and distributed like $\mathcal{W}_\emptyset$. We only consider the $\mathcal{V}_{\mathbf{i}}$ since the $\mathcal{W}_{\mathbf{i}}$ are i.i.d. within one generation. That is, $M_{\mathbf{i}}(t) = \log \mathcal{V}_{\mathbf{i}}$ for all $t \geq 0$. The characteristic χ is used to count generations. If we put $\chi_{\mathbf{i}}(t) = \mathbb{I}_{[0,1)}(t)$, where $\mathbb{I}$ is the indicator function, then the population at time t is given by $S_t = \sum_{\mathbf{i}\in U} \chi_{\mathbf{i}}(t - \sigma_{\mathbf{i}})$.

The right-most particle at time t is defined as

$$B_t = \max_{\mathbf{i}\in U} \chi_{\mathbf{i}}(t - \sigma_{\mathbf{i}})(\sigma_{\mathbf{i}} + M_{\mathbf{i}}(t - \sigma_{\mathbf{i}})) = \max_{\mathbf{i}\in \Upsilon^{GW}_{\lfloor t \rfloor}} \log(\rho_{\mathbf{i}}\mathcal{V}_{\mathbf{i}}).$$

Biggins [23] gives conditions for the existence of a γ such that $B_t/t \to \gamma$ a.s. If $\gamma < 0$ then clearly $\max_{\mathbf{i}\in\Upsilon^{GW}_n} \rho_{\mathbf{i}}\mathcal{V}_{\mathbf{i}} \to 0$ a.s. as $n \to \infty$, and thus $\max_{\mathbf{i}\in\Upsilon^{GW}_n} \rho_{\mathbf{i}}\mathcal{W}_{\mathbf{i}} \to 0$ a.s., which is our desired result.

Biggins' results on the right-most particle require that the triple (Q, M, χ) is 'well-regulated'. In our case if we put $m(\theta) = \mathbb{E}\sum_{j=1}^{Z_\emptyset} \rho_j(\emptyset)^{-\theta}$ and $\alpha(\theta) = \log m(\theta)$, then the triple is well-regulated if for all $\theta \leq 0$,

$$\mathbb{E}\Big(\sup_t \big\{ e^{-\alpha(\theta)t} e^{-\theta M(t)} \chi(t) \big\} \Big) < \infty.$$

Let $c(\theta) = \sup_{t\in[0,1)} e^{-\alpha(\theta)t}$ then our triple is well regulated if $c(\theta)\mathbb{E}\mathcal{W}_\emptyset^{-\theta} < \infty$ for all $\theta \leq 0$. That is, if $\mathcal{W}_\emptyset$ admits moments of all orders.

To finish the proof we just need $\gamma < 0$. Let $\alpha^*(a) = \inf_{\theta<0}\{a\theta + \alpha(\theta)\}$ then from [23]

$$\gamma = \inf\{a \mid \alpha^*(a) < 0\}.$$

γ will be strictly negative if for some $\epsilon < 0$,

$$\inf_{\theta<0}\{\epsilon\theta + \alpha(\theta)\} < 0. \tag{3.9}$$

As ϵ is arbitrary, this will hold if for some $\theta_0 < 0$, $m(\theta_0) < 1$. Now m is convex over $(-\infty, 0]$, since $m''(\theta) = \mathbb{E}\sum_{j=1}^{Z_\emptyset} \rho_j(\emptyset)^{-\theta}(\log \rho_j(\emptyset))^2 > 0$. Also, from Assumption 1, $m(-1) = 1$ and $m'(-1) = \mathbb{E}\sum_{j=1}^{Z_\emptyset} \rho_j(\emptyset)\log \rho_j(\emptyset) < 0$. Thus there exists $\theta_0 \in (-1, 0)$ such that $m(\theta_0) < 1$, and the result follows. □

We finish this section with some verifiable conditions that imply $\mathcal{W}_\emptyset$ has moments of all orders.

Assumption 3. *Assume that*

$$\rho_i \in (0, 1],\ \mathbb{P}(\sum_{i=1}^{Z_\emptyset} \rho_i(\emptyset) = 1) < 1 \text{ and}$$
$$\text{for all } p > 1,\ \mathbb{E}\Big[\Big(\sum_{i=1}^{Z_\emptyset} \rho_i(\emptyset)\Big)^p\Big] < \infty.$$

Liu [78] shows that if p is regular and Assumptions 1 and 3 hold then $\mathbb{E}\mathcal{W}_\emptyset^p < \infty$ for all $p \geq 0$.

Corollary 3. *If Assumptions 1, 2 and 3 hold, then $Y(t)$ and $\mathcal{M}(t)$ are continuous.*

3.3 On-line simulation

There are many ways we could make a multifractal time-change of a CEBP. However, by defining a time-change via the crossing tree, we obtain a fast on-line algorithm to simulate the process. In this section, we present a new algorithm which is a generalization and a simplification of the one proposed by Jones in [69] for simulating CEBP. It is a generalisation since it incorporates the time change and a simplification since the length of the code is reduced by generating the crossing types and the weights using a vector notation. We first explain the key representation of CEBP behind this algorithm using vector notation, and then explain how it can be modified to incorporate the time change.

3.3.1 Markov representation

3.3.1.1 Markov representation of CEBP

When constructing CEBP processes we started with a level-0 crossing then generated consistent sequences of crossings at finer and finer scales. When simulating CEBP we specify a minimum level, without loss of generality level-0, and then generate crossings at level 0 and above.

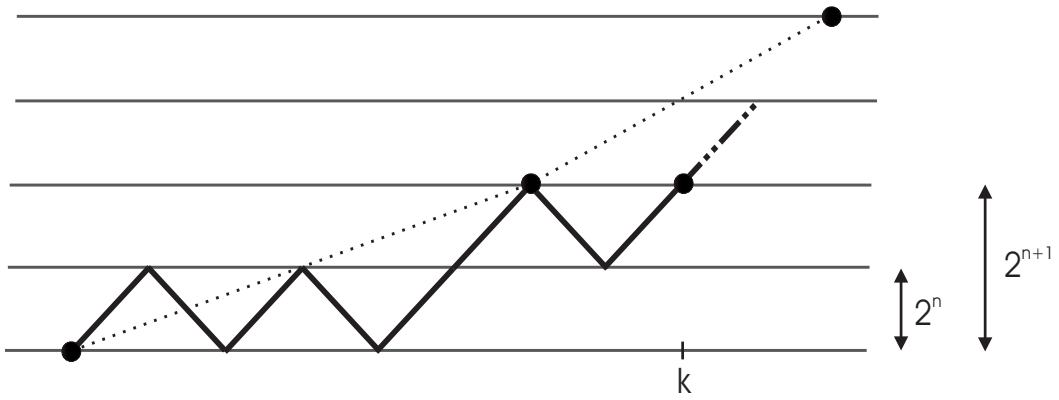


Figure 3.8: Markov representation of CEBP processes X^n. We have joined the points T_k^n at level n and $n+1$. The process X is not represented. Consider the k-th crossing of level n. Then $\kappa(n,n,k) = k$, $S^n_{\kappa(n,n,k)} = 3$, $Z^{n+1}_{\kappa(n,n+1,k)} = 4$. Also $\kappa(n,n+1,k) = 2$ and $\kappa(n,n+2,k) = 1$ so that $n_{\max} = n+2$ here.

Let X^m be the random walk on $2^m\mathbb{Z}$, defined by $X^m(k) = X(T_k^m)$ for $k = 1, 2, \ldots$ Let C_k^n be the k-th level n crossing (see Definition 4), which is divided into Z_k^n subcrossings. For $0 \leq m \leq n$, let $\kappa(m,n,k)$ be the number such that $X^m(k)$ belongs to crossing $C^n_{\kappa(m,n,k)}$. Clearly $\kappa(n,n,k) = k$. Also, denote by S_k^n the position of crossing C_k^n within its parent crossing $C^{n+1}_{\kappa(n,n+1,k)}$. Thus $1 \leqslant S_k^n \leqslant Z^{n+1}_{\kappa(n,n+1,k)}$. Finally, note that a crossing can be one of 6 different types, depending on its direction (up or down) and where it starts. Consider a level n crossing whose parent crossing starts at $k2^{n+1}$: the 6 types are 0^+, 0^-, 1^+, 1^-, -1^+ and -1^- where type i^+ denotes a crossing from $k2^{n+1} + i2^n$ to $k2^{n+1} + (i+1)2^n$ and type i^- is a crossing from $k2^{n+1} + i2^n$ to $k2^{n+1} + (i-1)2^n$. Let α_k^n be a vector of size Z_k^n whose components are the types of the Z_k^n subcrossings of C_k^n. Denote the distribution of α_k^n by p_c, which can depend on the number of subcrossings, the type of the parent crossing and the position of the current crossing with its parent crossing. We make explicit this dependence by writing $p_{c|Z^{n+1}_{\kappa(n,n+1,k)},i^{+/-}}$ or $p_{c|Z^{n+1}_{\kappa(n,n+1,k)},S^n_k,i^{+/-}}$ when appropriate.

Define

$$\mathcal{X}^n(k) = (\kappa(0,n,k), S^n_{\kappa(0,n,k)}, Z^{n+1}_{\kappa(0,n+1,k)}, \alpha^{n+1}_{\kappa(0,n+1,k)})$$

then $\mathcal{X}^n(k)$ describes the level n crossing of X^0 at time k.

DEFINITION 7. *$\mathcal{X}(k) = \{\mathcal{X}^n(k)\}_{n\geqslant 0}$ is called the crossing state of X^0 at time k.*

It is shown in [69] that $\mathcal{X}$ is a Markov chain, so knowing $\mathcal{X}(k)$ is enough to generate $\mathcal{X}(k+1)$. The proof consists of showing procedure `Increment` described in the next section only requires $\mathcal{X}(k)$ to generate $\mathcal{X}(k+1)$, which is tedious but straightforward. Initially this may not seem useful, since $\mathcal{X}(k)$ is infinite, however it turns out that we do not need all of $\mathcal{X}(k)$, just the truncation $\overline{\mathcal{X}}(k) := \{\mathcal{X}^0(k), \ldots, \mathcal{X}^{n_{\max}}(k)\}$, where $n_{\max}$ is such that $\kappa(0, n_{\max}, k) = \kappa(0, n_{\max}, 1) = 1$. This is illustrated in Figure 3.8 where we consider the k-th level n crossing of X^n.

Given a sequence $\{\mathcal{X}(k)\}_{k=1}^{N}$, from $\{\alpha_k^1\}_{k=1}^{\kappa(0,1,N)}$ we can generate a process X^0 with steps of size 1, which are the level-0 crossing points of our CEBP process. Noting that $\overline{\mathcal{X}}(N) = \{\mathcal{X}^0(N), \ldots, \mathcal{X}^{n_{\max}}(N)\}$ contains all the information we need about $\{X^0(k)\}_{k=1}^{N}$ in order to simulate $X^0(N+1)$, we see that this is a parsimonious encoding of the past of X^0, since $n_{\max}$ grows like $\log N$: a finite Galton-Watson tree with $n_{\max}$ levels has on average $\mu^{n_{\max}}$ leaves. Then $N \approx \mu^{n_{\max}}$ and $n_{\max} = O(\log N)$ follows.

For the CEBP process, the crossing durations of X^0 are i.i.d. with distribution W. In general we can not simulate W exactly, though we can do so approximately, and in practice we usually just set the level-0 crossing durations to be constant. We could evaluate the error induced by this assumption by simulating a few crossings below the scale 0. We did not do this study here.

3.3.1.2 Markov representation of MEBP

To simulate an MEBP we simulate a CEBP as well as weights for the crossing tree, which we use to generate the crossing durations. We extend $\mathcal{X}^n(k)$ to include weights $P^{n+1}_{\kappa(0,n+1,k)}$ where P_k^n is a vector of size Z_k^n containing the random weights attached to the Z_k^n branches of C_k^n. To enable us to truncate the crossing state of X^0 at time k, for each n we norm $\{P^{n+1}_{\kappa(0,n+1,k)}\}_{k\geq 1}$ by $P^{n+1}_{\kappa(0,n+1,1)}(S^n_{\kappa(0,n,1)}) = P_1^{n+1}(S_1^n)$, which is the weight of the first crossing generated at level n. Let $\tilde{P}^{n+1}$ be the first weight of generation n: $\tilde{P}^{n+1} = P_1^{n+1}(S_1^n)$. Since $\kappa(0,n,k) = 1$ for all $n \geq n_{\max}$ we have $P^{n+1}_{\kappa(0,n+1,k)}(S^n_{\kappa(0,n,k)})/\tilde{P}^{n+1} = 1$ for all $n \geq n_{\max}$ and we can truncate without loss of information. It follows that with the new definition of $\mathcal{X}^n(k)$,

$$\mathcal{X}^n(k) = (\kappa(0,n,k), S^n_{\kappa(0,n,k)}, Z^{n+1}_{\kappa(0,n+1,k)}, \alpha^{n+1}_{\kappa(0,n+1,k)}, P^{n+1}_{\kappa(0,n+1,k)}, \tilde{P}^{n+1}).$$

$\overline{\mathcal{X}}(k) := \{\mathcal{X}^0(k), \ldots, \mathcal{X}^{n_{\max}}(k)\}$ is still Markov. We denote the distribution of the vector of weights by q_z, which can be discrete or continuous. The duration of crossing C_k^0 is given by

$$\mathcal{W}_k^0 \prod_{n=0}^{n_{\max}-1} P^{n+1}_{\kappa(0,n+1,k)}(S^n_{\kappa(0,n,k)})/\tilde{P}^{n+1}$$

where the $\mathcal{W}_k^0$ are i.i.d. with the distribution of $\mathcal{W}$. As for the CEBP we can not in general simulate the $\mathcal{W}_k^0$ exactly, and in practice set them to be constant. We sometimes write $S^{m,n}(k)$, $Z^{m,n+1}(k)$, $\alpha^{m,n}(k)$ and $P^{m,n+1}(k)$ for $S^n_{\kappa(m,n,k)}$, $Z^{n+1}_{\kappa(m,n+1,k)}$, $\alpha^n_{\kappa(m,n,k)}$ and $P^{n+1}_{\kappa(m,n+1,k)}$.

3.3.2 On-line algorithm

Let X be an MEBP and let $X_p(k)$ and $X_t(k)$ be respectively the position and time of sample k. In this section, we give the simulation algorithm in detail. It can be divided into three main procedures: `Increment`, `Expand` and `Simulation`.

The procedure `Simulation` loops the procedures `Expand` and `Increment` to produce a new sample on demand. It returns sample position $X_p(k+1)$, sample time $X_t(k+1)$ and crossing state $\overline{\mathcal{X}}(k+1)$ given $X_p(k)$, $X_t(k)$ and $\overline{\mathcal{X}}(k)$. `Increment`

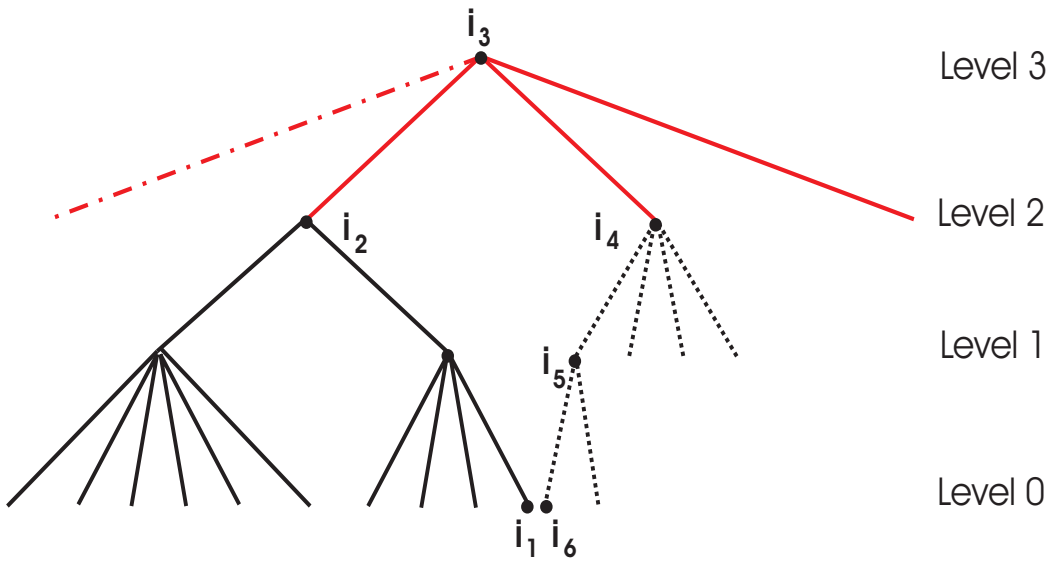


Figure 3.9: Description of the procedures `Increment`, `Expand` and `Simulation`. Suppose we have generated the tree in solid black lines only, so that $n_{\max} = 2$. When we reach node $\mathbf{i}_1$, we are at the end of level 0,1 and 2 crossings. To generate the next sample, one needs to increase $n_{\max}$ by at least 1, which is the role of the procedure `Expand`. `Expand` generates the level 3 of the crossing tree (in orange on the figure), generates the number of children of $\mathbf{i}_3$, the position of $\mathbf{i}_2$ within this family and the vector of types and weights. If $\mathbf{i}_2$ is at the end of level 3 crossing (which is not the case here), then `Expand` increases $n_{\max}$ again by 1. Here, $n_{\max} = 3$. Once we are sure we have all the levels needed to generate the next sample, procedure `Increment` goes down the tree and generates the family size, type of crossings and weights at nodes $\mathbf{i}_4$ and $\mathbf{i}_5$, at levels 2 and 1, which contains all the information to generate the next sample, corresponding to node $\mathbf{i}_6$. The procedure `Simulation` loops `Expand` and `Increment` to generate as many samples as required.

generates $\overline{\mathcal{X}}(k+1)$ from $\overline{\mathcal{X}}(k)$ assuming that you have enough levels in the crossing tree to do so. `Expand` makes sure that this is the case, by increasing $n_{\max}$ when we reach the end of level $n_{\max}$ crossing. This is illustrated in Figure 3.9. A Matlab implementation is available from the web page of Jones [68].

Procedure Increment. We start this section by giving the procedure `Increment`.

`Procedure Increment` $\mathcal{X}^n(k)$
(Assume that $X^0(k)$ is at the end of level n crossing)
(This is always the case for $n = 0$)
$\kappa(0, n, k+1) = \kappa(0, n, k) + 1$
† `If` $S^n_{\kappa(0,n,k)} = Z^{n+1}_{\kappa(0,n+1,k)}$ `Then`
`Increment` $\mathcal{X}^{n+1}(k)$
$S^n_{\kappa(0,n,k+1)} = 1$
Generate $Z^{n+1}_{\kappa(0,n+1,k+1)}$ using distribution p
Generate $P^{n+1}_{\kappa(0,n+1,k+1)}$ using distribution $q_{Z^{n+1}_{\kappa(0,n+1,k+1)}}$
`If` $\alpha^{n+2}_{\kappa(0,n+2,k+1)}(S^{n+1}_{\kappa(0,n+1,k+1)}) = i^+$ `Then`
†† Generate $\alpha^{n+1}_{\kappa(0,n+1,k+1)}$ using distribution $p_{c|Z^{n+1}_{\kappa(0,n+1,k+1)},i^+}$
`Else`
Generate $\alpha^{n+1}_{\kappa(0,n+1,k+1)}$ using distribution $p_{c|Z^{n+1}_{\kappa(0,n+1,k+1)},i^-}$
`End If`
††† `Else`
$\mathcal{X}^q(k+1) = \mathcal{X}^q(k)$ `for` $q = n+1, \ldots, n_{max}$
$S^n_{\kappa(0,n,k+1)} = S^n_{\kappa(0,n,k)} + 1$
`End If`
`End Procedure`

To generate $\overline{\mathcal{X}}(k+1)$ given $\overline{\mathcal{X}}(k)$, apply procedure `Increment` to $\mathcal{X}^0(k)$. Then, `Increment` is applied to all level q crossings $\mathcal{X}^n(k)$ such that $X^0(k)$ is at the end of level $q+1$ crossing; that is, $S^q_{\kappa(m,q,k)} = Z^{q+1}_{\kappa(m,q+1,k)}$, for all $0 \leqslant q < n$. $\mathcal{X}^n(k)$ remains unchanged for all n larger than this. We give a few comments on the procedure `Increment`:

At †, $X^0(k)$ is at the end of level $n+1$ crossing. Information at level $n+1$ needs to be updated, so `Increment` is applied to $\mathcal{X}^{n+1}(k)$.

On the other hand, at †††, $X^0(k)$ is not at the end of level $n+1$ crossing. Information at all coarser scales remain unchanged; only the position of crossing C^n_k within its parent crossing is updated.

At ††, the type $\alpha^{n+2}_{\kappa(0,n+2,k+1)}$ of the parent crossing is already known to be $i^{+/-}$ so the vector of crossing types $\alpha^{n+1}_{\kappa(0,n+1,k+1)}$ at the finer scale must be generated accordingly, hence the notation.

Procedure Expand. We give a procedure `Expand` to increase $n_{\max}$ when needed, that is when $X^0(k)$ is at the end of level n crossing for $n = 0 \ldots n_{\max}$. We distinguish two cases:

- Simulation of a CEBP process: The position of the crossing within its parent crossing is chosen in a random manner, as described in Figure 3.9.
- Simulation of an MEBP process: Assume $X^0(0)$ belongs to the first level m crossing, for $m = 1, 2, \ldots$. In Figure 3.9, this is equivalent to generate the level 3 of the crossing tree and to place node $\mathbf{i}_2$ at the beginning of the level 3 generation.

The need to separate these two situations will become clear later.

Suppose we want to simulate a CEBP process. Two preliminary results are needed. The first one concerns the local structure of CEBP processes and provides a way to determine the distribution of the direction of the parent crossing when we go from scale n to $n+1$. The second one gives an explicit expression of the offspring distribution to consider when we move upward in the crossing tree.

LEMMA 4. *Let Υ be the crossing tree of a CEBP process with regular offspring distribution and Υ_n^{GW} its n-th generation, for some finite integer n. Pick a node uniformly on Υ_n^{GW}. Then, in the limit $n \to \infty$, the associated crossing has direction $+$(up) or $-$(down) with equal probability.*

Proof. Let X be a CEBP process and consider its crossing tree together with the type $+$ or $-$ of each branch. Denote by Z_n^+ and Z_n^- the number of types $+/-$ at generation n. Then the size of generation n is $Z_n = Z_n^+ + Z_n^-$. Consider M, the matrix of first moments, whose entries are

$$m_{++} = \mathbb{E}(Z_1^+ | Z_0 \text{ type } +) = 2 + \mathbb{E}((Z_n - 2)/2) = 1 + \mu/2$$

$$m_{-+} = \mathbb{E}(Z_1^+ | Z_0 \text{ type } -) = \mathbb{E}((Z_n - 2)/2) = -1 + \mu/2$$

since a $+$ crossing is decomposed in $(Z_n - 2)/2$ excursions followed by two $+$ subcrossings and similarly for a $-$ crossing, which is decomposed into $(Z_n - 2)/2$ excursions followed by two $-$ subcrossings. It is easy to see that $m_{++} = m_{--}$ and $m_{+-} = m_{-+}$ so that

$$M = \begin{pmatrix} 1 + \frac{\mu}{2} & \frac{\mu}{2} - 1 \\ \frac{\mu}{2} - 1 & 1 + \frac{\mu}{2} \end{pmatrix}.$$

It is a famous result in the theory of multitype Galton-Watson processes that under certain conditions (see below) the left unit eigenvector ϖ corresponding to the maximum eigenvalue λ of the mean matrix M gives the relative fixed proportions of types in the limit $n \to \infty$. By unit vector we mean its components sum up to 1. Suppose $\lambda > 1$ and that the process is positive regular, that is, powers of M have all positive components. This is true here since $\mu > 2$ from which the positivity of M follows. Finally, we assume that the random variable representing the number of particles of type j produced by a particle of type i, for $(i, j) \in \{+, -\}$, has finite mean, which holds for CEBP processes with regular offspring distribution. Then, almost surely on non-extinction, we have (Theorem 2 in [76])

$$\lim_{n \to \infty} \frac{(Z_n^+, Z_n^-)}{Z_n} = \nu.$$

Crossing trees of CEBP processes possess at least two children per node and therefore extinction cannot occur. A straightforward calculation yields $\lambda = \mu > 2$ and $\varpi = (1/2, 1/2)$ which concludes the proof. □

COROLLARY 4. *Let X be a CEBP process with regular offspring distribution and $T > 0$. Observe X at some t uniformly distributed over $[0, T)$, and let k be such that $X(t) \in C_k^n$. Then the type of C_k^n is $+$ or $-$ with equal probability, in the limit $n \to \infty$.*

Motivated by this result, if the current crossing is such that $S^n_{\kappa(0,n,k)} = Z^{n+1}_{\kappa(0,n+1,k)}$ for all $n = 0 \dots n_{\max}$, then we choose the type $\alpha^{n_{\max}+2}_{\kappa(0,n_{\max}+2,k)}$ of the parent crossing $C^{n_{\max}+2}_{\kappa(0,n_{\max}+2,k)}$ to be $+$ or $-$ with equal probability. We do not know the error induced by this approximation.

We now derive a result concerning the offspring distribution and the position of a current crossing within its parent crossing when we move from level n to level $n+1$: we cannot sample directly from p when we move from a fine scale to a coarser one. To illustrate this fact, consider individuals which have a random number of offspring according to some distribution $\mathcal{L}$ and pick randomly a certain number of them to estimate $\mathcal{L}$. Since it is more likely to pick individuals coming from a large family, the distribution we are estimating is not $\mathcal{L}$ but some transformation of it. A similar issue arises when generating CEBP processes. When choosing a time t randomly, we are more likely to hit a large crossing and therefore the distribution giving the number of subcrossing of the parent crossing is not p but something else, which we now make explicit. To do so, let $X_{1,1}, X_{1,2}, \dots, X_{1,N(1)}, X_{2,1}, \dots, X_{k,N(k)}, \dots$ be a sequence of i.i.d. non-negative random variables arranged into families, with $P(X_{i,j} \leqslant x) = F(x)$ and $P(N(i) = n) = p(n)$. Then partition $[0, \infty)$ into adjacent intervals with length $X_{i,j}$ ordered as above, choose a point x uniformly in $[0, T)$ for some $T > 0$ and consider the sizes of the interval and of the family that contain x.

LEMMA 5. *Let $\mathcal{P}$ be the partition above, let X^* and N^* be the interval length and family size of a uniformly chosen point on $[0, T)$ for some $T > 0$, and let J^* be the position of the chosen interval within its family. If $\mu = \sum_x x\, p(x)$ and $m = \mathbb{E}X_{i,j}$ are finite then with probability 1, for $1 \leq l \leq n$, as $T \to \infty$,*

$$P(N^* = n,\, J^* = l,\, x \leq X^* < x + dx \,|\, \mathcal{P}) \to \frac{np(n)}{\mu} \frac{1}{n} \frac{x}{m} dF(x).$$

Note that X^* will have atoms at the same points as $X_{i,j}$.

Proof. We give a correction to the proof in [69], which is flawed.

Let $T_k = \sum_{i=1}^{k} \sum_{j=1}^{n(i)} X_{i,j}$ and let $\mathcal{P}_k$ be the partition of $[0, T_k)$ given by $X_{1,1}, \dots, X_{k,N(k)}$. Given $\mathcal{P}_k$, choose x uniformly on $[0, T_k)$ and let X_k^* and N_k^* be the interval length and family size of x.

Let $S_k(n) = \#\{i : 1 \leq i \leq k,\, N(i) = n\}$ then sending $k \to \infty$

$$\begin{aligned}
&P(N_k^* = n,\, J^* = l,\, X_k^* \leq x \,|\, \mathcal{P}) \\
&= \sum_{i=1}^{k} \sum_{j=1}^{N(i)} P(N_k^* = n,\, J^* = l,\, X_k^* \leq x \,|\, N_k^* = N(i),\, X_k^* = X_{i,j},\, \mathcal{P}) \frac{X_{i,j}}{T_k} \\
&= \sum_{i=1}^{k} \sum_{j=1}^{n} I_{\{n\}}(N(i)) I_{\{l\}}(j) I_{[0,x]}(X_{i,j}) \frac{X_{i,j}}{T_k} \\
&= \frac{k}{T_k} \frac{S_k(n)}{k} \frac{1}{S_k(n)} \sum_{i=1}^{k} I_{\{n\}}(N(i)) X_{i,l} I_{[0,x]}(X_{i,l})
\end{aligned}$$

$$\begin{aligned}
&\to \frac{1}{\mu m} p(n) \mathbb{E} X_{i,j} I_{[0,x]}(X_{i,j}) \text{ with probability 1} \\
&= \frac{np(n)}{\mu} \frac{1}{n} \int_0^x \frac{y}{m} dF(y).
\end{aligned}$$

Differentiating w.r.t. x gives the result.

By integrating/summing out the other terms, one can easily show that the marginal distributions of N^* and X^* are given by

$$\begin{aligned}
P(N^* = n) &= \frac{np(n)}{\mu}; \\
P(x \leq X^* < x + dx) &= \frac{x}{m} dF(x).
\end{aligned}$$

Similarly, the conditional distribution of J^* given N^* is given by

$$P(J^* = l \,|\, N^* = n) = \frac{1}{n} \text{ for } 1 \leq l \leq n.$$

It follows that N^* and J^* are independent of X^*. □

The consequence of this result for the simulation of CEBP processes is stated in the following corollary.

COROLLARY 5. *Let X be a CEBP process. Observe X at some t uniformly distributed over $[0, T)$, and let k be such that $X(t) \in C_k^n$. Then $Z_{\kappa(n,n+1,k)}^{n+1}$ has distribution $\overline{p}(x) = xp(x)/\mu$ and S_k^n is uniformly distributed over $\{1, \ldots, Z_{\kappa(n,n+1,k)}^{n+1}\}$ in the limit $T \to \infty$.*

Finally, the sampling distributions of S_k^n and $Z_{\kappa(n,n+1,k)}^{n+1}$ are independent of the length of C_k^n.

This corollary does not hold for the generation of MEBP processes. Indeed, a fundamental assumption in the derivation of Lemma 5 is the independence of the variables $X_{i,j}$, which correspond to crossing durations. For CEBP processes, crossing durations are independent of each other, but this is no longer true for MEBP processes, where correlation exists. It would be interesting to investigate how this

correlation modifies Lemma 5. We do not pursue this here. What we do is, instead of generating randomly the position of the crossing within its parent crossing for MEBP processes, we assume that the position of the first level m crossing is 1 for all $m \geqslant 0$. Consequently, $Z^{n+1}_{\kappa(n,n+1,k)}$ is still sampled according to p and $S^n_k = 1$. When simulating MEBP processes, replace $\overline{p}$ by p and $S^{0,n_{\max}+1}(k) \sim U\{1,\ldots,Z^{0,n_{\max}+2}(k)\}$ by $S^n_k = 1$ in the following procedure:

```
Procedure Expand X̄(k)
    While S^{0,n_max}(k) = Z^{0,n_max+1}(k) Do
        κ(0, n_max + 1, k) = 1
        Generate Z^{0,n_max+2}(k) using distribution p̄
        Generate S^{0,n_max+1}(k) ~ U{1,...,Z^{0,n_max+2}(k)}
        Generate P^{0,n_max+2}(k) using distribution q_{Z^{n_max+2}_{κ(0,n_max+2,k)}}
        If α^{n_max+1}_{κ(0,n_max+1,k)}(S^{n_max}_{κ(0,n_max,k)}) = 1^+ Then
            Generate α^{n_max+2}_{κ(0,n_max+2,k)} using distribution p_{c|Z^{n_max+2}_{κ(0,n_max+2,k)},S^{n_max+1}_{κ(0,n_max+1,k)},1^+}
        Else
            Generate α^{n_max+2}_{κ(0,n_max+2,k)} using distribution p_{c|Z^{n_max+2}_{κ(0,n_max+2,k)},S^{n_max+1}_{κ(0,n_max+1,k)},-1^-}
        End If
        P_1(n_max + 2) = P^{n_max+2}_{κ(0,n_max+2,k)}(S^{n_max+1}_{κ(0,n_max+1,k)})
        n_max = n_max + 1
    End While
End Procedure
```

In `Expand`, the vector of types $\alpha^{n_{\max}+2}_{\kappa(0,n_{\max}+2,k)}$ is sampled from $p_{c|z}$ after conditioning on the size of the parent crossing and the type $1^+/-1^-$ of the current crossing $S^{n_{\max}}_{\kappa(0,n_{\max},k)}$. The type of the parent crossing is then $+$ or $-$ with equal probability, according to Corollary 4. Once the type of the parent crossing is decided, then one can use a rejection algorithm that generates vectors $\alpha^{n_{\max}+2}_{\kappa(0,n_{\max}+2,k)}$ until the type of crossing $S^{n_{\max}}_{\kappa(0,n_{\max},k)}$ is $1^+/-1^-$.

Procedures Simulate and Initialise. Finally we can give the simulation algorithm, which loops the procedures `Expand` and `Increment` to generate a new sample.

```
Procedure Simulate
    (Given X̄(k), X^0_p(k) and X^0_t(k) (respectively the crossing state, the position
    and time of the process at sample k) returns X̄(k+1), X^0_p(k+1) and X^0_t(k+1))
    If α^1_{κ(0,1,k+1)}(S^0_{κ(0,0,k+1)}) = i^+ Then
        X^0_p(k + 1) = X^0_p(k) + 1
    Else
        X^0_p(k + 1) = X^0_p(k) − 1
    End If
    For j = 0 : n_max
        P_t(j + 1) = P^{j+1}_{κ(0,j+1,k+1)}(S^j_{κ(0,j,k+1)})
    End
†   X_t(k + 1) = X_t(k) + ∏ P_t/P_1
    Expand X̄(k)
```

```
        Increment 𝒳^0(k)
End Procedure
```

At †, we renormalise each weight in P_t by the first weight of each generation, contained in P_1.

To initialise the algorithm, the procedure `Initialise` can be used:

```
Procedure Initialise 𝒳̄^m(k)
        n_max = 0
        κ(0,0,1) = 1
        Generate Z^{0,1}(1) using distribution p̄
        Generate S^{0,0}(1) ~ U{1,...,Z^{0,1}(1)}
        Generate P^{0,1}(1) using distribution q_{Z^{0,1}(1)}
        Generate α^{0,1}(1) using distribution p_{c|Z^{0,1}(1)}
        P_1 = P^1_{κ(0,1,1)}(S^0_{κ(0,0,1)})
End Procedure
```

Similarly, when generating MEBP processes, $\overline{p}$ is replaced by p and $S^{0,0}(1) \sim U\{1,\dots,Z^{0,1}(1)\}$ by $S^{0,0}(1) = 1$.

3.3.3 Efficiency

Generation n of the crossing tree has on average μ^n leaves. Therefore, X^0 starts a new crossing after μ^n steps and it follows that $n_{\max}(k) = O(\log k)$. Thus, generating N steps requires $O(\log N)$ storage.

The procedure `Expand` is finite and does not depend on $n_{\max}(k)$ and we go through `Increment` up to $n_{\max}(k)$ times. Therefore, the number of operations computed by `Simulation` is of order $N \log N$ since $\sum_{k=1}^{n} \log k = O(N \log N)$.

The type of signal obtained with this algorithm is illustrated in Figure 3.10 where we have represented an MEBP process with its CEBP. p, q_z and $p_{c|z}$ are given in the Figure caption.

3.4 Fractional Brownian Motion

In this section, we propose to construct empirically an MEBP process which imitates a fractional Brownian motion with Hölder exponent H. The idea of this section is mainly to provide a description of fBm in terms of its crossing tree. In the particular case $H = 0.5$, we can fully describe a Brownian motion in terms of its crossing tree. Consider the k-th level n crossing and let D_k^n and Z_k^n denote the crossing length and the number of subcrossings respectively. V_j^n is the crossing excursion (let $V_j^n = 0$ if the j-th level n excursion is up-down and $V_j^n = 1$ if it is down-up). Then it is known that [52]

THEOREM 9. *Brownian motion is the unique continuous process B for which*

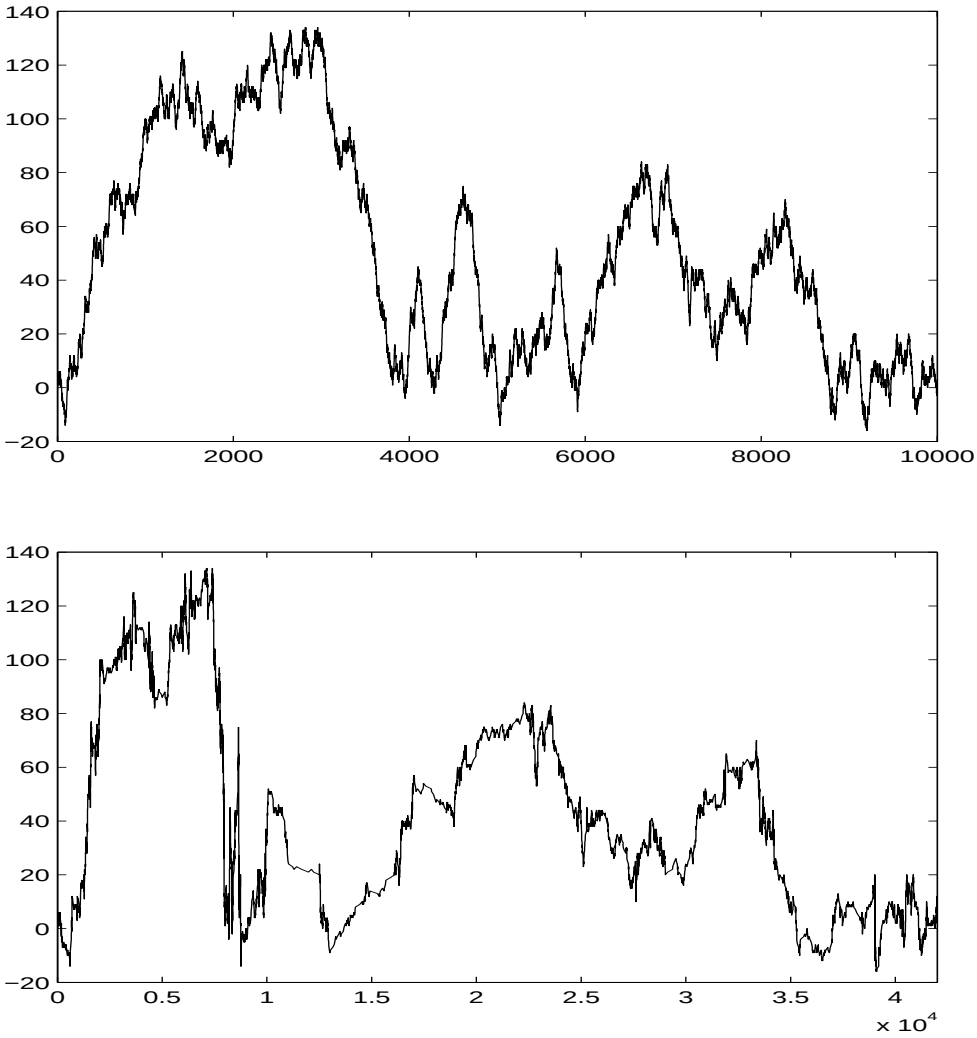


Figure 3.10: Top figure: CEBP process with offspring distribution $2\,\text{geometric}_1(0.6)$ and excursions $+/-$ or $-/+$ with equal probability, independent of Z_k^n. Bottom figure: MEBP process obtained from a multifractal time change of the top CEBP process, with i.i.d. gamma distributed weights.

- $B(0) = 0$ *and* $B(1)$ *has unit variance;*
- *For each* n*, the* D_k^n*,* $k = 1, 2, \ldots$ *are i.i.d. with Laplace transform* $\mathbb{E}e^{-sD_k^n} = 1/\cosh(2^{-n+1/2}s^{1/2})$*;*
- *The* Z_k^n *are i.i.d. for all* n *and* k*, with* $\mathbb{P}(Z_k^n = 2i) = 2^{-i}$*,* $i = 1, 2, \ldots$*;*
- *The* V_j^n *are i.i.d. for all* n *and* j*, with* $\mathbb{P}(V_j^n = 0) = \mathbb{P}(V_j^n = 1) = 1/2$*.*

For fractional Brownian motion, there is no equivalent result concerning its branching structure. We propose to estimate empirically the crossing durations and the offspring distribution of an fBm with parameter $H = 0.7$ and then construct an MEBP process that imitates this behaviour. We then estimate the correlation structure, the marginals and the partition function of the MEBP process and compare it with a theoretical fBm. We generate an fBm using the method proposed by Wood and Chan [116]. This algorithm returns samples at regular times and differs from the MEBP algorithm which simulates crossings and returns samples on an irregular lattice.

Consider a node $\mathbf{i} \in \Upsilon_{n-1}$ on generation $n-1$ of the crossing tree starting from a level 0 crossing and its k-th child $\mathbf{i}k$. Denote by $D_{\mathbf{i}k}^n$ the crossing duration associated with node $\mathbf{i}k$. Recall that $Z_{\mathbf{i}}$ denotes the number of offspring of node $\mathbf{i}$. We compute the average duration of subcrossings for an fBm after conditioning on the family size of the parent crossing, and renormalizing by a factor α:

$$f(z) := \alpha \mathbb{E}(D_{\mathbf{i}k}^n \mid Z_{\mathbf{i}} = z) \tag{3.10}$$

where α ensures $\sum z f(z) \mathbb{P}(Z = z) = 1$. The motivation for scaling by α is explained later. The expectation is estimated over one level of the crossing tree only, so that we do not have to worry about scaling the crossing durations amongst different levels. The choice of the level is important however. If we pick up a level at a large scale, the variance of the estimation is large since we have fewer crossings available. If the level corresponds to a fine scale, then one can miss crossings using the Chan-Wood algorithm since the samples are generated on a regular vertical lattice. This is illustrated in Figure 3.11.

Now we turn to the estimation of $f(z)$ of an fBm with Hölder exponent $H = 0.7$. Denote it by $\hat{f}_{\text{fBm}}(z)$. The results presented in Figure 3.13(a) are obtained by averaging 80 estimates of $f(z)$ for the Brownian motion and for the fBm (H=0.7), each of length 2^{20}. The expectation is estimated using the level 6 crossing (crossings of size 64), following the previous remarks. After normalization, we obtain an average duration of 0.26 for Brownian motion (bottom curve), which is close to the theoretical value $1/\mu = 0.25$. For fBm with $H = 0.7$, it is interesting to note that the average crossing duration is much shorter when $Z = 2$ and seems to reach a constant value of 0.45 for larger Zs ('o'). The previous observation remains true for other values of $H > 0.5$, but we do not present the curves here.

We also estimate the offspring distribution of an fBm with $H = 0.7$, say $\hat{p}$. The table below summarizes the numerical values obtained for $\hat{p}$ and $\hat{f}_{\text{fBm}}(z)$.

z	2	4	6	8	10	12	> 12
$\hat{p}(z)$	0.758	0.171	0.050	0.015	0.004	0.001	0.001
$\hat{f}_{\text{fBm}}(z)$	0.328	0.419	0.434	0.444	0.451	0.450	0.459

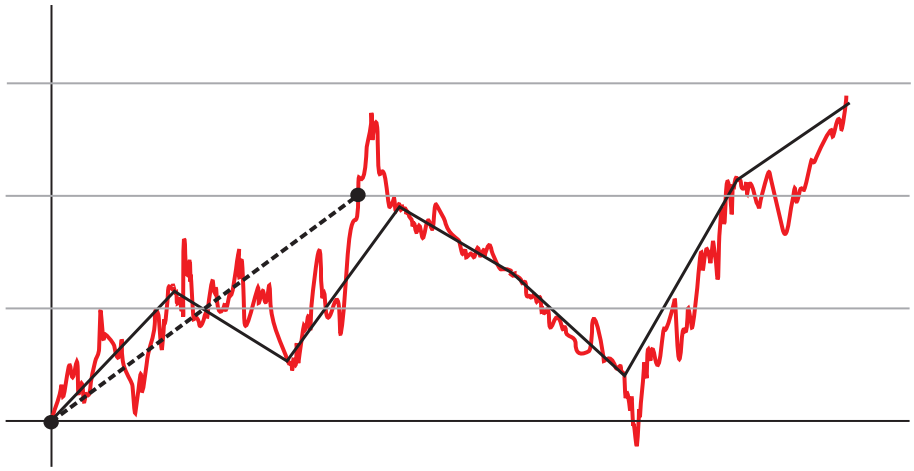


Figure 3.11: The figure displays the first few samples of an fBm generated on a regular lattice (black line), which is a discrete version of a realisation of a continuous fBm (in orange). Since the samples are generated at regular times and not on a regular horizontal lattice, when looking at crossings at the resolution of the process, we are likely to miss some of them. This is illustrated by the crossing in dot line which is missed by the discrete process.

Based on the results of the Brownian motion, one may hope that the offspring distribution of an fBm with Hölder exponent H is geometrical with parameter $p = 2^{1-1/H}$, parameter obtained by solving $H = \log 2/\log \mu$ where $\mu = 2/p$ for a $2\,\text{geometric}_1(p)$ distribution. Figure 3.12 shows the deviation for fBm from this distribution, as H and the number of offspring increases.

To imitate the fBm, we define weights ρ on the crossing tree as follows

$$\rho \mid (Z = z) = \hat{f}_{\text{fBm}}(z).$$

Weights are deterministic after conditioning on Z. With this definition, scaling by α in (3.10) ensures $\mathbb{E}\sum_{i=1}^{Z} \rho_i = 1$, which is a key assumption in the existence of MEBP processes (see Assumption 1 in the previous chapter). For the Brownian motion, weights are constant equal to $1/\mu$, as suggested by the empirical $f(z)$. By defining weights that way, the expected subcrossing duration given the family size of the parent crossing of the MEBP process is then

$$\mathbb{E}(\rho_{\mathbf{i}k}\mathcal{W}_{\mathbf{i}k} \mid Z_{\mathbf{i}} = z) = \hat{f}_{\text{fBm}}(z)\,[\mathbb{E}\rho]^n \tag{3.11}$$

and (3.11) and (3.10) match.

We now generate an MEBP process which imitates an fBm with $H = 0.7$. Let $\hat{f}_{\text{fBm}}(12) = \hat{f}_{\text{fBm}}(14) = 0.453$ and use $\hat{p}$ for the offspring distribution. In our simulations, we use the approximation $\mathbb{P}(Z > 12) = \mathbb{P}(Z = 14) = 0.001$.

Figure 3.13(b) presents $\hat{f}_{\text{fBm}}(z)$ and $\hat{f}_{\text{MEBP}}(z)$, which is the estimation of $f(z)$ for the MEBP process. We consider MEBP processes of length 2^{20}, estimate $f(z)$ and build confidence intervals using 80 independent realisations. Estimation is good for $z = 2$ and $z = 4$ but the average crossing duration tends to be larger than expected

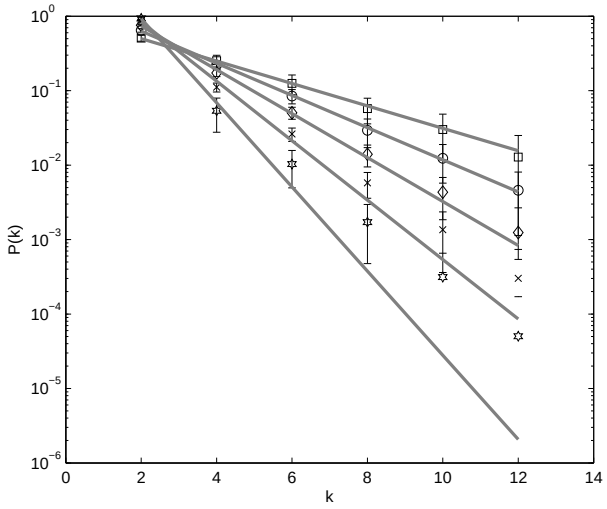
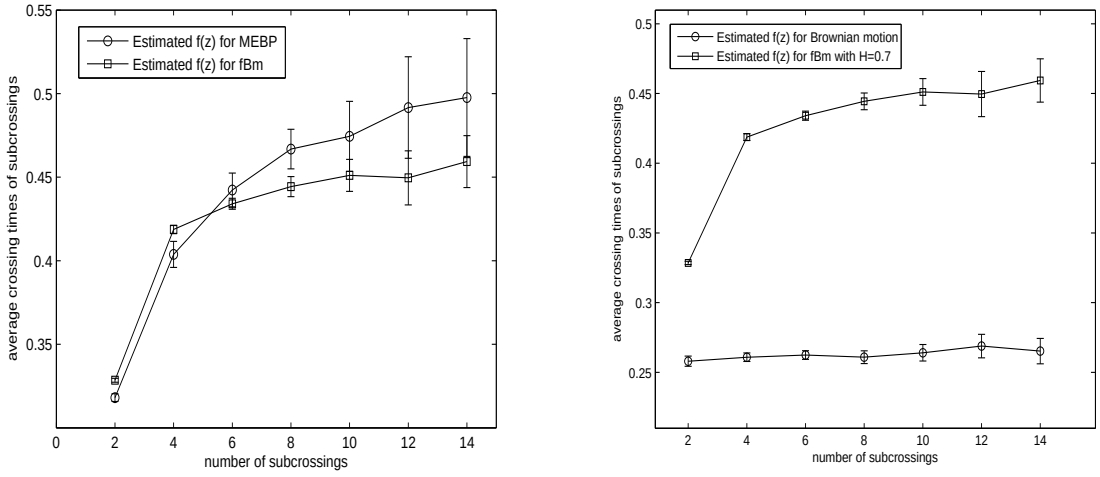


Figure 3.12: Estimated offspring distribution from fBm time series. From top to bottom, $H = 0.5$ (square), $H = 0.6$ (circle), $H = 0.7$ (diamond), $H = 0.8$ (cross) and $H = 0.9$ (star). The solid grey lines are the geometrical distribution $P(Z = k) = p(1-p)^{(k-2)/2}$, $k = 2, 4, 6, \ldots$ with parameter $p = 2^{1-1/H}$.



(a) (b)

Figure 3.13: Estimation of the average crossing times of subcrossings for fBm and MEBP. Figure (a) presents $\hat{f}_{\mathrm{fBm}}(z)$ for $H = 0.7$ (top curve) and $H = 0.5$ (bottom curve). For Brownian motion ($H = 0.5$), $\hat{f}_{\mathrm{fBm}}(z)$ is constant and close to the theoretical value $1/\mu = 0.25$. The case $H = 0.7$ shows that the average crossing time of subcrossings is smaller when $Z = 2$ and tends to reach a constant value as Z increases. Figure (b) displays the plots $\hat{f}_{\mathrm{fBm}}(z)$ and $\hat{f}_{\mathrm{MEBP}}(z)$ for a fitted MEBP. The two curves follow the same trend, with small differences for larger values of Z.

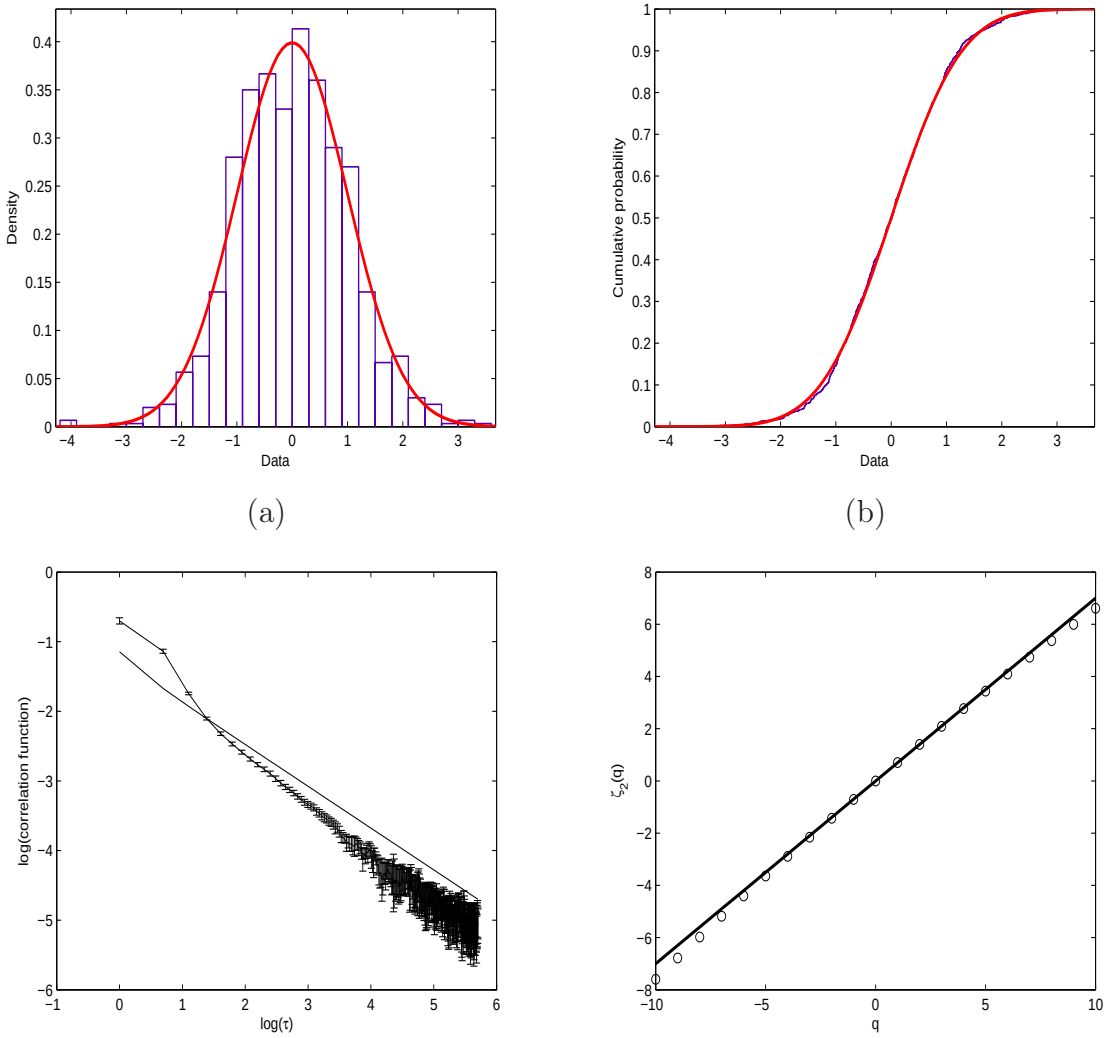


Figure 3.14: The figures present the marginal distribution, the correlation function and the partition function of an MEBP process which imitates an fBm with $H = 0.7$. Figure (a) displays the empirical marginal distribution with a fitted normal distribution. Figure (b) plots the empirical cdf with the fitted one. In Figure (c) we show the estimated corrolation function of the MEBP together with the exact correlation structure of an fBm with $H = 0.7$. Finally, Figure (d) presents the partition function of the MEBP estimated using wavelet leaders (dots) with the straight line with slope 0.7, which is the partition function of an fBm with $H = 0.7$.

for larger values of z, with larger confidence intervals. This can be due to imprecisions in the offspring distribution we used, but also from the dependency structure for MEBP and fBm, which are different.

We now estimate the marginal distribution of the MEBP process at a fixed time t. Since the MEBP algorithm returns samples on an irregular time grid, we do a linear interpolation between two consecutive samples. We arbitrarily set $t = 500$ and we denote by $X_1, \dots, X_n$ the random sample we collect. Note that the conclusions of this paragraph remain unchanged if we set $t = 1000$ and $t = 10000$. We have not tested the dependence of the results on n. Let $\bar{X}$ be the sample mean and S^2 the sample variance. We calculate for $i = 1, \dots, n$

$$Y_i = \frac{X_i - \bar{X}}{S}$$

and we fit a normal distribution, shown in Figure 3.14(a). To test the closeness of the sample to a normal distribution, we consider two tests based on differences between the empirical distribution function $F_n(x)$ and the target distribution $\Phi(x)$: the Kolmogorov-Smirnov and Anderson-Darling tests. Figure 3.14(b) also displays the empirical and fitted cumulative distribution functions, which match closely.

Kolmogorov-Smirnov Test. Consider the maximum absolute difference D_n between $F_n(x)$ and $\Phi(x)$,

$$D_n = \sup_x |F_n(x) - \Phi(x)|.$$

Then the exact distribution of D_n can be derived and is independent of the target distribution (see Kolmogorov [75]). Consider the modified statistic

$$\mathcal{D}_n = D_n(\sqrt{n} - 0.01 + 0.85/\sqrt{n}).$$

Let $F(x)$ be the cdf of the sample. We test the hypothesis H_0: $F(x) = \Phi(x)$ against the two sided alternative H_1: $F(x) \neq \Phi(x)$. We accept H_0 if $\mathcal{D}_n \leqslant z_\alpha$ where $P(\mathcal{D}_n \leqslant z_\alpha) = \alpha$. Values of z_α can be found in [34] for example. For $n = 1000$, $\alpha = 0.95$, with unknown mean and variance we have $z_\alpha = 0.895$. With the data presented in Figure 3.14(a), we obtain $\mathcal{D}_n = 0.723$ and we accept H_0 with 95% confidence.

Anderson-Darling Test. This test measures the discrepancy by assigning weights to the squared difference $(F_n(x) - \Phi(x))^2$. It gives more importance to the tails of the distribution than the previous test. Let $X_1 \leqslant \ldots \leqslant X_n$ be the ordered statistic and consider $U_i = \Phi(X_i)$. The asymptotic distribution of

$$W_n^2 = -n - \frac{1}{n}\sum_{j=1}^{n}(2j-1)[\log U_j + \log(1 - U_{n-j+1})]$$

is known [5]. For the case of unknown mean and variance, consider the modified statistic

$$\mathcal{W}_n^2 = W_n^2(1 + 0.75/n + 2.25/n^2).$$

We test the same two hypothesis H_0 and H_1 as before and want a 0.95 confidence level. Then $z_{0.95} = 0.752$ [34] and the data yields $\mathcal{W}_n^2 = 0.558$. Here again we accept H_0 with 95% confidence.

Figure 3.14(c) presents the correlation structure of the increments of the MEBP process on a log-log plot and the theoretical correlation of a fGn $\Delta B_H(t) := B_H(t+1) - B_H(t)$ where $B_H(t)$ is an fBm with $H = 0.7$. Its correlation structure is given by [89]

$$\mathbb{E}[\Delta B_H(t)\Delta B_H(t+\tau)] = \frac{1}{2}V_H(|\tau+1|^{2H} + |\tau-1|^{2H} - 2|\tau|^{2H}) \tag{3.12}$$

where $V_H = \mathrm{Var}[B_H(1)]$. For $H > 0.5$ and large τ, the correlation is positive and decays like τ^{2H-2}. The estimated correlation is averaged over 100 estimates, for signals of length 1000. The estimated correlation is displayed for $\tau = 1 \ldots 300$, together with the theoretical one given by Equation (3.12) and 95% confidence intervals. It appears to decay linearly with $\log \tau$, with a small departure from the slope $2H - 2$. This small departure shows the different corrolation structure of fBm and MEBP processes. Maybe we could consider in another study a more general model and allow correlation in the Z_k^n to take this observation into account.

Finally, Figure 3.14(d) estimates the partition function $\zeta_2(q)$ with wavelet leaders. We generate 100 independent samples of length 2^{14}, and perform its wavelet transform from scale 3 to 12. We use Daubechies wavelets with two vanishing moments and estimate the spectrum for $q \in [-10, 10]$. We used a linear interpolation to go from the irregular time grid generated by the MEBP algorithm to a regular lattice. The partition function appears linear, which is expected for monofractal processes such as fBm. In this case, the slope of the partition function gives the Hölder exponent of the fBm (see the end of Section 4.1 in the next chapter). A mean square regression estimate of the slope of the partition function gives $\hat{H} = 0.709$, close to the theoretical value. The slight departure from linearity observed for large negative values of q probably comes from the linear interpolation. We will notice again in the next chapter the high sensitivity of the estimation to the step of linear interpolation. Oversampling or undersampling can discard or add extra information about the fractal structure of the process, which has an impact when estimating the spectrum.

Overall, the statistical tests performed indicate the MEBP process here defined is a good approximation of an fBm with parameter $H = 0.7$, with the advantage of an on-line simulation. To illustrate, two realisations of an fBm using the Chan-Wood algorithm and with the MEBP algorithm are presented in Figure 3.15. The advantage of our algorithm is that it provides an on-line simulation of an approximated fBm.

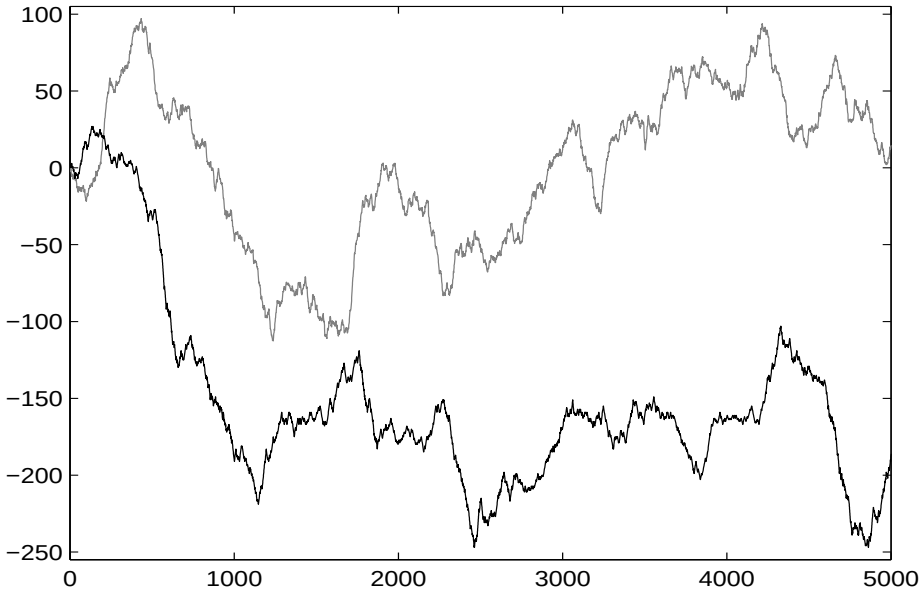


Figure 3.15: A realisation of fBm generated using the Chan-Wood algorithm (in grey) and a realisation of an MEBP process imitating an fBm (in black).

Chapter 4

Multifractal formalism for CEBP and MEBP processes

> "Abstract painting is abstract. There was a reviewer a while back who wrote that my pictures didn't have any beginning or any end. He didn't mean it as a compliment, but it was."
>
> Jackson Pollock

In practice, we are interested in deciding the nature of the fractal structure of a random process, and in particular, we want to know whether a stochastic process is mono or multifractal. The aim of this chapter is to answer this question theoretically for both the CEBP and MEBP processes. In the first section, we explain the terminology mono and multifractal with the definition of the Hausdorff spectrum. In the second, we prove that CEBP processes have a constant local Hölder regularity along their sample paths. In the third section, we introduce a multifractal formalism adapted to the study of MEBP and derive an upper bound of its Hausdorff spectrum. Simulations using the wavelet leader technique support the theory.

4.1 Monofractals versus Multifractals

As stated in the introduction, the local Hölder exponent is an efficient tool to measure the local fluctuations of a process, by comparing it to a power law. Unless stated otherwise, we work on a realisation of a continuous process. Random variables take values in $\mathcal{C}(\mathbb{R})$, the space of continuous functions of $\mathbb{R}$. The local Hölder exponent at time t is defined as

Definition 8. [102] *If there exists a polynomial P_t such that*

$$|X(u) - P_t(u)| = O(|u-t|^h) \text{ as } u \to t$$

we say that the function X is in C_t^h. The local Hölder exponent of X at time t is defined to be the supremum of those h such that X is in C_t^h:

$$H(t) := \sup\{h \mid X \in C_t^h\}.$$

When the polynomial P_t is constant with respect to u, $P_t(u) = X(t)$. This motivates the following definition

DEFINITION 9. *Let* $\epsilon > 0$, $I_\epsilon = [t-\epsilon, t+\epsilon]$ *and* $X \in \mathcal{C}(\mathbb{R})$

$$\begin{aligned} h(t) &:= \liminf_{\epsilon \to 0} \frac{1}{\log(2\epsilon)} \log \sup_{|u-t|<\epsilon} |X(u) - X(t)| \\ &= \liminf_{\epsilon \to 0} \frac{1}{\log(2\epsilon)} \log \sup_{(u,u') \in I_\epsilon} |X(u) - X(u')| \end{aligned}$$

with the convention that $\log(0) = -\infty$.

When X is a non decreasing process, $h(t)$ reduces to

$$h(t) = \liminf_{\epsilon \to 0} \frac{1}{\log(2\epsilon)} \log[X(t+\epsilon) - X(t-\epsilon)]. \tag{4.1}$$

Note that we write $h_X(t)$ to indicate $h(t)$ for a function or path X.

When the polynomial P_t is known to be constant, then $h(t) = H(t)$. In fact, Riedi proved the stronger result that when $h(t)$ is not an integer, then P_t is constant and $h(t) = H(t)$ [102]. In the remainder we call $h(t)$ the Hölder exponent. An interesting geometrical characteristic for irregular processes is the multifractal spectrum or Hausdorff spectrum. The spectrum describes for each $a > 0$, the size of the set of points Θ^a where X has local Hölder exponent a

$$\Theta^a := \{t \mid h(t) = a\}. \tag{4.2}$$

The measurement of the size of Θ^a is provided by the Hausdorff dimension, hence the definition of the multifractal spectrum

DEFINITION 10. *The Hausdorff spectrum* f_H *of the function* X *at the point* a *is the Hausdorff dimension* $D_{\mathcal{H}}$ *of* Θ^a

$$D(a) := D_{\mathcal{H}} \Theta^a.$$

A stochastic process is said to have a multifractal structure if the sets Θ^a are highly interwoven. The spectrum thus describes the connection between the local fluctuations of a process and the measure of the set of points that have the same local variability. Generally, sets Θ^a are dense in the support of the process. It follows that measuring the size of sets Θ^a using box dimension would be inappropriate because $\dim_B \Theta^a$ would equal the box dimension of the support of the process [45]. When the Hausdorff spectrum degenerates to a single point, we say that the process is monofractal, since it possesses only one Hölder exponent.

In general, it is hard to calculate or estimate directly the Hausdorff spectrum of a function X since it is not possible to locate precisely the points of a process with a given regularity, and it is difficult to estimate the Hausdorff dimension of a set. In practice, the spectrum is calculated using a multifractal formalism, which relates the spectrum to the Legendre Fenchel (LF) transform of a partition function defined

as the power exponent of a structure function $T_X(a,t)$ in the limit of small scales. $T_X(a,t)$ is obtained from a comparison between the original process and a reference pattern dilated and located at different positions. The use of wavelets is useful for the analysis of a process at different scales and one can use for example the wavelets coefficients $d_X(a,t_k)$ or wavelet leaders $L_X(a,t_k)$ for $T_X(a,t_k)$. The characterization of the Hölder exponent using wavelets requires a regularity hypothesis that is stronger than continuity: X is said uniform Hölder if (Definition 1, Chapter 1)

$$\exists \epsilon > 0 \ \exists C > 0 \ \forall t, s \in \mathbb{R}, \ |X(t) - X(s)| \leqslant C|t-s|^{\epsilon}.$$

Recall the definition of $S_1(q,j)$ and $S_2(q,j)$ given in the Introduction,

$$S_1(q,j) = \frac{1}{n_j}\sum_{k=1}^{n_j} |d_X(j,t_k)|^q \qquad S_2(q,j) = \frac{1}{n_j}\sum_{k=1}^{n_j} |L_X(j,t_k)|^q.$$

We identify their power law behaviour by defining the partition functions $\zeta_1(q)$ and $\zeta_2(q)$

$$\zeta_1(q) = \liminf_{j\to-\infty} \big(\frac{\log_2 S_1(q,j)}{j}\big) \qquad \zeta_2(q) = \liminf_{j\to-\infty} \big(\frac{\log_2 S_2(q,j)}{j}\big). \tag{4.3}$$

Also, we have introduced earlier the LF transform $\xi(q)$ of the spectrum $D(a)$

$$\xi(q) = 1 + \inf_a (qa - D(a)). \tag{4.4}$$

Then it is known that for all uniform Hölder functions or processes, the LF transform of $\zeta_2(q)$ provides an upper bound for the Hausdorff spectrum

$$D(a) \leqslant \inf_q (1 + qa - \zeta_2(q)) \tag{4.5}$$

since $\zeta_2(q) = \xi(q)$ for all $q \in \mathbb{R}$ [3]. When the Hausdorff spectrum is concave, then the LF transform is involutive, and inequality in (4.5) is replaced by equality. This is true for most multiplicative cascades. Also, when X possesses no oscillating singularities, then $\zeta_1(q) = \xi(q)$ for all $q \geqslant 0$. When the inequality in (4.5) is an equality, the multifractal formalism is said to hold. When deciding whether a process is mono or multifractal, we can use either the spectrum or the partition function. Assuming the multifractal formalism holds and the partition function is reduced to a linear function of q, $D(h)$ reduces to a single point and the stochastic process is said to be monofractal. In this case, the slope of the partition function gives the Hölder exponent of the process. When the partition function is non linear, the spectrum is non trivial and the process is said to be multifractal.

4.2 The CEBP process

In this section we prove the monofractal structure of CEBP processes. First, we need a preliminary result on the random variables $W_{\mathbf{i}}$, introduced in the previous chapter, end of Section 3.1.1, as the limit of a martingale defined on a Galton-Watson tree. We keep the same notations as before.

ASSUMPTION 4. *Suppose for all $p > 0$, $\mathbb{E}Z_\emptyset^p < \infty$.*

LEMMA 6. *Let Υ be a Galton-Watson tree with regular offspring distribution. Under Assumption 4,*

$$\log \max_{\mathbf{i}\in\Upsilon_n^{GW}} W_\mathbf{i} = o(n) \qquad \textit{and} \qquad \log \min_{\mathbf{i}\in\Upsilon_n^{GW}} W_\mathbf{i} = o(n)$$

almost surely with respect to trees.

Proof. The proof uses an embedded branching random walk (BRW) [23]. We recall that a BRW is described by a triple (Q, M, χ), whose elements respectively describe the reproduction, movement and importance of each individual. We label individuals using elements of $U = \cup_{n\geqslant 0}\mathbb{N}^n$ (as for the crossing tree). With each individual $\mathbf{i}$ we associate an independent version of (Q, M, χ), denoted $(Q_\mathbf{i}, M_\mathbf{i}, \chi_\mathbf{i})$. Q is a point process on $\mathbb{R}\times\mathbb{R}^+$, its first coordinate gives a child's displacement from its parent's birth position and the second gives the parent's age at that child's birth. M is a stochastic process which describes an individual's movement during its lifetime. $\chi : \mathbb{R}^+ \to \mathbb{R}$ is the so-called characteristic function.

We construct a BRW as follows. The point process $Q_\mathbf{i}$ generates $Z_\mathbf{i}$ points at time 1, with no displacement. The birth time of individual $\mathbf{i}$ is just $|\mathbf{i}|$. Individual movement consists of a single jump at birth of size $\log \mathcal{V}_\mathbf{i}$, where the $\mathcal{V}_\mathbf{i}$ are i.i.d. and distributed like $W_\emptyset$. That is, $M_\mathbf{i}(t) = \log \mathcal{V}_\mathbf{i}$ for all $t \geq 0$. The characteristic χ is used to count generations: $\chi_\mathbf{i}(t) = \mathbb{I}_{[0,1)}(t)$.

The right-most particle of the BRW at time t is defined as

$$B_t = \max_{\mathbf{i}\in\Upsilon_{\lfloor t\rfloor}} \log(\mathcal{V}_\mathbf{i}).$$

Biggins [23] studied the behaviour of the right-most particle under the condition that the BRW is 'well-regulated' and derived conditions of the existence of a γ such that $B_t/t \to \gamma$ a.s. In our case, $m(\theta,\phi) := \mathbb{E}\int e^{-\theta z-\phi\tau}Q(dz, d\tau) = e^{-\phi}$ and $\alpha(\theta) := \inf\{\phi \mid m(\theta,\phi) \leqslant 1\} = 0$. (Q, M, χ) is said to be well-regulated if for all $\theta \leqslant 0$,

$$\mathbb{E}\Big(\sup_t \big\{e^{-\alpha(\theta)t}e^{-\theta M(t)}\chi(t)\big\}\Big) = \mathbb{E}\Big(\sup_t \big\{e^{-\theta M(t)}\chi(t)\big\}\Big) < \infty.$$

With the present definition of characteristic, this reduces to checking that $\mathbb{E}\mathcal{V}_\emptyset^{-\theta}$ for all $\theta \leqslant 0$. The regularity condition on the offspring distribution ensures that $\mathbb{E}\mathcal{V}_\emptyset = 1$. Moreover, Bingham and Doney [27] have shown that for all $p > 1$, $\mathbb{E}Z_\emptyset^p$ and $\mathbb{E}\mathcal{V}_\emptyset^p$ converge or diverge together. The process is thus well regulated under Assumption 4.

Let $\alpha^*(a) := \inf_{\theta<0}\{a\theta + \alpha(\theta)\}$, that is

$$\alpha^*(\mathrm{a}) = \begin{cases} 0 & \text{if } a \leqslant 0 \\ -\infty & \text{if } a > 0 \end{cases}$$

Then from [23],

$$\gamma = \inf\{a \mid \alpha^*(a) < 0\} = 0$$

and $\log\max_{\mathbf{i}\in\Upsilon_n^{GW}} W_\mathbf{i} = o(n)$ follows.

To prove that $\log \min W_{\mathbf{i}} = o(n)$, we note that $\min(\log W_{\mathbf{i}}) = -\max(-\log W_{\mathbf{i}})$. We consider the same BRW as before, but with jumps of magnitude $-\log \mathcal{V}_{\mathbf{i}}$. We still have $\gamma = 0$, provided that the branching random walk is 'well-regulated', which is the case if $\mathbb{E}(\mathcal{V}^{\theta}) < \infty$ for all $\theta < 0$. This follows from Biggins and Bingham [24], Theorem 3, which shows that the left tail of $\mathcal{V}_{\emptyset}$ decays exponentially as long as the family size is greater than 2, which is verified for the crossing tree of CEBP processes. □

We can now state the main result of this section.

THEOREM 10. *Let X be a CEBP with regular offspring distribution p and mean family size μ. Suppose the number of offspring is bounded, that is there exists M such that $p(x) = 0$ for all $x \geqslant M$. Also, suppose Assumption 4 holds. Then, with probability* 1*, for all t,*

$$h(t) = \frac{\log 2}{\log \mu}. \tag{4.6}$$

In other words, X is monofractal.

Proof. Let $\epsilon > 0$. It is always possible to find a positive $n^*(\epsilon)$ such that

$$2^{-(n^*+1)} < \sup_{\substack{|u-t|<\epsilon \\ u,t\in[0,W_{\emptyset}]}} |X(u) - X(t)| \leqslant 2^{-n^*}. \tag{4.7}$$

Reciprocally, given n^*, ϵ is not unique, hence we define $I_{n^*} := \{\epsilon \mid n(\epsilon) = n^*\}$. We need to derive bounds for I_{n^*}.

Suppose

$$\epsilon \leqslant \min_{\mathbf{j}\in\Upsilon^{GW}_{n^*+2}} \mu^{-(n^*+2)} W_{\mathbf{j}} \tag{4.8}$$

and consider crossings of size $2^{-(n^*+2)}$. If ϵ satisfies (4.8), then the interval (u, t) intersects at most two intervals corresponding to generation $n^* + 2$ and it follows that

$$\sup_{\substack{|u-t|<\epsilon \\ u,t\in[0,W_{\emptyset}]}} |X(u) - X(t)| \leqslant 2.2^{-(n^*+2)} = 2^{-(n^*+1)}$$

which contradicts (4.7). Therefore, if $\epsilon \in I_{n^*}$ then

$$\epsilon > \min_{\mathbf{j}\in\Upsilon^{GW}_{n^*+2}} \mu^{-(n^*+2)} W_{\mathbf{j}}.$$

Suppose

$$\epsilon > \mu^{-n^*} \max_{\mathbf{i}\in\Upsilon^{GW}_{n^*}} [W_{\mathbf{i}} + \ldots + W_{\mathbf{i}+M}, W_{\mathbf{i}} + \ldots + W_{\mathbf{i}-M}]$$

where $\mu^{-n^*} W_{\mathbf{i}+M}$ is the length of the M-th consecutive interval to the right of $\mu^{-n^*} W_{\mathbf{i}}$, corresponding to level n^* crossing, and $\mu^{-n^*} W_{\mathbf{i}-M}$ the length of its M-th consecutive interval to the left. When these intervals do not exist (which can be the case at the two ends of the process), then omit them. Thus

$$\sup_{\substack{|u-t|<\epsilon \\ u,t\in[0,W_{\emptyset}]}} |X(u) - X(t)| \geqslant 2^{-n^*}$$

θ	0.2	0.5	0.7
Theoretical slope	0.3010	0.5	0.6603
Estimated slope	0.2957	0.4909	0.6675

Table 4.1: Estimated slope of the partition function of a CEBP for different values of θ, using a least squares regression. Estimates are close to the theoretical value $\log 2/\log \mu$ with $\mu = 2/\theta$.

which contradicts (4.7). Thus if $\epsilon \in I_{n^*}$ then

$$\epsilon \leqslant \mu^{-n^*} \max_{\mathbf{i} \in \Upsilon_{n^*}^{GW}} [W_{\mathbf{i}} + \ldots + W_{\mathbf{i}+M}, W_{\mathbf{i}} + \ldots + W_{\mathbf{i}-M}] \leqslant M\mu^{-n^*} \max_{\mathbf{j} \in \Upsilon_{n^*}^{GW}} W_{\mathbf{j}}.$$

In summary, for $\epsilon \in I_{n^*}$, we have

$$\min_{\mathbf{j} \in \Upsilon_{n^*+2}^{GW}} \mu^{-(n^*+2)} W_{\mathbf{j}} < \epsilon \leqslant M\mu^{-n^*} \max_{\mathbf{j} \in \Upsilon_{n^*}^{GW}} W_{\mathbf{j}}.$$

Under Assumption 4, it follows that

$$-(n^*+2)\log \mu + \log 2 + o(n^*) < \log(2\epsilon) < -n^* \log \mu + \log(2M) + o(n^*). \quad (4.9)$$

Taking the logs of (4.7) and combining it with (4.9) yields

$$\frac{-(n^*+1)\log 2}{-n^* \log \mu + \log(2M) + o(n^*)} < \frac{\log \sup |X(u) - X(t)|}{\log(2\epsilon)} < \frac{-n^* \log 2}{-(n^*+2)\log \mu + \log 2 + o(n^*)}.$$

From the almost sure continuity of X, $n^*(\epsilon) \to \infty$ as $\epsilon \to 0$, and the result follows. □

Figure 4.1 displays the estimation of the partition function $\tau(q)$ of CEBP processes with geometric offspring distribution, with parameter $\theta = 0.2, 0.5$ and 0.7,

$$P(k) = \theta(1-\theta)^{(k-2)/2}, \ k = 2, 4, 6, \ldots$$

Note that the case $\theta = 0.5$ corresponds to the Brownian motion. The estimation is performed using the wavelet leader technique described earlier, assuming the CEBP process is Hölder uniform. In each case, the results are obtained by averaging 100 realisations, each realisation of length 2^{17}. We used Daubechies wavelets with 3 vanishing moments. The wavelet leaders $L_X(j,k)$ are calculated over the range of scales $j_1 \leqslant j \leqslant j_2$ with $j_1 = 3$ and $j_2 = 14$. As predicted, $\tau(q)$ is linear in q, with slope $\log 2/\log \mu$. The slope is estimated using a least squares regression and presented in table 4.1.

In the simulations, we have used a geometric offspring distribution which is not bounded. The numerical results show that this assumption on the offspring distribution can probably be weakened.

To further test the monofractal behaviour of CEBP processes, we perform a multifractality test proposed by Wendt and Abry [114], briefly described in section 2.5.3,

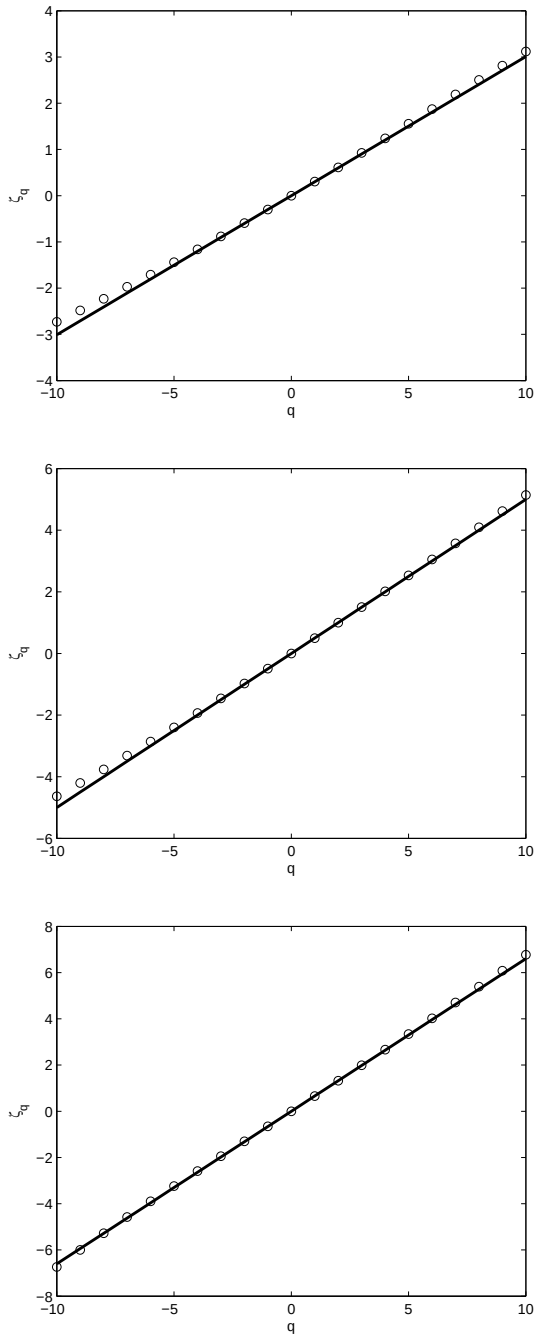


Figure 4.1: **Partition function of CEBP processes.** Offspring distribution is $P(Z = k) = \theta(1-\theta)^{(k-2)/2}$, with $k = 2, 4, 6, \ldots$ and $\theta = 0.2, 0.5$ and 0.7, from top to bottom figures. Estimation ('o') is from an average of 100 realisations. The theoretical function is the solid black line. The partition function is linear in q, which characterizes a monofractal stochastic processes. Note that confidence intervals are too small to be plotted.

θ	0.2	0.5	0.7
$\hat{\alpha}$ for c_2	0.12	0.06	0.08
$\hat{\alpha}$ for c_3	0.11	0.10	0.11

Table 4.2: Rejection of hypothesis $c_2 = 0$ and $c_3 = 0$ at 10% significance.

using the Matlab toolbox developed by Wendt in his PhD [113]. Consider the partition function $\zeta_i(q)$, $i = 1$ or 2 whose definitions are given by Equations (4.3), and their polynomial expansion $\zeta_i(q) = c_1 q + c_2 q^2/2 + c_3 q^3/3 + \dots$ The test allows us to decide whether $c_p = 0$ or not. In particular, assuming the multifractal formalism holds, the case $p \geqslant 2$ distinguishes between mono and multifractal behaviour.

Of the six tests proposed by Wendt and Abry, we chose the one called 'Percentile Bootstrap Test'. We test the null hypothesis 100 times for $p = 2$ and 3 at a significance level of $\alpha = 0.1$ and report the rejection rate. We use wavelet coefficients, however wavelet leaders would give similar results. Results are given in table 4.2. In the three considered cases, the average rejection rate $\hat{\alpha}$ is close to the theoretical value, agreeing with the monofractal behaviour of CEBP processes established in Theorem 10.

4.3 Multifractal formalism for MEBP

4.3.1 Multifractal formalism

There are many technical difficulties when trying to estimate the Hausdorff spectrum of a measure. Not only is it impossible to calculate the Hausdorff dimension of a set in any real-world situation, but also the local regularity of a multifractal stochastic process at each point cannot be estimated numerically. From these difficulties appear indirect methods for calculating the spectrum.

An alternative description of the singular structure of multifractals is provided by the coarse theory which splits the support of the function or the measure with r-mesh cubes, counts how many of them have a measure with Hölder exponent of order a, and then lets $r \to 0$. The coarse theory is more easily computed theoretically. In practice, one needs a partition function to estimate it.

Let X be a stochastic process defined on $\mathbb{R}$, whose realisations belong to $\mathcal{C}(\mathbb{R})$, the space of continuous functions on $\mathbb{R}$. In the coarse theory, we are talking about path by path realisation of the process. Consider the increment process ΔX_r defined over the r-mesh cubes $I_r^n = [nr, (n+1)r)$. If $t \in I_r^n$, then $\Delta X_r(n) = X((n+1)r) - X(nr)$. Then, the coarse spectrum of X is defined as [44]

$$f^\dagger(a) := \lim_{\epsilon \to 0} \lim_{r \to 0} \frac{1}{-\log r} \log[\#\{I_r^n \mid t \in I_r^n,\ a - \epsilon \leqslant \frac{\log |\Delta X_r(t)|}{\log r} \leqslant a + \epsilon\}]. \quad (4.10)$$

When the limit $\lim_{r \to 0}$ fails to exist, one can consider alternative definitions of $f^\dagger$ using $\liminf$ and $\limsup$, called respectively the lower and upper coarse spectrum. It is shown in [44], Lemma 11.1, that the fine and coarse spectra are related by

$$D(a) \leqslant f^\dagger(a).$$

Estimating the coarse spectrum directly from its definition is hard, particularly because of the presence of the double limit. In practice, one prefer to relate the coarse spectrum to the LF transform of the power law exponents of moment sums, which is faster and more manageable numerically. Let $S^r(q) = \sum^* |\Delta X_r(n)|^q$ where the $\sum^*$ is taken over all I_r^n for which $|\Delta X_r(n)| > 0$. Let $\beta_1(q) = \lim_{r\to 0} \log S^r(q)/\log r$, and call $\beta_1(q)$ a partition function. Then it is proved in [44], Lemma 11.2, that for any function in $\mathcal{C}(\mathbb{R})$,

$$f^\dagger(a) \leqslant f_L^\dagger(a) := \inf_{-\infty<q<\infty} (aq - \beta_1(q)).$$

The LF transform of $\beta_1(q)$ provides an upper bound of the Hausdorff spectrum, but there are many functions for which equality holds.

Consider an averaging of $S^r(q)$ over all realisations of X and denote this averaging by $\mathbb{E}$. $\beta_2(q) = \lim_{r\to 0} \log \mathbb{E}S^r(q)/\log r$ is called the deterministic partition function. Then, for any random process, with probability one (Lemma 3.9 in [102]),

$$\beta_1(q,\omega) \geqslant \beta_2(q) \text{ for all } q \text{ with } \beta_2(q) < \infty.$$

In the previous inequality, we have made the randomness of β_1 explicit by writing its dependence on ω. Let $\beta_2^*(a) := \inf_q(qa - \beta_2(q))$ be the LF transform of β_2. Since $\beta_1(q,\omega) \geqslant \beta_2(q)$ almost surely, we get $f_L^\dagger(a) \leqslant \beta_2^*(a)$. Relations between the different spectra introduced so far can be summarized as:

$$D(a) \leqslant f^\dagger(a) \leqslant f_L^\dagger(a) \leqslant \beta_2^*(a). \tag{4.11}$$

Examples where the previous inequalities are strict inequalities can be found in [101] and [104]. In the remainder, by multifractal formalism we mean a formalism where we calculate the Legendre-Fenchel transform of a partition function ($\beta_1(q)$ or $\beta_2(q)$ for example) which provides an upper bound for the multifractal spectrum, and we say that the multifractal formalism holds when the inequality between the two spectra is replaced by an equality.

Finally, define the deterministic coarse spectrum $F^\dagger$ of X as an averaged version of $f^\dagger$

$$F^\dagger(a) := \lim_{\epsilon\to 0}\lim_{r\to 0} \frac{1}{-\log r}\mathbb{E}\log[\#\{I_r^n \mid t \in I_r^n,\ a-\epsilon \leqslant \frac{\log \Delta X_r(n)}{\log r} \leqslant a+\epsilon\}]. \tag{4.12}$$

Theorem 3.13 and 3.14 in [102] state that $f^\dagger(a) \leqslant F^\dagger(a) \leqslant \beta_2^*(a)$.

In the classical definitions presented so far and encountered in the literature, coarse spectra are derived using a regular subdivision of the real line: intervals I_r^n all have length r. The mathematical study of the singularities of the chronometer $\mathcal{M}$ introduced in the previous chapter requires an irregular subdivision of its support. Recall that $\mathcal{M}$ is defined as the integral of a measure ζ, where $\zeta(R_{\mathbf{i}}) = \rho_{\mathbf{i}}\mathcal{W}_{\mathbf{i}}$ for all $\mathbf{i} \in \Upsilon_n^{GW}$. The length of interval $R_{\mathbf{i}}$ is given by $\mu^{-n}W_{\mathbf{i}}$ and varies from one $\mathbf{i}$ to another. In order to be able to obtain theoretical results about the coarse spectra of $\mathcal{M}$, it will be convenient to partition the real line with intervals $R_{\mathbf{i}}$. In section 4.3.3, we introduce definitions of coarse spectra and partition functions with an irregular subdivision of the real line. We show that inequalities (4.11) still hold for those definitions, and derive explicitly an upper bound for the fine spectrum of $\mathcal{M}$. First, a discrete version of the local Hölder exponent adapted to the study of $\mathcal{M}$ is needed.

4.3.2 Discretization of the local Hölder exponent

We reuse notations from the previous chapter. Let p be a regular offspring distribution and suppose Assumptions 1, 2 and 3 hold. Then the chronometer $\mathcal{M}$ has continuous sample paths and we consider its local Hölder exponent

$$h_{\mathcal{M}}(t) := \liminf_{\epsilon \to 0} \frac{1}{\log(2\epsilon)} \log(\mathcal{M}(t+\epsilon) - \mathcal{M}(t-\epsilon))$$

for all $t \in [0, W_\emptyset]$.

In the remainder, we always assume a regular offspring distribution and that Assumptions 1, 2 and 3 hold, which ensures that realisations of $\mathcal{M}$ belong to $\mathcal{C}(\mathbb{R})$.

Let $\mathbf{i} \in \Upsilon_n^{GW}$. Let $\psi(\mathbf{i})$ be the position of $\mathbf{i}$ in generation n. Denote by $\mathbf{i}-$ and $\mathbf{i}+$ respectively the left and right neighbours of $\mathbf{i}$ in generation n, so that $\psi(\mathbf{i}-) + 1 = \psi(\mathbf{i}) = \psi(\mathbf{i}+) - 1$.

Definition 11. *Let $\mathbf{i} \in \Upsilon_n^{GW}$ and $s_{\mathbf{i}} = \rho_{\mathbf{i}-}\mathcal{W}_{\mathbf{i}-} + \rho_{\mathbf{i}}\mathcal{W}_{\mathbf{i}} + \rho_{\mathbf{i}+}\mathcal{W}_{\mathbf{i}+}$. Define $h^n_{\mathcal{M}}(\mathbf{i})$ as*

$$h^n_{\mathcal{M}}(\mathbf{i}) := -\frac{1}{n} \log_\mu s_{\mathbf{i}}.$$

The right and left neighbours of $\mathbf{i}$ play a central role in the discretization of the Hölder exponent by avoiding boundary problems which can arise when t is a hitting time. In fact, we have

Proposition 5. *Let p be a regular offspring distribution and suppose Assumptions 1, 2, 3 and 4 hold. Let $t \in [0, W_\emptyset]$ and $\mathbf{i} \in \partial\Upsilon$ be such that $\bigcap_n R_{\mathbf{i}|_n} = \{t\}$. Then, for all $\omega \in \Omega$,*

$$\liminf_{n\to\infty} h^n_{\mathcal{M}}(\mathbf{i}|_n) = h_{\mathcal{M}}(t).$$

Proof. Take $\epsilon > 0$. $(s_{\mathbf{i}|_n})$ is a nested sequence, therefore decreases with n. Thus, it is always possible to find a positive $n = n(\epsilon, t)$ large enough that

$$\log s_{\mathbf{i}|_{n+1}} \leqslant \log[\mathcal{M}(t+\epsilon) - \mathcal{M}(t-\epsilon)] < \log s_{\mathbf{i}|_n}. \tag{4.13}$$

The procedure is now similar to that used in the proof of Theorem 10. Given n and t, let $I_n(t) = \{\epsilon | n(\epsilon, t) = n\}$ and suppose $\epsilon \in I_n(t)$. We need bounds for $I_n(t)$.

Suppose ϵ is larger than $\mu^{-n}W_{\mathbf{i}|n} + \mu^{-n}W_{\mathbf{i}-|n}$ and $\mu^{-n}W_{\mathbf{i}|n} + \mu^{-n}W_{\mathbf{i}+|n}$. Then it follows that

$$|\mathcal{M}(t+\epsilon) - \mathcal{M}(t-\epsilon)| \geqslant s_{\mathbf{i}|_n}$$

which contradicts (4.13). This is illustrated in Figure 4.2. Consequently, $\epsilon < \mu^{-n}\max[W_{\mathbf{i}|n} + W_{\mathbf{i}-|n}, W_{\mathbf{i}|n} + W_{\mathbf{i}+|n}]$ for all $\epsilon \in I_n(t)$.

Similarly, if ϵ is strictly smaller than $\mu^{-(n+1)}W_{\mathbf{i}-|n+1}$ and $\mu^{-(n+1)}W_{\mathbf{i}+|n+1}$ then

$$|\mathcal{M}(t+\epsilon) - \mathcal{M}(t-\epsilon)| < s_{\mathbf{i}|_{n+1}}$$

which also contradicts (4.13). Hence, $\epsilon \geqslant \mu^{-(n+1)}\min[W_{\mathbf{i}-|n+1}, W_{\mathbf{i}+|n+1}]$ for all $\epsilon \in I_n(t)$.

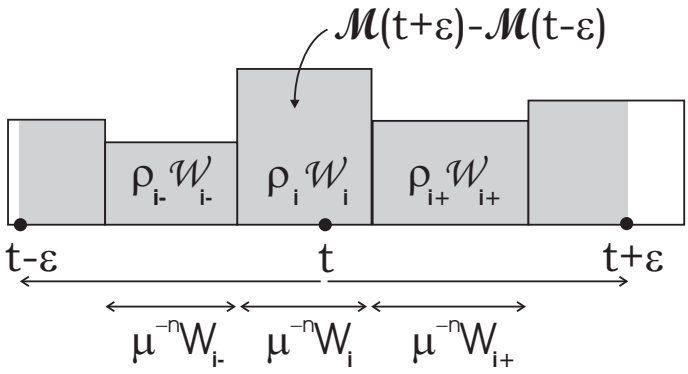


Figure 4.2: Bounds for $I_n(t)$. If ϵ is larger than $\mu^{-n}W_{\mathbf{i}}+\mu^{-n}W_{\mathbf{i}-}$ and $\mu^{-n}W_{\mathbf{i}}+\mu^{-n}W_{\mathbf{i}+}$ for $\mathbf{i} \in \Upsilon_n^{GW}$, then $\mathcal{M}(t+\epsilon)-\mathcal{M}(t-\epsilon)$ colored in grey is necessarily larger than $s_{\mathbf{i}}$, the measure of the interval containing t and its right and left neighbours.

Under Assumption 4, it follows from Lemma 6 that for $\epsilon \in I_n(t)$,

$$-(n+1)\log\mu + \log 2 + o(n) \leqslant \log 2\epsilon < -n\log\mu + \log 4 + o(n).$$

Now dividing each member of the double inequality (4.13) by $\log 2\epsilon < 0$ and using these bounds, we obtain

$$\liminf_{\epsilon\to 0}\left[\frac{\log|\mathcal{M}(t+\epsilon)-\mathcal{M}(t-\epsilon)|}{\log(2\epsilon)}\right] \leqslant \liminf_{n\to+\infty}\left[\frac{\log s_{\mathbf{i}|_{n+1}}}{-n\log\mu + o(n)}\right]$$

and

$$\liminf_{n\to+\infty}\left[\frac{\log s_{\mathbf{i}|_n}}{-(n+1)\log\mu + o(n)}\right] \leqslant \liminf_{\epsilon\to 0}\left[\frac{\log|\mathcal{M}(t+\epsilon)-\mathcal{M}(t-\epsilon)|}{\log(2\epsilon)}\right]$$

which yields

$$\liminf_{n\to+\infty}\left[\frac{n}{n+1}h^n_{\mathcal{M}}(\mathbf{i}|_n)\right] \leqslant h_{\mathcal{M}}(t) \leqslant \liminf_{n\to+\infty}\left[\frac{n+1}{n}h^{n+1}_{\mathcal{M}}(\mathbf{i}|_{n+1})\right]$$

and the result follows. □

4.3.3 An upper bound for the spectrum of $\mathcal{M}$

As explained earlier, we need to introduce new definitions to derive theoretically the spectrum of $\mathcal{M}$. Motivated by its cascade construction, the results and proofs presented here follow the methodology presented by Riedi [102] who gives conditions for the multifractal formalism to hold and presents results for the binary cascade. We introduce definitions of coarse spectrum (section 4.3.3.1) and deterministic coarse spectrum (section 4.3.3.2) adapted to the study of $\mathcal{M}$, which possess strong similarities with the spectra presented in section 4.3.1. The derivation of an upper bound for the Hausdorff spectrum of $\mathcal{M}$ (section 4.3.3.6) is obtained using a large deviations approach, whose theory is briefly summarized in section 4.3.3.3.

4.3.3.1 Coarse spectrum

In this section, we present the definition of a coarse spectrum adapted to $\mathcal{M}$ and show that it provides an upper bound of the fine spectrum $D(a)$.

DEFINITION 12. *Let f be the coarse spectrum of $\mathcal{M}$ defined as*

$$f(a) := \lim_{\epsilon \to 0} \limsup_{n \to +\infty} \frac{1}{n} \log_\mu [\#\{\mathbf{i} \in \Upsilon_n^{GW} \mid a - \epsilon \leqslant h_{\mathcal{M}}^n(\mathbf{i}) \leqslant a + \epsilon\}]. \tag{4.14}$$

When convenient, we write $N^{(n)}(a, \epsilon)$ for $\#\{\mathbf{i} \in \Upsilon_n^{GW} \mid a - \epsilon \leqslant h_{\mathcal{M}}^n(\mathbf{i}) \leqslant a + \epsilon\}$. The main difference between f and $f^\dagger$ (Equation 4.10) comes from the irregular grid used to partition the real line instead of r-mesh cubes. The 2 spectra are a priori different and it is not fully understood whether the global properties of $\mathcal{M}$ are similar when described by f or $f^\dagger$. A way to show equality between the 2 spectra would be to obtain bounds for the cardinality of Υ_n^{GW} and to compare them with the number of μ^{-n}-mesh cubes used in the definition of $f^\dagger$, with $r = \mu^{-n}$. However, we do not pursue this study here.

Also, note that neither f nor $f^\dagger$ can be understood as the box dimension of $\mathcal{M}$, as explained in [45], Chapter 17 for $f^\dagger$, since intervals covering the set of points with regularity a do not form a nested sequence. That is, intervals not counted when n is small can contain an interval counted when n is large, which contradicts the definition of box dimension.

The classical result that $f^\dagger$ is an upper bound of the fine spectrum also holds for f, as stated in the next theorem.

THEOREM 11. *Let p be a regular offspring distribution and suppose Assumptions 1, 2, 3 and 4 hold. Then,*

$$D(a) \leqslant f(a)$$

almost surely with respect to trees.

Proof. Let $a \in \mathbb{R}$ and $\gamma > f(a)$ be arbitrary. We want to show that $f(a)$ is an upper bound for the Hausdorff dimension of $\Theta^a = \{t \mid h(t) = a\}$. To do so, we find a covering of Θ^a and show that its γ-Hausdorff measure is zero for all $\gamma > f(a)$. The result follows by sending γ to $f(a)$.

Let $\eta > 0$ such that $\gamma > f(a) + 2\eta$. By definition of $f(a)$, there exists $\epsilon_0 > 0$ and m_0 such that for all $n \geqslant m_0$ and $0 < \epsilon \leqslant \epsilon_0$,

$$f(a) + \eta \geqslant \frac{1}{n} \log_\mu N^{(n)}(a, \epsilon)$$

that is

$$N^{(n)}(a, \epsilon) \leqslant \mu^{n[f(a)+\eta]}. \tag{4.15}$$

Note that this inequality is true for all $n \geqslant m_0$ since f is defined with a lim sup. Defining f using a lim inf would lead to an inequality true only for infinitely many $n \geqslant m_0$.

We now find a covering for Θ^a. Let $t \in [0, W_\emptyset]$ and let $\mathbf{i} \in \partial\Upsilon$ be such that $\bigcap_n R_{\mathbf{i}|_n} = \{t\}$. If $t \in \Theta^a$, then $\lim_{n \to +\infty} h_{\mathcal{M}}^n(\mathbf{i}|_n) = a$, that is for any $0 < \epsilon \leqslant \epsilon_0$,

there exists $N \in \mathbb{N}$ such that for all $n \geqslant N$, $a - \epsilon \leqslant h^n_{\mathcal{M}}(\mathbf{i}|_n) < a + \epsilon$. Since $t \in [T^{-n}_{\psi(\mathbf{i}|_n)-1}\ T^{-n}_{\psi(\mathbf{i}|_n)}[:= R_{\mathbf{i}|_n}$ for all n, we conclude that for any m

$$\bigcup_{n \geqslant m} \bigcup^{\dagger} R_{\mathbf{i}|_n}$$

is a covering of Θ^a, where $\bigcup^{\dagger}$ is the union taken over $\mathbf{i} \in \partial\Upsilon$ such that $a - \epsilon \leqslant h^n_{\mathcal{M}}(\mathbf{i}|_n) \leqslant a+\epsilon$. Note that the number of nodes $\mathbf{i}$ in generation n of the tree such that $a - \epsilon \leqslant h^n_{\mathcal{M}}(\mathbf{i}|_n) \leqslant a + \epsilon$ is exactly $N^{(n)}(a, \epsilon)$. Also note that $|R_{\mathbf{i}|_n}| \leqslant \max\limits_{\mathbf{j} \in \Upsilon^{GW}_n} \mu^{-n} W_{\mathbf{j}}$. It follows that if $m \geqslant m_0$ and $0 < \epsilon \leqslant \epsilon_0$, then

$$\sum_{n \geqslant m} \sum^{\dagger} |R_{\mathbf{i}|_n}|^{\gamma} \leqslant \sum_{n \geqslant m} |\max_{\mathbf{j} \in \Upsilon^{GW}_n} \mu^{-n} W_{\mathbf{j}}|^{\gamma} \mu^{n[f(a)+\eta]} = \sum_{n \geqslant m} |\max_{\mathbf{j} \in \Upsilon^{GW}_n} W_{\mathbf{j}}|^{\gamma} \mu^{-n[\gamma - f(a) - \eta]}$$

where $\sum^{\dagger}$ is taken over all $\mathbf{i} \in \partial\Upsilon$ such that $a - \epsilon \leqslant h^n_{\mathcal{M}}(\mathbf{i}|_n) \leqslant a + \epsilon$. Since $\gamma > f(a) + 2\eta$, then $\mu^{-n\eta} > \mu^{-n[\gamma - f(a) - \eta]}$ and

$$\sum_{n \geqslant m} \sum^{\dagger} |R_{\mathbf{i}|_n}|^{\gamma} \leqslant \sum_{n \geqslant m} |\max_{\mathbf{j} \in \Upsilon^{GW}_n} W_{\mathbf{j}}|^{\gamma} \mu^{-n\eta}.$$

Under Assumption 4, for any $\epsilon > 0$ there exists n_0 such that for all $n \geqslant n_0$,

$$\max_{\mathbf{j} \in \Upsilon^{GW}_n} W_{\mathbf{j}} \leqslant e^{n\varepsilon} \quad \text{a.s. with respect to trees.}$$

It follows for all $m > \max(m_0, n_0)$:

$$\sum_{n \geqslant m} \sum^{\dagger} |R_{\mathbf{i}|_n}|^{\gamma} \leqslant \sum_{n \geqslant m} e^{n\varepsilon\gamma} \mu^{-n\eta} = \sum_{n \geqslant m} e^{n(\varepsilon\gamma - \eta \log \mu)}.$$

For $\varepsilon < \eta(\log \mu)/\gamma$, the exponent is strictly negative and the sum is finite. Hence, sending m to infinity, we obtain

$$\sum_{n \geqslant m} \sum^{\dagger} |R_{\mathbf{i}}|^{\gamma} \to 0 \quad \text{as } m \to +\infty.$$

The γ-Hausdorff measure of a covering of Θ^a is zero, therefore its Hausdorff dimension is smaller than γ. Since it is true for all $\gamma > f(a)$, the result follows. □

4.3.3.2 Deterministic coarse spectrum

The coarse spectrum f is obtained for one realisation of $\mathcal{M}$. Here, we propose to average the spectrum over the whole sample space, leading to the deterministic coarse spectrum. Recall that $\mathbb{P}$ is the probability measure defined over the space of extended Galton-Watson trees and $\mathbb{E}$ is the expectation under $\mathbb{P}$.

Definition 13. *The deterministic coarse spectrum F of $\mathcal{M}$ is given by*

$$F(a) := \lim_{\epsilon \to 0} \limsup_{n \to +\infty} \frac{1}{n} \log_{\mu} \mathbb{E}[\#\{\mathbf{i} \in \Upsilon^{GW}_n \mid a - \epsilon \leqslant h^n_{\mathcal{M}}(\mathbf{i}) < a + \epsilon\}].$$

Using classical definitions of section 4.3.1, we saw that the deterministic coarse spectrum $F^\dagger$ gives an upper bound for the coarse spectrum $f^\dagger$. This relation also holds with f and F, as stated below

THEOREM 12. *Let p be a regular offspring distribution and suppose Assumptions 1, 2 and 3 hold. Then*

$$f(a) \leqslant F(a)$$

$\mathbb{P}$-almost surely.

Proof. Given $a \in \mathbb{R}$, $\epsilon > 0$ let n_0 be such that for all $n \geqslant n_0$,

$$\mathbb{E}N^{(n)}(a,\epsilon) \leqslant \mu^{n(F(a)+\epsilon)}.$$

Consider $\mathbb{E}\limsup_{n\to\infty} \mu^{-n(F(a)+2\epsilon)} N^{(n)}(a,\epsilon)$. This can be bounded by

$$\mathbb{E}\sum_{n\geqslant n_0} \mu^{-n(F(a)+2\epsilon)} N^{(n)}(a,\epsilon) \leqslant \sum_{n\geqslant n_0} \mu^{-n(F(a)+2\epsilon)} \mu^{n(F(a)+\epsilon)} = \sum_{n\geqslant n_0} \mu^{-n\epsilon} < \infty$$

so that $\limsup_{n\to\infty} \mu^{-n(F(a)+2\epsilon)} N^{(n)}(a,\epsilon)$ is finite $\mathbb{P}$-almost surely. Thus using the definition of $f(a)$, $F(a) + 2\epsilon \geqslant f(a)$. Sending ϵ to zero gives the desired result. □

One could also try to define the deterministic coarse spectrum by averaging $f(a)$, that is we could define $F^*(a) := \mathbb{E}f(a)$. However, there is no obvious relationship between the f and F^*. In particular, using Fatou's lemma and Jensen's inequality,

$$\begin{aligned} F^*(a) = \mathbb{E}\lim_{\epsilon\to 0}\limsup_{n\to+\infty} \frac{1}{n}\log_\mu N^{(n)}(a,\epsilon) &\geqslant \lim_{\epsilon\to 0}\limsup_{n\to+\infty} \frac{1}{n}\mathbb{E}\log_\mu N^{(n)}(a,\epsilon) \\ &\leqslant \lim_{\epsilon\to 0}\limsup_{n\to+\infty} \frac{1}{n}\log_\mu \mathbb{E}N^{(n)}(a,\epsilon) = F(a). \end{aligned}$$

No clear relationship can be derived between F and $F^\dagger$, for the same reasons as for f and $f^\dagger$.

The presence of the double limit in the definition of the coarse spectra is problematic in practice. Usually, we estimate it via the Legendre Fenchel transform of an auxiliary function, as explained in section 4.3.1. We introduce auxiliary functions for the coarse spectra and deterministic coarse spectra in the framework of the theory of large deviations. Before doing so, in the next section we recall the main result of Gärtner and Ellis on large deviations.

4.3.3.3 Large deviations

The theory of large deviation studies the occurrence of rare events. Consider a sequence of random variables $S_1, S_2, \ldots$ converging in probability to some constant c. Then, $P(|S_n - c| > \epsilon) \to 0$ as $n \to \infty$. Often, the convergence is at an exponential rate, hence we can write

$$\lim_{n\to\infty} \frac{1}{n}\log(P(|S_n - c| > \epsilon)) = -I(\epsilon, c)$$

where $I(\cdot,\cdot)$ is a non negative function. When the previous limit exists, we say that the sequence $\{S_n\}$ satisfies a large deviation principle. The theory of large deviations is mainly concerned about determining $I(\cdot,\cdot)$. It can be seen as a generalization of the law of large numbers since it not only says that certain probabilities go to zero, but also find rates of convergence. First results on large deviations are attributed to Cramer who derived an expression for $I(\cdot,\cdot)$ in the case of i.i.d. random variables on the real line [33]. Since then, the theory has been considerably broadened. Gärtner [51] and then Ellis [40, 41] have obtained similar results in a more general context, where S_n can be vectors, not necessarily averages and not necessarily defined over the same probability space. A good introduction to the theory of large deviations can be found in [29].

We will state the Gärtner-Ellis theorem here, then use it in the following sections. We need the following definitions.

DEFINITION 14. *A function $f(x)$ is said to be closed if the set $\{x \in \mathbb{R} \mid f(x) \leqslant b\}$ is closed, for all $b \in \mathbb{R}$.*

DEFINITION 15. *Let f be a convex function on $\mathbb{R}$, whose domain $\mathfrak{D} = \{x \mid f(x) < \infty\}$ has a non-empty interior. Assume also that f is differentiable on the interior of $\mathfrak{D}$. We say that f is steep at x for x a boundary point of $\mathfrak{D}$ if $f'(x_i) \to \infty$ whenever $x_1, x_2, \ldots$ is a sequence of points of the interior of $\mathfrak{D}$ converging to x. f is steep if it is steep at all boundary points.*

THEOREM 13. (Gärtner-Ellis theorem [40, 41])
Let:

- *Y_n be a sequence of random variables defined on probability spaces $(\Omega_n, \mathcal{F}_n, \rho_n)$ for $n \geqslant 1$.*
- *$a_n^{-1} Y_n$ be a sequence of random variables where $a_n \to \infty$ as $n \to \infty$.*
- *$\mathbb{P}_n$ be a sequence of probability measure defined by $\mathbb{P}_n(A) = \rho_n\{\omega \in \Omega_n \mid a_n^{-1} Y_n \in A\}$ for any $A \in \mathcal{B}(\mathbb{R})$. Denote by $\mathbb{E}_n$ the expectation under $\mathbb{P}_n$.*

We assume the following

- *$C(q) = \lim\limits_{n \to +\infty} C_n(q)$ where $C_n(q) = \frac{1}{a_n} \log \mathbb{E}_n e^{q Y_n}$ exists for all $q \in \mathbb{R}$. The value $+\infty$ is allowed both for $C(q)$ and $C_n(q)$. We define $C(q) = +\infty$ if $C_n(q) = +\infty$ for n larger than some $N \in \mathbb{N}$.*
- *$C(q)$ is a closed convex function on $\mathbb{R}$.*
- *The domain $\mathfrak{D} = \{q \in \mathbb{R} \mid C(q) < \infty\}$ of $C(q)$ has a non empty interior containing $q = 0$.*
- *$C(q)$ is differentiable on its domain and is steep.*

Let $I(y) := \sup_{q \in \mathbb{R}}(qy - C(q))$ be the LF transform of $C(q)$. Then, $I(y)$ is convex, closed and non-negative. A Borel subset A of $\mathbb{R}$ is called I-continuity set if

$$\inf_{y \in clA} I(y) = \inf_{y \in intA} I(y)$$

If A is an I-continuity set, then

$$\lim_{n\to+\infty} \frac{1}{a_n} \log \mathbb{P}_n(a_n^{-1} Y_n \in A) = -\inf_{y\in A} I(y).$$

Remark: This theorem is the combination of Theorem II.3.4 in [41] and Theorem II.2 in [40]. Theorem II.3.4 gives the previous result for general $I(y)$, not necessarily defined as the Legendre Fenchel transform of an auxiliary function, under the assumption that I has a large deviation property (see Definition II.3.1. in [41]). Theorem II.2 states that when the conditions of the theorem hold, then I indeed satisfies the large deviation properties, and we can apply Theorem II.3.4 in this context.

4.3.3.4 Partition function τ_h

We use the theorem of Gärtner and Ellis to derive an upper bound of the coarse spectrum f. In this section, we work on one realisation of $\mathcal{M}$. All quantities introduced are therefore deterministic. Let $\mathbf{i} \in \Upsilon_n^{GW}$ and

$$a_n = n \log \mu \qquad \text{and} \qquad Y_n = \log s_{\mathbf{i}}.$$

Then $a_n \to \infty$ as $n \to \infty$. Also,

$$-a_n^{-1} Y_n = -\frac{1}{n} \log_\mu s_{\mathbf{i}} = h^n_{\mathcal{M}}(\mathbf{i}).$$

Let $A_\epsilon = [-a-\epsilon, -a+\epsilon] \to \{-a\}$ as $\epsilon \to 0$ and let $\mathbb{P}_n$ be the uniform distribution on $\{1, \ldots, Z_n\}$. $\mathbb{E}_n$ is the expectation under $\mathbb{P}_n$. Here, the randomness comes by choosing one of the Z_n intervals and then checking whether or not the singularity exponent is in a neighborhood of a.

Assumption 5. *The random variable Y_n and probability measure $\mathbb{P}_n$ satisfy the conditions of the Gärtner-Ellis theorem.*

Under Assumption 5 one have

$$-C_n(q) = -\frac{1}{n \log \mu} \log \mathbb{E}_n \exp[q \log s_{\mathbf{i}}] \to \tau_h(q) = -C(q) \tag{4.16}$$

for all $q \in \mathbb{R}$ with $-C(q)$ a closed, concave, differentiable on its domain and steep function. Thus we can apply Theorem 13 to get

$$\frac{1}{n \log \mu} \log \mathbb{P}_n[a_n^{-1} Y_n \in A_\epsilon] \to -\inf_{z\in A_\epsilon} \sup_{q\in\mathbb{R}} [qz + \tau_h(q)].$$

Then,

$$\begin{aligned} \lim_{\epsilon\to 0} \lim_{n\to+\infty} \frac{1}{n} \log_\mu \mathbb{P}_n[h^n_{\mathcal{M}}(\mathbf{i}) \in -A_\epsilon] &= -\lim_{\epsilon\to 0} \inf_{z\in -A_\epsilon} \sup_{q\in\mathbb{R}} [-qz + \tau_h(q)] \\ &= -\sup_{q\in\mathbb{R}} [-qa + \tau_h(q)] \\ &= \inf_{q\in\mathbb{R}} [qa - \tau_h(q)]. \end{aligned} \tag{4.17}$$

The probability $\mathbb{P}_n[h^n(\mathbf{i}) \in -A_\epsilon]$ can be re-expressed as its relative frequency in the limit as $n \to \infty$. Thus,

$$\begin{aligned}\lim_{\epsilon\to 0}\lim_{n\to+\infty}\frac{1}{a_n}\log \mathbb{P}_n[h^n_{\mathcal{M}}(\mathbf{i}) \in -A_\epsilon] &= \lim_{\epsilon\to 0}\lim_{n\to+\infty}\frac{1}{n}\log_\mu \frac{\#[h^n_{\mathcal{M}}(\mathbf{i}) \in -A_\epsilon]}{Z_n} \\ &= f(a) - \limsup_{n\to+\infty}\frac{1}{n}\log_\mu Z_n. \end{aligned} \tag{4.18}$$

Since $\mu^{-n}Z_n$ converges almost surely to a non-degenerate random variable $W_\emptyset$, $1/n \log_\mu Z_n \to 1$ as $n \to \infty$ when $\mathbb{E}Z_1 \log Z_1 < \infty$ [73], re-expressing (4.16) and combining results from Equations (4.17) and (4.18), we obtain

THEOREM 14. *Let p be a regular offspring distribution and suppose Assumptions 1, 2, 3 and 5 hold. Then, as $n \to \infty$, the limit*

$$-\frac{1}{n}\log_\mu\Big[\frac{1}{Z_n}\sum_{\mathbf{i}\in\Upsilon_n^{GW}}|s_\mathbf{i}|^q\Big] \to \tau_h(q) \tag{4.19}$$

exists for all $q \in \mathbb{R}$ with τ_h a closed, concave, differentiable on the interior of its domain and steep function. Let $\tilde{\tau}_h(q) = \tau_h(q) - 1$ and let $\tilde{\tau}_h^*(a)$ be its LF transform

$$\tilde{\tau}_h^*(a) := \inf_{q\in\mathbb{R}}[qa - \tilde{\tau}_h(q)].$$

Then,

$$f(a) = \inf_{q\in\mathbb{R}}[1 + qa - \tau_h(q)] = \tau_h^*(a)$$

almost surely on trees.

We remark that when $\tau_h(q) < \infty$ for all $q \in \mathbb{R}$, its domain has no boundary point and we do not need to check the steep property of τ.

Assumption 5 is hard to check. As f only gives an upper bound on the Hausdorff spectrum D, it is sufficient to have an upper bound on f, which is what the next theorem gives.

THEOREM 15. *Let p be a regular offspring distribution and suppose Assumptions 1, 2 and 3 hold. Define*

$$\begin{aligned}\tau_h(q) &:= \liminf_{n\to+\infty} -\frac{1}{n}\log_\mu\Big[\frac{1}{Z_n}\sum_{\mathbf{i}\in\Upsilon_n^{GW}}|s_\mathbf{i}|^q\Big] \\ &= \liminf_{n\to+\infty} -\frac{1}{n}\log_\mu\Big[\frac{1}{Z_n}\sum_{\mathbf{i}\in\Upsilon_n^{GW}}\mu^{-nqh^n_{\mathcal{M}}(\mathbf{i})}\Big]\end{aligned} \tag{4.20}$$

and $\tilde{\tau}_h(q) = \tau_h(q) - 1$. Then, for any $a \in \mathbb{R}$,

$$f(a) \leqslant \tilde{\tau}_h^*(a) := \inf_{q\in\mathbb{R}}[qa - \tilde{\tau}_h(q)]$$

almost surely on trees

Remark. We have used $\tau_h(q)$ for the limit and lim inf, but we will understand it to be the lim inf in what follows.

This theorem provides a way to estimate an upper bound for the coarse spectrum of $\mathcal{M}$, based on its increments. The method is similar to classical estimators of the partition function using a wavelet decomposition on a Haar basis. The difference here is that we use an irregular grid to derive increments. The grid is provided by the crossing tree and is therefore adapted to the signal.

Proof. Let $a \in \mathbb{R}$ be such that $f(a) > -\infty$. Let $\gamma < f(a)$, then there exists $\epsilon_0 > 0$ and m_0 such that for all $0 < \epsilon \leqslant \epsilon_0$ and $n \geqslant m_0$,

$$\gamma \leqslant \frac{1}{n} \log_\mu N^{(n)}(a, \epsilon)$$

or equivalently $N^{(n)}(a, \epsilon) \geqslant \mu^{\gamma n}$. The sum $\sum_{\mathbf{i} \in \Upsilon_n^{GW}} \mu^{-nqh^n_{\mathcal{M}}(\mathbf{i})}$ can be bounded below by considering only singularity coefficients which are roughly equal to a:

$$\sum_{\mathbf{i} \in \Upsilon_n^{GW}} \mu^{-nqh^n_{\mathcal{M}}(\mathbf{i})} \geqslant \sum_{\substack{\mathbf{i} \in \Upsilon_n^{GW} \\ |h^n_{\mathcal{M}}(\mathbf{i}) - a| \leqslant \epsilon}} \mu^{-nqh^n_{\mathcal{M}}(\mathbf{i})}.$$

There are exactly $N^{(n)}(a, \epsilon) \geqslant \mu^{\gamma n}$ of those coefficients. Furthermore, for such a coefficient $h^n_{\mathcal{M}}(\mathbf{i})$, we have

$$\begin{aligned} &\text{For } q \geqslant 0, \quad -nqh^n_{\mathcal{M}}(\mathbf{i}) \geqslant -nqa - nq\epsilon \\ &\text{For } q < 0, \quad -nqh^n_{\mathcal{M}}(\mathbf{i}) \geqslant -nqa + nq\epsilon. \end{aligned}$$

So that $\mu^{-nqh^n_{\mathcal{M}}(\mathbf{i})} \geqslant \mu^{-n(qa+|q|\epsilon)}$ for all q. It follows that

$$\sum_{\mathbf{i} \in \Upsilon_n^{GW}} \mu^{-nqh^n_{\mathcal{M}}(\mathbf{i})} \geqslant N^{(n)}(a, \epsilon) \mu^{-n(qa+|q|\epsilon)} \geqslant \mu^{-n(qa-\gamma+|q|\epsilon)}.$$

From the definition of τ_h,

$$\begin{aligned} \tau_h(q) &= \liminf_{n \to +\infty} -\frac{1}{n} \log_\mu \frac{1}{Z_n} \sum_{\mathbf{i} \in \Upsilon_n^{GW}} \mu^{-nqh^n_{\mathcal{M}}(\mathbf{i})} \\ &= 1 + \liminf_{n \to +\infty} -\frac{1}{n} \log_\mu \sum_{\mathbf{i} \in \Upsilon_n^{GW}} \mu^{-nqh^n_{\mathcal{M}}(\mathbf{i})} \\ &\leqslant 1 + qa - \gamma + |q|\epsilon. \end{aligned}$$

Now let $\epsilon \to 0$ and $\gamma \to f(a)$. Then $\tau_h(q) \leqslant 1 + qa - f(a)$, or $f(a) \leqslant qa - \tilde{\tau}_h(q)$. This inequality is obviously true if $f(a) = -\infty$. Since this inequality holds for all $(a, q) \in \mathbb{R}^2$,the result follows. □

4.3.3.5 Deterministic partition function T_h

We introduce a deterministic partition function T_h by averaging across all sample paths. τ_h and f depend on one realisation of the chronometer $\mathcal{M}$ whereas T_h and F (the deterministic coarse spectrum introduced in section 4.3.3.2) are path independent. Recall that Ω is the sample space of extended Galton-Watson trees.

To define τ, we defined a probability distribution $\mathbb{P}_n$ to be the uniform distribution on level n of a single realisation of the tree and applied the Gärtner-Ellis theorem. Here, we introduce a new probability $\mathcal{P}$ on Ω to perform averages within and across all sample paths.

Let $X : \Omega \to \mathbb{R}$. We are particularly interested in the case when X is the weight assigned to a particular node of a tree, that is when $X(\omega) = \rho_{\mathbf{i}}\mathcal{W}_{\mathbf{i}}$ for $\omega \in \Omega$. We define $\mathcal{P}$ as follows:

$$\mathcal{P}(\rho_{\mathbf{i}}\mathcal{W}_{\mathbf{i}} > x) := \mu^{-n}\mathbb{E}\#\{\mathbf{i} \in \Upsilon_n^{GW} \mid \rho_{\mathbf{i}}\mathcal{W}_{\mathbf{i}} > x\}.$$

Denote by $\mathcal{E}$ the expectation under $\mathcal{P}$.

$$\mathcal{E}[\rho_{\mathbf{i}}\mathcal{W}_{\mathbf{i}}] = \int_{\mathbb{R}_+} \mathcal{P}(\rho_{\mathbf{i}}\mathcal{W}_{\mathbf{i}} > x)dx = \mu^{-n}\int_{\mathbb{R}_+} \mathbb{E}\#\{\mathbf{i} \in \Upsilon_n^{GW} \mid \rho_{\mathbf{i}}\mathcal{W}_{\mathbf{i}} > x\}dx.$$

We can rewrite the expression inside the expectation as a sum of indicator functions. Also, since all terms are non-negative, we can swap sum and integral:

$$\mathcal{E}[\rho_{\mathbf{i}}\mathcal{W}_{\mathbf{i}}] = \mu^{-n}\mathbb{E}\int_{\mathbb{R}_+}\sum_{\mathbf{i}\in\Upsilon_n^{GW}} \mathbf{1}_{\{\rho_{\mathbf{i}}\mathcal{W}_{\mathbf{i}}>x\}}dx = \mu^{-n}\mathbb{E}\sum_{\mathbf{i}\in\Upsilon_n^{GW}}\int_0^{\rho_{\mathbf{i}}\mathcal{W}_{\mathbf{i}}} dx = \mu^{-n}\mathbb{E}\sum_{\mathbf{i}\in\Upsilon_n^{GW}}\rho_{\mathbf{i}}\mathcal{W}_{\mathbf{i}}.$$

We prove the following in the same way we proved Theorem 14 and 15. We keep the same definitions for a_n and Y_n and we use the probability measure $\mathcal{P}$.

Assumption 6. *The random variable Y_n and probability measure $\mathcal{P}$ satisfy the conditions of the Gärtner-Ellis theorem.*

Under Assumption 6 one have

$$-\frac{1}{n\log\mu}\log\mathcal{E}\exp[q\log s_{\mathbf{i}}] = -\frac{1}{n}\log_\mu\left[\mu^{-n}\mathbb{E}\sum_{\mathbf{i}\in\Upsilon_n^{GW}}|s_{\mathbf{i}}|^q\right] \to T_h(q)$$

for all $q \in \mathbb{R}$ with $-T_h(q)$ a closed, concave, differentiable on its domain and steep function. Let $A_\epsilon = [a-\epsilon, a+\epsilon]$ be defined as in section 4.3.3.4. Theorem 13 states that

$$\frac{1}{n}\log_\mu \mathcal{P}[a_n^{-1}Y_n \in A_\epsilon] \to -\inf_{z\in A_\epsilon}\sup_{q\in\mathbb{R}}[qz + T_h(q)].$$

Sending $n \to \infty$ and $\epsilon \to 0$, the left hand side in (4.3.3.5) becomes:

$$\lim_{\epsilon\to 0}\lim_{n\to+\infty}\frac{1}{a_n}\log\mathcal{P}[h^n(\mathbf{i}) \in -A_\epsilon] = F(a) - 1$$

and the right hand side is given by

$$\inf_{q\in\mathbb{R}}[qa - T_h(q)].$$

Let $\tilde{T}_h(q) = T_h(q) - 1$ and $\tilde{T}_h^*(a) = \inf_{q\in\mathbb{R}}[qa - \tilde{T}_h(q)]$ its LF transform. We summarize the result of applying the Gärtner-Ellis theorem to $\mathcal{P}$ in the following theorem

THEOREM 16. *Let p be a regular offspring distribution and suppose Assumptions 1, 2, 3 and 6 hold. Then, as $n \to \infty$,*

$$-\frac{1}{n}\log_\mu\left[\mu^{-n}\mathbb{E}\sum_{\mathbf{i}\in\Upsilon_n^{GW}}|s_{\mathbf{i}}|^q\right] \to T_h(q) \tag{4.21}$$

exists for all $q \in \mathbb{R}$ with T_h a closed, concave, differentiable on the interior of its domain and steep function. Let $\tilde{T}_h(q) = T_h(q) - 1$ and let $\tilde{T}_h^*(a)$ be its LF transform

$$\tilde{T}_h^*(a) := \inf_{q\in\mathbb{R}}[qa - \tilde{T}_h(q)].$$

Then, for all $a \in \mathbb{R}$,

$$F(a) = \inf_{q\in\mathbb{R}}[qa - \tilde{T}_h(q)] = \tilde{T}_h^*(a).$$

When one cannot check Assumption 6, we can obtain an upper bound for F, the same way Theorem 15 provides an upper bound for f.

THEOREM 17. *Let p be a regular offspring distribution and suppose assumptions 1, 2 and 3 hold. Define*

$$T_h(q) := \liminf_{n\to+\infty} -\frac{1}{n}\log_\mu\left[\mu^{-n}\mathbb{E}\sum_{\mathbf{i}\in\Upsilon_n^{GW}}|s_{\mathbf{i}}|^q\right] \tag{4.22}$$

and $\tilde{T}_h(q) = T_h(q) - 1$. Then, for any $a \in \mathbb{R}$,

$$F(a) \leqslant \tilde{T}_h^*(a) := \inf_{q\in\mathbb{R}}[qa - \tilde{T}_h(q)].$$

Remark. We have used $T_h(q)$ for the limit and lim inf, but we will understand it to be the lim inf in what follows.

We now investigate the relationship between $\tilde{\tau}_h^*(a)$ and $\tilde{T}_h^*(a)$. For convenience, we introduce the following notation

$$\begin{aligned} S_n^h(q) &= \frac{1}{Z_n}\sum_{\mathbf{i}\in\Upsilon_n^{GW}}|s_{\mathbf{i}}|^q \\ \mathcal{S}_n^h(q) &= \mu^{-n}\sum_{\mathbf{i}\in\Upsilon_n^{GW}}|s_{\mathbf{i}}|^q. \end{aligned}$$

Equations (4.20) and (4.22) can be rewritten

$$\begin{aligned} \tau_h(q) &:= \liminf_{n\to+\infty} -\frac{1}{n}\log_\mu S_n^h(q) \\ T_h(q) &:= \liminf_{n\to+\infty} -\frac{1}{n}\log_\mu \mathbb{E}\mathcal{S}_n^h(q) \end{aligned}$$

or equivalently

$$\begin{aligned}\tilde{\tau}_h(q) &:= \liminf_{n\to+\infty} -\frac{1}{n}\log_\mu Z_n S_n^h(q)\\ \tilde{T}_h(q) &:= \liminf_{n\to+\infty} -\frac{1}{n}\log_\mu \mathbb{E}\mu^n \mathcal{S}_n^h(q).\end{aligned}$$

The following result can be derived, following the steps of Riedi:

THEOREM 18. *Let p be a regular distribution and suppose Assumptions 1, 2 and 3 hold. Then, for all $a \in \mathbb{R}$,*

$$\tau_h^*(a) \leqslant T_h^*(a)$$

almost surely with respect to trees.

Proof. Let $\epsilon > 0$. Then $T_h(q) - \epsilon < \liminf -\frac{1}{n}\log_\mu \mathbb{E}\mathcal{S}_n^h(q)$ and there exists a positive integer m_0 such that $\forall n \geqslant m_0$,

$$\mathbb{E}\mathcal{S}_n^h(q) \leqslant \mu^{-n(T_h(q)-\epsilon)}.$$

Now consider $\mathbb{E}[\limsup \mu^{n(T_h(q)-2\epsilon)}\mathcal{S}_n^h(q)]$. This quantity is bounded above by

$$\mathbb{E}\sum_{n\geqslant m_0} \mu^{n(T_h(q)-2\epsilon)}\mathcal{S}_n^h(q) = \sum_{n\geqslant m_0} \mu^{n(T_h(q)-2\epsilon)}\mathbb{E}\mathcal{S}_n^h(q) \leqslant \sum_{n\geqslant m_0} \mu^{-n\epsilon} < \infty.$$

It follows that $\limsup \mu^{n(T_h(q)-2\epsilon)} Z_n \mu^{-n} S_n^h(q) < \infty$ almost surely with respect to trees, or $\limsup \mu^{n(T_h(q)-2\epsilon-\tau(q))}\mu^{-n}Z_n$ is finite with probability 1. Since $\mu^{-n}Z_n$ tends almost surely to some non degenerate random variable $W_\emptyset < \infty$ under the assumption that p is regular, the last inequality holds if

$$T_h(q) - 2\epsilon \leqslant \tau_h(q).$$

Sending $\epsilon \to 0$, we obtain with probability 1, $\tau_h(q) \geqslant T_h(q)$ for all q such that $T_h(q) < \infty$. The result now follows from taking the LF transform of $\tau_h(q)$ and $T_h(q)$. □

We summarize the relationships among the various spectra we have introduced so far in the following theorem. The set of inequalities is similar to the previous ones (4.11).

THEOREM 19. *Given Assumptions 1, 2, 3 and 4 hold, we have almost surely with respect to trees, for all $a \in \mathbb{R}$,*

$$D(a) \leqslant f(a) \leqslant \tilde{\tau}_h^*(a) \leqslant \tilde{T}_h^*(a) \tag{4.23}$$

where D, f and $\tilde{\tau}_h$ are defined pathwise. Also,

$$D(a) \leqslant f(a) \leqslant F(a) \leqslant \tilde{T}_h^*(a). \tag{4.24}$$

4.3.3.6 Deterministic coarse spectrum of $\mathcal{M}$

In this section we derive an expression for an upper bound of the Hausdorff spectrum in terms of the cascade parameters using the Theorem 19. The first step is to introduce another discrete local exponent $\gamma^n_{\mathcal{M}}(\mathbf{i})$, closely related to $h^n_{\mathcal{M}}(\mathbf{i})$ introduced in Section 4.3.2. Next, we show that using $h^n_{\mathcal{M}}(\mathbf{i})$ or $\gamma^n_{\mathcal{M}}(\mathbf{i})$ does not change the definition of $\tau_h(q)$ and $T_h(q)$ (Proposition 6) and then express $T_h(q)$ as a function of the offspring distribution parameters and the weight moments.

DEFINITION 16. *Let* $\mathbf{i} \in \Upsilon^{GW}_n$. *Define* $\gamma^n_{\mathcal{M}}(\mathbf{i})$ *as*

$$\gamma^n_{\mathcal{M}}(\mathbf{i}) := -\frac{1}{n}\log_\mu \rho_{\mathbf{i}}\mathcal{W}_{\mathbf{i}}.$$

We introduce further notation

$$\begin{aligned} S^\gamma_n(q) &= \frac{1}{Z_n}\sum_{\mathbf{i}\in\Upsilon^{GW}_n} |\rho_{\mathbf{i}}\mathcal{W}_{\mathbf{i}}|^q \\ \mathcal{S}^\gamma_n(q) &= \mu^{-n}\sum_{\mathbf{i}\in\Upsilon^{GW}_n} |\rho_{\mathbf{i}}\mathcal{W}_{\mathbf{i}}|^q. \end{aligned}$$

Define $\tau_\gamma(q)$ and $T_\gamma(q)$ as

$$\begin{aligned} \tau_\gamma(q) &:= \liminf_{n\to+\infty} -\frac{1}{n}\log_\mu S^\gamma_n(q) \\ T_\gamma(q) &:= \liminf_{n\to+\infty} -\frac{1}{n}\log_\mu \mathbb{E}\mathcal{S}^\gamma_n(q). \end{aligned} \tag{4.25}$$

Also put $\tilde{\tau}_\gamma(q) = \tau_\gamma(q) - 1$ and $\tilde{T}_\gamma(q) = T_\gamma(q) - 1$. Using the same arguments, it is possible to obtain the same set of inequalities derived in Theorem 19 using the exponents $\gamma^n_{\mathcal{M}}(\mathbf{i})$. However, $\gamma^n_{\mathcal{M}}(\mathbf{i})$ does not tend a priori to the local Hölder exponent $h_{\mathcal{M}}(t)$ of $\mathcal{M}(t)$, so $\tilde{\tau}_\gamma(q)$ and $\tilde{T}_\gamma(q)$ are not necessarily the same as in Theorem 19. The following proposition gives conditions such that $\tau_h(q) = \tau_\gamma(q)$ and $T_h(q) = T_\gamma(q)$. Its proof follows the same lines as Lemma 5.5 in [102].

PROPOSITION 6. *Let p be a regular distribution. Assume the crossing tree has bounded family size and that Assumptions 1, 2 and 3 hold. Then $\gamma^n(\mathbf{i})$ and $h^n(\mathbf{i})$ lead to the same partition function and deterministic partition function. That is, for all $q \in \mathbb{R}$,*

$$\tau_\gamma(q) = \tau_h(q) \qquad T_\gamma(q) = T_h(q). \tag{4.26}$$

The first equality holds for all q almost surely with respect to trees.

Proof. **First step.** $\mathcal{M}$ is a non decreasing process, therefore $\rho_{\mathbf{i}}\mathcal{W}_{\mathbf{i}} \leqslant s_{\mathbf{i}} = \rho_{\mathbf{i}-}\mathcal{W}_{\mathbf{i}-} + \rho_{\mathbf{i}}\mathcal{W}_{\mathbf{i}} + \rho_{\mathbf{i}+}\mathcal{W}_{\mathbf{i}+}$ or equivalently $h^n_{\mathcal{M}}(\mathbf{i}) \leqslant \gamma^n_{\mathcal{M}}(\mathbf{i})$. Consider first the case $q \geqslant 0$. Then $\mu^{-nqh^n_{\mathcal{M}}(\mathbf{i})} \geqslant \mu^{-nq\gamma^n_{\mathcal{M}}(\mathbf{i})}$. By summing over all intervals at generation n, we obtain $Z_n S^h_n(q) \geqslant Z_n S^\gamma_n(q)$. It follows that for positive q, $\tau_h(q) \leqslant \tau_\gamma(q)$ and similarly for the deterministic envelopes $T_h(q) \leqslant T_\gamma(q)$. It is straightforward to check that the reverse inequalities hold for negative values of q. In summary, we have:

$$\begin{aligned} &\text{For } q \geqslant 0, \quad \tau_h(q) \leqslant \tau_\gamma(q) \quad T_h(q) \leqslant T_\gamma(q) \\ &\text{For } q < 0, \quad \tau_h(q) \geqslant \tau_\gamma(q) \quad T_h(q) \geqslant T_\gamma(q). \end{aligned}$$

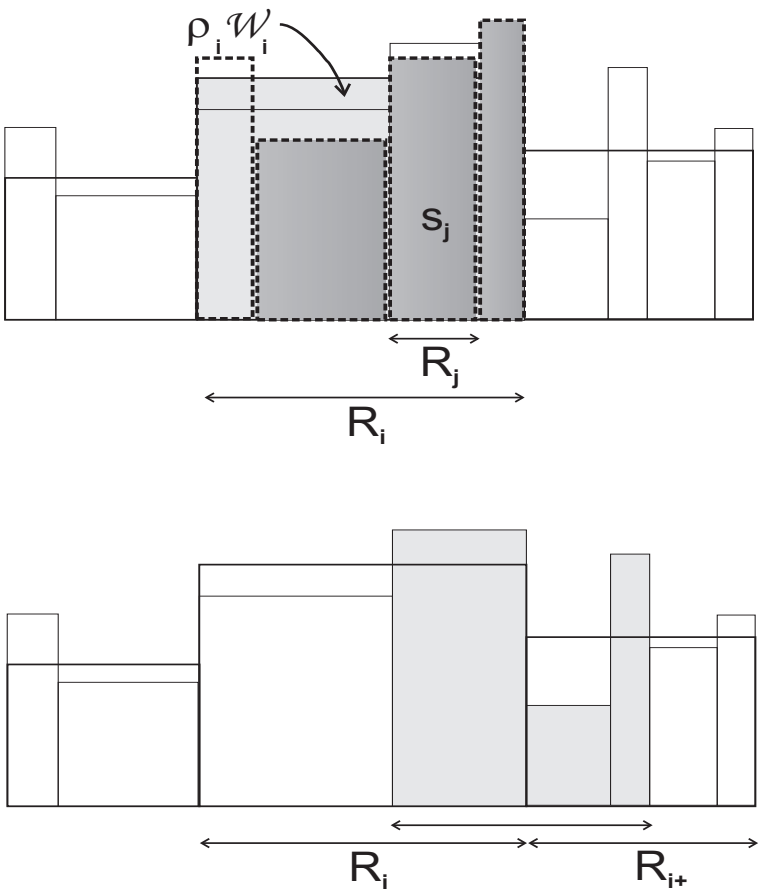


Figure 4.3: Top figure: Each interval $R_{\mathbf{i}}$ of generation n is decomposed into at least two intervals in generation $n+1$, which are also subdivided into at least two intervals $R_{\mathbf{j}}$ in generation $n+2$. By conservation of mass, $\rho_{\mathbf{i}}\mathcal{W}_{\mathbf{i}}$, in light grey on the figure, is split here into four intervals at generation $n+2$. The measure $s_{\mathbf{j}}$ (dark grey) is therefore smaller than $\rho_{\mathbf{i}}\mathcal{W}_{\mathbf{i}}$. This geometric argument illustrates why it is always possible to find $\mathbf{j} \in \Upsilon_{n+2}$ such that $R_{\mathbf{j}} \subset R_{\mathbf{i}}$ for $\mathbf{i} \in \Upsilon_n^{GW}$ and such that $s_{\mathbf{j}} \leqslant \rho_{\mathbf{i}}\mathcal{W}_{\mathbf{i}}$. Bottom figure: each interval of the form $[T_{k-2}^{-(n+1)}\ T_{k+1}^{-(n+1)})$ (light grey) can only intersect two intervals $R_{\mathbf{i}}$ at generation n, since each interval of generation n is divided in at least two intervals at generation $n+1$.

Second step. Let $q < 0$. The crossing tree has at least two children per node. Also, for a particular realisation, the intervals $R_{\mathbf{i}}$ form a nested sequence. For a given $\mathbf{i} \in \Upsilon_n^{GW}$, this ensures the existence of a node $\mathbf{j} \in \Upsilon_{n+2}$ such that

$$R_{\mathbf{j}} \subset R_{\mathbf{i}} \text{ and } s_{\mathbf{j}} \leqslant \rho_{\mathbf{i}} \mathcal{W}_{\mathbf{i}}. \tag{4.27}$$

It follows by the definition of $S_n^{\gamma}(q)$ and $S_{n+1}^{h}(q)$ that:

$$Z_n S_n^{\gamma}(q) = \sum_{\mathbf{i} \in \Upsilon_n^{GW}} |\rho_{\mathbf{i}} \mathcal{W}_{\mathbf{i}}|^q < \sum_{\mathbf{j} \in \Upsilon_{n+2}} |s_{\mathbf{j}}|^q = Z_{n+2} S_{n+2}^{h}(q)$$

since $q < 0$. The inequality is strict since there are only Z_n of $\mathbf{j} \in \Upsilon_{n+2}$ for which (4.27) holds while $S_{n+2}^{h}(q)$ is defined over the whole generation $n+2$ of size $Z_{n+2} > Z_n$. By taking the $\log_{\mu}$ and sending n to infinity, it follows that

$$\text{For } q < 0, \quad \tau_h(q) \leqslant \tau_{\gamma}(q) \quad T_h(q) \leqslant T_{\gamma}(q).$$

Let $q \geqslant 0$. Each interval of the form $[T_{k-2}^{-(n+1)}\ T_{k+1}^{-(n+1)})$ intersects at most two $R_{\mathbf{i}}$ at generation n. Let $s(k)$ denote the largest increment $\rho_{\mathbf{i}} \mathcal{W}_{\mathbf{i}}$ over the two. Let $\mathbf{i} \in \Upsilon_{n+1}$ such that $\psi(\mathbf{i}) = k$ (recall that ψ gives the position of a node within its generation) and $\mathbf{j} \in \Upsilon_n^{GW}$ such that $\psi(\mathbf{j}) = s(k)$. Then:

$$s_{\mathbf{i}} \leqslant 2 \rho_{\mathbf{j}} \mathcal{W}_{\mathbf{j}}.$$

We assume a bounded maximum number of children M_1. Then each interval $R_{\mathbf{j}}$ of generation n will only intersect a finite number of intervals $s_{\mathbf{i}}$ of generation $n+1$. Denote by M_2 the largest such number. M_1 and M_2 are related to each other by $M_2 = M_1 + 2$. It follows:

$$Z_{n+1} S_{n+1}^{h}(q) := \sum_{\mathbf{i} \in \Upsilon_{n+1}} |s_{\mathbf{i}}|^q \leqslant 2^q M_2 \sum_{\mathbf{j} \in \Upsilon_n^{GW}} |\rho_{\mathbf{j}} \mathcal{W}_{\mathbf{j}}|^q = 2^q M_2 Z_n S_n^{\gamma}(q).$$

The extra multiplicative factor $2^q M_2$ disappears when taking logs, dividing by n and sending n to infinity. Thus,

$$\text{For } q \geqslant 0, \quad \tau_h(q) \geqslant \tau_{\gamma}(q) \quad T_h(q) \geqslant T_{\gamma}(q).$$

□

LEMMA 7. *Let p be a regular distribution. Assume i.i.d. weights on the crossing tree distributed like ρ, bounded family size and that Assumptions 1, 2 and 3 hold. The deterministic partition function of $\mathcal{M}$ reduces to*

$$T_h(q) = -\log_{\mu} \mathbb{E}[\rho^q].$$

Proof. By definition of $T_{\gamma}(q)$,

$$\mu^{-n} \mathbb{E} \sum_{\mathbf{i} \in \Upsilon_n^{GW}} |\rho_{\mathbf{i}} \mathcal{W}_{\mathbf{i}}|^q = \mu^{-n} \mathbb{E} \mathbb{E} \left[\sum_{\mathbf{i} \in \Upsilon_n^{GW}} |\rho_{\mathbf{i}} \mathcal{W}_{\mathbf{i}}|^q \mid Z_n \right]$$

where $\rho_{\mathbf{i}}\mathcal{W}_{\mathbf{i}}$ are independent and identically distributed. Since $\mathcal{W}_{\mathbf{i}}$ is distributed like $\mathcal{W}_\emptyset$ and $\rho_{\mathbf{i}} = \prod_{k=1}^{n} \rho_{i_k}(\mathbf{i}|_{k-1})$, where $\rho_{i_1}(\emptyset), \ldots, \rho_{i_n}(\mathbf{i}|_{n-1})$ are i.i.d. random variables, we have

$$\mu^{-n}\mathbb{E} \sum_{\mathbf{i}\in\Upsilon_n^{GW}} |\rho_{\mathbf{i}}\mathcal{W}_{\mathbf{i}}|^q = \mu^{-n}\mathbb{E}[Z_n\mathbb{E}|\rho_1 \times \ldots \times \rho_n\mathcal{W}_\emptyset|^q]$$

where $\rho_1 \times \ldots \times \rho_n$ is a product of n i.i.d. random variables distributed like, say, ρ. Thus:

$$\mu^{-n}\mathbb{E} \sum_{\mathbf{i}\in\Upsilon_n^{GW}} |\rho_{\mathbf{i}}\mathcal{W}_{\mathbf{i}}|^q = \mu^{-n}(\mathbb{E}Z_n)(\mathbb{E}\rho^q)^n(\mathbb{E}\mathcal{W}_\emptyset^q) = (\mathbb{E}\rho^q)^n(\mathbb{E}\mathcal{W}_\emptyset^q) \tag{4.28}$$

where $\mathbb{E}\mathcal{W}_\emptyset^q < \infty$ under Assumption 3. The deterministic envelope T_γ can be computed by taking logs, dividing by n and then sending n to infinity:

$$\lim_{n\to+\infty} -\frac{1}{n}\log_\mu \left[\mu^{-n}\mathbb{E} \sum_{\mathbf{i}\in\Upsilon_n^{GW}} |\rho_{\mathbf{i}}\mathcal{W}_{\mathbf{i}}|^q\right] = -\log_\mu \mathbb{E}\rho^q.$$

Assuming bounded family size, the result holds for T_h from Proposition 6. □

Corollary 6. *Let p be a regular offspring distribution. Assume bounded family size, i.i.d. weights and Assumptions 1, 2, 3 and 4 hold. Let $\tilde{T}_h(q) = -1 - \log_\mu \mathbb{E}\rho^q$. An upper bound for the Hausdorff spectrum of $\mathcal{M}$ is given by*

$$D_{\mathcal{H}}(\Theta^a) \leqslant \inf_{q\in\mathbb{R}}(qa - \tilde{T}_h(q)). \tag{4.29}$$

Proof. It follows from Theorem 19, Proposition 6 and Lemma 7. □

Remarks.

- If we consider constant weights, $\rho_{\mathbf{i}}\mathcal{W}_{\mathbf{i}}$ become $\mu^{-n}W_{\mathbf{i}}$. The time change is then linear (hence monofractal) and $T_h(q) = q$.
- By definition, when $\mathbb{E}Z_1 \log Z_1 < \infty$, $\mu^{-n}Z_n$ converges to a non-degenerate limit and $T_\gamma(0) = 0$. Furthermore, since $\mathcal{M}$ is a non decreasing process, it has positive increments and

  $$Z_n S_n^\gamma(1) = \sum_{\mathbf{i}\in\Upsilon_n^{GW}} \rho_{\mathbf{i}}\mathcal{W}_{\mathbf{i}} = \sum_{\mathbf{i}\in\Upsilon_n^{GW}} \rho_{\mathbf{i}} \left(\sum_{j=1}^{Z_{\mathbf{i}}} \rho_j(\mathbf{i})\mathcal{W}_{\mathbf{i}j}\right) = \sum_{\mathbf{i}\in\Upsilon_{n+1}^{GW}} \rho_{\mathbf{i}}\mathcal{W}_{\mathbf{i}} = Z_{n+1}S_{n+1}^\gamma(1).$$

  Thus $-(1/n)\log_\mu Z_n S_n^\gamma(1) \to 0$ as n tends to infinity, which implies $T_\gamma(1) = 1$. Similarly, $\tilde{T}_\gamma(0) = -1$ and $\tilde{T}_\gamma(1) = 0$ and the results also hold for T_h and $\tilde{T}_h$.
- $S_n^{1,h}(0)$ counts the number of finite $h_{\mathcal{M}}^n(\mathbf{i})$. Thus for all n, $(1/n)\log_\mu \mathbb{E}N^{(n)}(a,\epsilon) \leqslant (1/n)\log_\mu \mathbb{E}S_n^{1,h}(0)$. If all the $h_{\mathcal{M}}^n(\mathbf{i})$ are finite, sending n to infinity yields

  $$F(a) \leqslant -\tilde{T}_h(0) = 1. \tag{4.30}$$

Under conditions of Theorem 16, we have equality between F and $\widetilde{T}_h$ (the LF transform of T_h), and T_h is differentiable. Therefore, there exists a^* such that $F(a^*) = 1$. The associated value of q, denoted $q(a^*)$ is 0. Thus,

$$F(a^*) = 1 = \inf_q(qa^* - \widetilde{T}_h(q)) = qa^* - \widetilde{T}_h(q)|_{q=q(a^*)}$$

where the infimum is reached at $a^* - \widetilde{T}'_h(q(a^*)) = 0$, or $a^* = \widetilde{T_h}'(0) = T'_h(0)$ since $q(a^*) = 0$.

- Under the conditions of Theorem 16, the equality between F and $\widetilde{T}_h$ holds, and T_h is differentiable. We saw before that $\widetilde{T}_\gamma(1) = 0$. It follows that for all a,
$$F(a) = \inf_q(qa - \widetilde{T}_h(q)) \leqslant qa - \widetilde{T}_h(q)|_{q=1} = a$$
with equality at $a^0 = \widetilde{T}'_h(1) = T'_h(1)$.

- Suppose μ is an integer. Then the upper bound derived for $\mathcal{M}$ whose crossing tree is a Galton-Watson tree with average family size μ is exactly the Hausdorff spectrum of a random cascade defined on a μ-ary tree with random weights [102]. This motivates us to believe that the upper bound derived for $\mathcal{M}$ is a tight upper bound.

We simulate 100 approximations of $\mathcal{M}$ of length 2^{15} by constructing finite realisations of Galton-Watson trees with eight levels and random weights and estimate its partition function using two techniques described below. We pick a geometric distribution with parameter 0.5 for p ($\mu = 4$) and we suppose weights are log normal with parameters $m = -0.05 - \ln(\mu)$ and $\sigma^2 = 0.1$ so that the mean value is $1/\mu = 0.25$. Straightforward calculations give

$$\tilde{T}_\gamma(q) = \tilde{T}_h(q) = -\frac{\sigma^2}{2\ln\mu}q^2 - \frac{m}{\ln\mu}q - 1. \tag{4.31}$$

The infimum of $qa - \tilde{T}_\gamma(q)$ is obtained for $q^* = -(m + a\ln\mu)/\sigma^2$. The two intersections a^* of the LF transform of $\tilde{T}_\gamma$ with the horizontal axis are such that $qa^* - \tilde{T}_\gamma(q)|_{q=q^*} = 0 = (-\ln\mu/2\sigma^2)(a^*)^2 - (m/\sigma^2)a^* + (1 - m^2/2\sigma^2\ln\mu)$ that is

$$a^* = -\frac{m}{\ln\mu} \pm \sigma\sqrt{\frac{2}{\ln\mu}}.$$

With the previous choice of m and σ^2, those values are 0.656 and 1.416. The corresponding values of q are ± 5.26, so there is no need to estimate $\tilde{T}_\gamma(q)$ for $|q| > 5.26$.

First, we use Equation (4.25) which provides a direct method for estimating the deterministic coarse spectrum $\tilde{T}_\gamma(q)$. The expectation is approximated by averaging over 100 realisations. Figure 4.4(a) presents the estimated and theoretical $T_\gamma(q) = \tilde{T}_\gamma(q) + 1$. Error bars are too small to be plotted. The two spectra match closely. Figure 4.4(b) is the LF transform of $T_\gamma(q)$. The 95% confidence intervals are derived by estimating $T_\gamma(q)$ 100 times.

Then, assuming $\mathcal{M}$ is Hölder uniform, we use wavelet leaders to estimate the partition function $\zeta_2(q)$ given in Equation (4.3), whose LF transform also provides an upper bound for the Hausdorff spectrum of $\mathcal{M}$. The spectrum is estimated using Daubechies wavelets with three vanishing moments. The scale of analysis ranges from $j_1 = 3$ to $j_2 = 13$. Since taking the LF transform of a function is a very unstable procedure, the estimated spectrum is obtained from the average of 100 estimations of the partition function. Here again, ζ_2 and T_γ match closely for positive and negative q as illustrated in Figure 4.4(c).

Recall that for all $q \in \mathbb{R}$, $\zeta_2(q) = \xi(q)$, where $\xi(q)$ is the LF transform of the Hausdorff spectrum D (Equation (4.4)). If D is concave, then the LF is involutive and the LF transform of $\xi(q)$ gives D back (see Appendix A). Assuming this is the case, which is true for many multiplicative cascades as we noticed earlier in the introduction chapter, then the LF transform of $\zeta_2(q)$ gives the Hausdorff spectrum of $\mathcal{M}$. From this remark, the close similarities between T_γ and ζ_2 tend to indicate empirically that the upper bound is in fact the exact spectrum of $\mathcal{M}$. Proving this result theoretically can be challenging. Deriving lower bounds for Hausdorff dimensions is usually a hard problem since it requires an optimal covering of the set, though techniques exist [45].

4.3.4 An upper bound for the spectrum of a subordinated Brownian motion.

MEBP processes are an example of monofractal processes time changed by a multifractal cascade. When the crossing tree has a geometric offspring distribution with parameter 0.5, CEBP reduces to a standard Brownian motion B [52] and therefore

$$Y = B \circ \mathcal{M}^{-1} \tag{4.32}$$

is a Brownian motion in multifractal time. Mandelbrot first proposed to model financial data by means of a Brownian motion in multifractal time [86, 87]. Thus MEBP processes can potentially have applications in finance. In this section, we are going to look at the multifractal spectrum of Y, defined as in Equation (4.32).

First, we want to relate the Hausdorff spectrum of the chronometer $\mathcal{M}$ with the inverse process $\mathcal{M}^{-1}$. Mandelbrot and Riedi [88, 103] made precise this relationship under general conditions on $\mathcal{M}$. We state their main result in the following lemma.

LEMMA 8. *Let $\mathcal{M}$ be a non decreasing continuous process and suppose it admits a continuous inverse $\mathcal{M}^{-1}$. Denote by $D_H^{\mathcal{M}}$ and $D_H^{\mathcal{M}^{-1}}$ the Hausdorff spectrum of $\mathcal{M}$ and $\mathcal{M}^{-1}$ respectively. Then,*

$$D_H^{\mathcal{M}^{-1}}(a) = aD_H^{\mathcal{M}}(1/a).$$

The next step is to relate the spectrum of $\mathcal{M}^{-1}$ and X when considering the composition $X \circ \mathcal{M}^{-1}$. We need a preliminary result due to Adler [4].

LEMMA 9. *Let I be a compact interval and B a Brownian motion.*

1- (Theorem 8.3.1, [4]) *$\forall \eta > 0$, $\exists \delta > 0$ and V an almost surely positive random variable such that $\forall (s,t) \in I$, $|s-t| \leqslant \delta$, with probability one,*

$$|B(t) - B(s)| \leqslant V|s-t|^{(1/2)-\eta}. \tag{4.33}$$

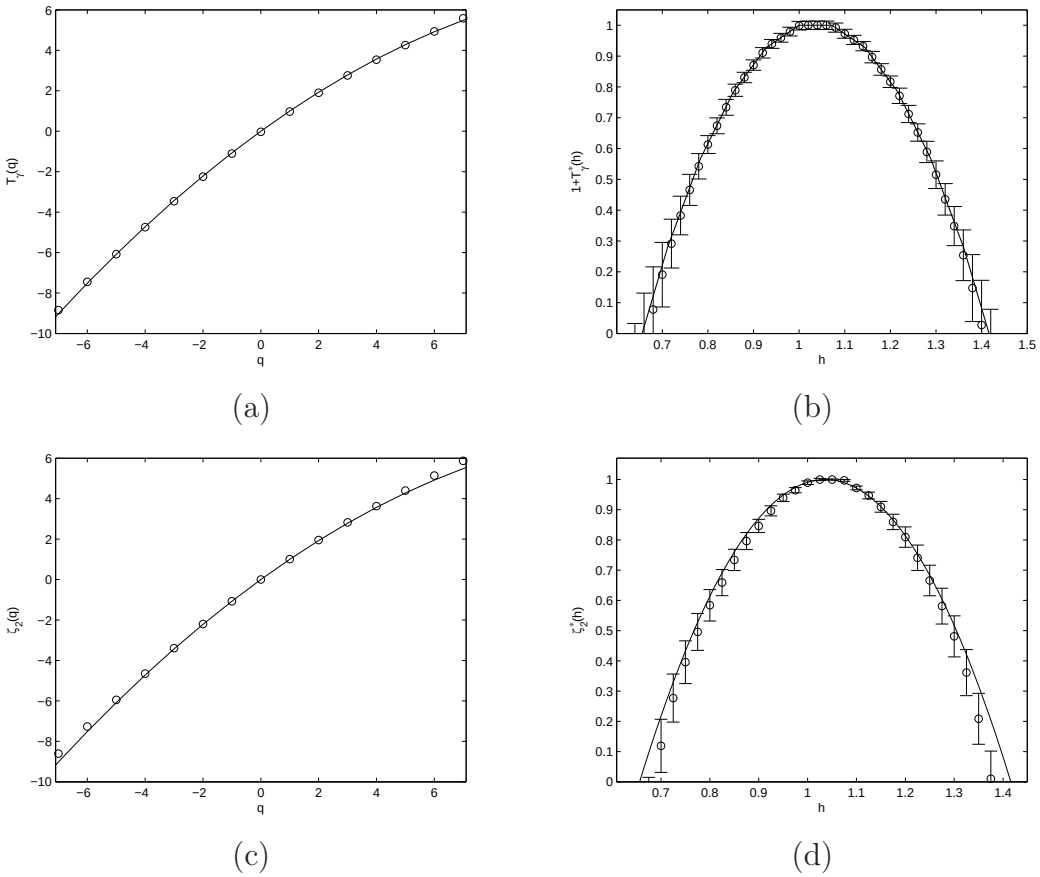


Figure 4.4: **Estimation of the partition function of $\mathcal{M}$.** We present in solid line in (a) and (c) the partition function $T_\gamma(q)$, and its Legendre Fenchel transform in (b) and (d), with 95% confidence intervals. The top figures also display the estimated partition function ('o' in (a)) and its LF transform (b) using a direct method, using Equation (4.25). The bottom two figures show the estimated partition function $\zeta_2(q)$ in the framework of wavelet leaders ('o' in (c)), and its LF transform (d). The close match between $T_\gamma(q)$ and $\zeta_2(q)$ tends to indicate that the multifractal formalism holds here, as explained in the text.

2- ((8.8.26), [4]) $\forall \eta > 0$, $\exists h > 0$,

$$\sup\{|B(t) - B(s)| \mid |s - t| \leqslant h\} \geqslant K h^{(1/2)+\eta} \tag{4.34}$$

for all t in I, all $K < \infty$, with probability one.

Riedi [102] used the previous result to relate the spectrum of a time changed fractional Brownian motion (fBm) to the spectrum of the time change, when the time change is independent of the fBm. However, an MEBP is a CEBP warped with a random cascade, both defined with the same crossing tree. It would therefore be surprising if the chronometer and its CEBP process are independent. As we shall see now, the following result holds nevertheless:

LEMMA 10. *Consider an MEBP process $Y = B \circ \mathcal{M}^{-1}$ obtained from a time changed CEBP process B, which reduces to a Brownian motion. Denote by D_H^Y and $D_H^{\mathcal{M}^{-1}}$ the Hausdorff spectrum of Y and $\mathcal{M}^{-1}$ respectively. Given Assumptions 1, 2 and 3 hold, $\mathcal{M}$ and $\mathcal{M}^{-1}$ are continuous and*

$$D_H^Y(a) = D_H^{\mathcal{M}^{-1}}(2a).$$

Proof. We recall the definition of the Hölder exponent h_Y of Y (Definition 9):

$$h_Y(t) := \liminf_{\epsilon \to 0} \frac{1}{\log(2\epsilon)} \log \sup_{(u,u') \in I_\epsilon} |Y(u) - Y(u')| \tag{4.35}$$

where $I_\epsilon = [t - \epsilon, t + \epsilon]$.

First step: Clearly

$$\begin{aligned} \sup_{(u,u') \in I_\epsilon} |Y(u) - Y(u')| &= \sup_{(u,u') \in I_\epsilon} |B \circ \mathcal{M}^{-1}(u) - B \circ \mathcal{M}^{-1}(u')| \\ &= \sup_{s,s'} \{|B(s) - B(s')| \mid \mathcal{M}^{-1}(t - \epsilon) \leqslant s' < s \leqslant \mathcal{M}^{-1}(t + \epsilon)\} \end{aligned}$$

where $u = \mathcal{M}(s)$ and $u' = \mathcal{M}(s')$.

Let $\eta > 0$ and $\delta > 0$. There exists ϵ_0 such that for all $\epsilon \leqslant \epsilon_0$, $\sup |\mathcal{M}^{-1}(u) - \mathcal{M}^{-1}(u')| \leqslant \delta$, where the supremum is over I_ϵ. Then, $|s - s'| \leqslant \delta$ and from Lemma 9, Part 1, there exists an almost surely finite positive random variable V such that

$$\sup_{(u,u') \in I_\epsilon} |Y(u) - Y(u')| \leqslant V \sup_{(u,u') \in I_\epsilon} |\mathcal{M}^{-1}(u) - \mathcal{M}^{-1}(u')|^{(1/2)+\eta}$$

Dividing each member of the inequality by $\log 2\epsilon < 0$ and letting $\epsilon \to 0$ and $\eta \to 0$, yields

$$2h_Y(t) \geqslant h_{\mathcal{M}^{-1}}(t).$$

Second step: To prove the reverse inequality, use Part 2 of Adler's result, following the same procedure as in Step 1. □

THEOREM 20. *Let $Y = B \circ \mathcal{M}^{-1}$ be an MEBP process, where B is a Brownian motion. Suppose Assumptions 1, 2 and 3 hold. Let $D_H^{\mathcal{M}}$ and D_H^Y be the Hausdorff spectra of Y and $\mathcal{M}$ respectively. Then*

$$D_H^Y(a) = 2a D_H^{\mathcal{M}}(1/2a).$$

Proof. This follows directly from Lemmas 8 and 10. □

We present in Figure 4.5 estimations of the partition function $\zeta_2(q)$ for various subordinated Brownian motions, assuming they are Hölder uniform. Assuming that the multifractal formalism holds, the upper bound of the spectrum of $\mathcal{M}$ is in fact its Hausdorff spectrum and the LF transform of $D_H^Y(a)$, which will be the same as $\zeta_2(q)$, can be computed given Theorem 20. We present on the same graph both estimations $\zeta_2(q)$ and the LF transform of $D_H^Y(a)$.

In the top and middle figure, we consider a time change obtained with lognormal weights, with parameters $m = -0.05 - \ln 4$ and $\sigma^2 = 0.1$. The figure shows an average of 100 estimations of the partition function, each obtained with an analysis of a process of length 2^{14} from scale $j_1 = 3$ to $j_2 = 11$. We used Daubechies wavelets with three vanishing moments. The original Brownian motion in multifractal time generated with the algorithm described in the previous chapter has a non constant sampling period. We need to resample it at a constant rate to estimate its partition function. In the top figure, we resampled every $\delta = 14$ units of time and in the middle one every $\delta = 6$ units of time. The results emphasize the issue of the resampling step in the estimation of the partition function for negative q's. While the estimation is correct for $\delta = 14$, the negative part of $\zeta_2(q)$ is wrongly estimated for $\delta = 6$.

To support this observation, we consider a Brownian motion in multifractal time, whose subordinator is a deterministic binomial cascade with weights $m_0 = 0.6$ and $m_1 = 0.4$ (see Section 1.1.8). The Hausdorff spectrum of binomial cascades is known and given by Equations 1.10 and 1.11. The estimated partition function presented on the bottom figure is obtained as an average of 100 estimations obtained with Daubechies wavelets with two vanishing moments. The length of the samples is 2^{15},the scale of analysis ranges from $j_1 = 3$ to $j_2 = 12$ and we set $\delta = 10^{-4}$. The choice $\delta = 10^{-4}$ is much different from the previous choices $\delta = 6, 14$ since we have two distinct time scales. This is of no importance here as we only want to check the impact of resampling when the resampling period is too small or too large with respect to the mean distance between two samples. The small discrepancies observed between the theory (known since the multifractal formalism holds for the deterministic binomial cascade) and practice mainly comes from the resampling period. If we decrease δ to 10^{-5}, the estimation gets much worse (Figure not presented here).

We explain the sensitivity of estimations to the resampling period as follows. By doing a linear interpolation, we modify information about the increments of the process. These errors are transmitted to the wavelet coefficients and therefore to the wavelet leaders. When raised to a negative power of q, errors are magnified from which the estimation of the negative part of the partition function is erroneously estimated. The difficulty to estimate the partition function for negative q's is a widely known problem [66, 101].

For the same reasons as before, the empirical results presented here motivate us to hope that the multifractal formalism holds for $Y = B \circ \mathcal{M}^{-1}$. We attribute estimation errors mainly to the resampling step, to which the estimation of $\zeta_2(q)$ is highly sensitive.

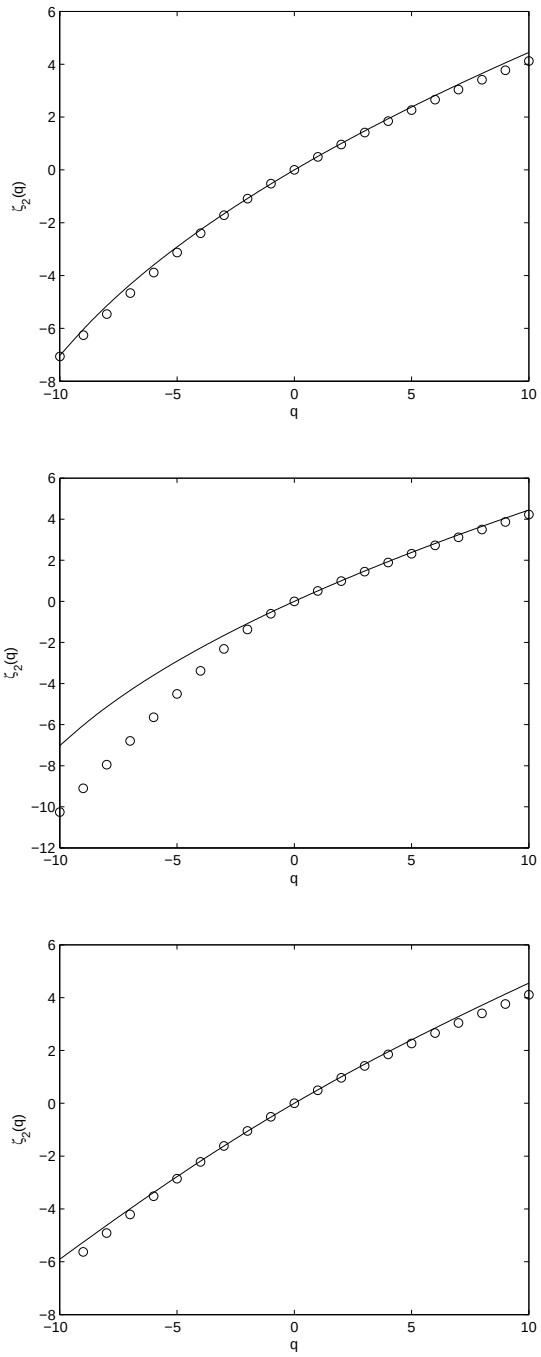


Figure 4.5: Estimation of $\zeta_2(q)$ for Brownian motions in multifractal time. Top and middle figure: lognormal weights with identical parameters. Resampling is respectively 14 and 6 units of time. Bottom figure: time change is a deterministic binomial cascade with weights 0.6 and 0.4. The solid line represents the LF transform of $D_H^Y(a)$. Error bars are too small to be plotted.

4.4 Special points on the spectrum of ν

In the previous sections, we have been interested in deriving the Hausdorff spectrum of $\mathcal{M}(t) = \zeta([0,t))$, where ζ is defined on the real line. ζ is constructed from the random measure ν whose discrete support is the boundary of the crossing tree (see the previous chapter for the definition of ν). When we consider Galton-Watson trees with fixed weights equal to $1/\mu$ where μ is the mean family size, ν reduces to a measure η known as the branching measure. Similarly, we call ν the weighted branching measure.

The study of the spectrum of measures defined on the boundary of a tree differs widely from standard techniques for measures defined on the real line. The underlying space is different (compact interval vs a subset of $\mathbb{N}^{\mathbb{N}}$) and the metric defined on it is not the same. Therefore, it is unlikely that ζ and ν have the same spectrum. Let Υ be a Galton-Watson tree and $\partial\Upsilon$ its boundary. We equip $\partial\Upsilon$ with the metric d given by

$$d(\mathbf{i},\mathbf{j}) = e^{-n(\mathbf{i},\mathbf{j})} \tag{4.36}$$

where $n(\mathbf{i},\mathbf{j})$ is such that $\mathbf{i}|k = \mathbf{j}|k$ for all $k \leqslant n$ and $\mathbf{i}|k \neq \mathbf{j}|k$ otherwise; it is the generation of the last common ancestor of $\mathbf{i},\mathbf{j} \in \partial\Upsilon$. Given the metric d, the local Hölder regularity of a measure ν at a point x is defined as

$$\lim_{r\to 0} \frac{\log \nu[B(x,r)]}{\log r}$$

where $B(x,r)$ is the open ball centered at x with radius r. For each $a > 0$, consider the set E_a of points $\mathbf{i} \in \partial\Upsilon$ at which the Hölder regularity of ν is a. When the sets E_a are non empty and fractal over a range of values of a, ν is said to be multifractal [44]. When the Hölder regularity is the same for all $\mathbf{i} \in \partial\Upsilon$, ν is called monofractal.

In Section 4.4.1, we give the definition of the branching measure η and recall some famous results on its Hausdorff spectrum. Then, we study particular points on the multifractal spectrum of ν and point out the similarities with the branching measure. Burd and Waymire [30] have studied independent random cascades on Galton-Watson trees, leading to the same definition of ν. They proved the existence of a Borel subset of the boundary of the tree which contains the whole mass ν and gave its Hausdorff dimension. Here, we obtain results about the Hölder exponent of points $\mathbf{i} \in \partial\Upsilon$, which differs from the study in [30].

4.4.1 The branching measure η

Consider a Galton-Watson tree Υ with regular offspring distribution p. Denote by $\emptyset$ its root node. Let Z_1 be the size of the first generation. $p(Z_1 = k) = p_k$. Let $\mu = \mathbb{E}Z_1$ be the mean family size and suppose $\mathbb{E}Z_1 \log Z_1 < \infty$. Equip the boundary of the tree $\partial\Upsilon$ with the metric (4.36). Denote by $C_{\mathbf{i}}$ the cylinder set defined by $C_{\mathbf{i}} = \{\mathbf{j} \in \partial\Upsilon \mid \mathbf{j}|_n = \mathbf{i}\}$. Let $\Upsilon_{\mathbf{i}}$ be defined as $\mathbf{i}$: $\Upsilon_{\mathbf{i}} = \{\mathbf{j} \in \Upsilon \mid |\mathbf{j}| \geqslant |\mathbf{i}| \text{ and } \mathbf{j}|_{|\mathbf{i}|} = \mathbf{i}\}$ and Υ_n^{GW} the n-th generation of Υ. Denote by $Z_n(\mathbf{i})$ the size of generation n in $\Upsilon_{\mathbf{i}}$. For each node $\mathbf{i} \in \Upsilon$, we define the martingale limit $W_{\mathbf{i}}$ as

$$W_{\mathbf{i}} = \lim_{n\to\infty} \frac{Z_n(\mathbf{i})}{\mu^{n-|\mathbf{i}|}}.$$

Then, almost surely on $\partial\Upsilon \neq \emptyset$, there exists a unique nonzero finite measure η on the boundary of the tree, called the branching measure such that

$$\eta(C_{\mathbf{i}}) = \frac{W_{\mathbf{i}}}{\mu^{|\mathbf{i}|}}.$$

Clearly,

$$\eta(\partial\Upsilon) = W_{\emptyset} = \mu^{-n} \sum_{\mathbf{i}\in\Upsilon_n^{GW}} W_{\mathbf{i}}.$$

The branching measure η has been studied by many authors and we refer to the works of Liu [79, 80], Mörters and Shieh [94], Shieh and Taylor [110], Hawkes [55]. Hawkes pioneered the study of the branching measure with the following result

THEOREM 21 (Theorem 1 [55]). *Suppose* $\sum_{k=2}^{\infty} p_k k(\log k)^2 < \infty$. *Then, almost surely on trees, for* $\mathbf{i} \in \partial\Upsilon$

$$\lim_{n\to\infty} \frac{\log \eta[B(\mathbf{i}|n, r)]}{-n} = \log \mu$$

except on an η *negligible set.*

The theorem states the local dimension of the branching measure at a boundary point is $\log \mu$ except possibly on a $\mathbb{P} \times \eta$ negligible set. There exists points $\mathbf{i} \in \partial\Upsilon$ which have a local regularity which differs from $\log \mu$. The condition $\sum_{k=2}^{\infty} p_k k(\log k)^2 < \infty$ is equivalent to $\mathbb{E}W_{\emptyset} \log W_{\emptyset} < \infty$ (see Athreya [9]). We want to know when Theorem 21 holds for all boundary points, or equivalently when the spectrum of η degenerates to a single point. This issue is addressed in Liu [80]

THEOREM 22 (Theorem 4.1 [80]). *Suppose that* $p_1 = 0$ *and* $\mathbb{E}Z_1^a < \infty$ *for all* $a > 1$. *Then, almost surely on trees,*

$$\lim_{n\to\infty} \frac{\log \eta[B(\mathbf{i}|_n, r)]}{-n} = -\log \mu$$

for all $\mathbf{i} \in \partial\Upsilon$.

Conditions of Theorem 22 are stronger than in Theorem 21. Bingham and Doney have shown [27] that $\mathbb{E}Z_1^a$ and $\mathbb{E}W_{\emptyset}^a$ converge or diverge together for all $a > 1$. Therefore, the condition $\mathbb{E}W_{\emptyset} \log W_{\emptyset} < \infty$ for Theorem 21 hold under the assumptions of Theorem 22. We now derive similar results for the weighted branching measure ν and show the correspondence with the branching measure η when we consider constant weights equal to $1/\mu$.

4.4.2 The weighted branching measure ν

Consider Υ an extended Galton-Watson tree whose branches are equipped with a random weight. Under Assumption 1, the martingale limit attached to node $\mathbf{i}$ of the tree

$$\mathcal{W}_{\mathbf{i}} = \lim_{n\to\infty} \sum_{\mathbf{j}\in\Upsilon_{\mathbf{i}}, |\mathbf{j}|=n} \rho_{\mathbf{j}}/\rho_{\mathbf{i}}.$$

exists in $(0,\infty)$ and is such that $\mathbb{E}\mathcal{W}_{\mathbf{i}} = 1$ (see Chapter 3 for further details). We define

$$\nu(C_{\mathbf{i}}) = \rho_{\mathbf{i}}\mathcal{W}_{\mathbf{i}}$$

and we call ν the weighted branching random measure.

First we derive an equivalent of Theorem 21 for ν. We need a result due to Biggins.

ASSUMPTION 7.

$$\mathbb{E}\Big[\Big(\sum_{i=1}^{Z_1}\rho_i\Big)\Big(\log\Big(\sum_{i=1}^{Z_1}\rho_i\Big)\Big)^2\Big] < \infty.$$

THEOREM 23 (page 28 of [22]). *Given Assumptions 1 and 7 hold,* $\mathbb{E}\mathcal{W}_\emptyset \log \mathcal{W}_\emptyset < \infty$.

THEOREM 24. *Given Assumptions 1 and 7 hold, then almost surely on trees,*

$$\lim_{n\to\infty}\frac{\log \nu[B(\mathbf{i}|_n, r)]}{-n} = \mathbb{E}\log\rho$$

for ν-almost all $\mathbf{i} \in \partial\Upsilon$.

Proof. The proof is similar to Theorem 21 [55]. It is obtained by replacing $\mu^n\eta[B(\mathbf{i}|_n, r)]$ by $\rho_{\mathbf{i}}^{-1}\nu[B(\mathbf{i}|_n, r)]$ everywhere in his proof. In the proof of Theorem 21, the condition $\mathbb{E}W_\emptyset \log W_\emptyset < \infty$ is needed. Similarly, one needs $\mathbb{E}\mathcal{W}_\emptyset \log W_\emptyset < \infty$ here, which holds under Assumption 7. Existence and non-degeneracy of $\mathcal{W}_\emptyset$ follows under Assumption 1. □

LEMMA 11. *Suppose Assumptions 1 and 3 hold. Then, almost surely on trees,*

$$\log \max_{\mathbf{i}\in\Upsilon_n^{GW}} \mathcal{W}_{\mathbf{i}} = o(n).$$

Proof. The proof is similar as the one found in Lemma 6. Just replace $W_{\mathbf{i}}$ by $\mathcal{W}_{\mathbf{i}}$. The result follows under the condition that $\mathcal{W}_{\mathbf{i}}$ admits moments of all orders, which holds under Assumption 3. □

THEOREM 25. *Given Assumptions 1 and 3,*

$$\lim_{n\to\infty}\frac{\log \nu[B(\mathbf{i}|_n, r)]}{-n} = \mathbb{E}\log\rho \tag{4.37}$$

for all $\mathbf{i} \in \partial\Upsilon$, with probability 1 with respect to trees.

Proof. Let $\mathbf{i} \in \partial\Upsilon$.

$$\frac{\log \nu[B(\mathbf{i}|_n, r)]}{n} = \frac{\log \rho_{\mathbf{i}|_n}\mathcal{W}_{\mathbf{i}|_n}}{n} = \frac{\log \rho_{\mathbf{i}|_n}}{n} + \frac{\log \mathcal{W}_{\mathbf{i}|_n}}{n}.$$

Under Assumption 3, the maximum of the $\mathcal{W}_{\mathbf{i}|_n}$ tends to 0 with n (Lemma 11). The second term of the right hand side tends to 0 as n tends to infinity. The first term can be reexpressed

$$\frac{\log \rho_{\mathbf{i}|_n}}{n} = \frac{1}{n}\sum_{k=1}^{n}\log\rho_{i_k}(\mathbf{i}|_k) \to \mathbb{E}\log\rho \text{ a.s.}$$

from the strong law of large numbers. □

From Theorem 25, adding weights 'regular' enough (in the sense that they admit moments of all orders) does not destroy the monofractal property of the branching measure η.

Set constant weights $\rho = 1/\mu$. The assumptions of Theorem 25 reduce to assumptions of Theorem 22 and the 2 theorems do not contradict themselves. This study present an alternative proof of the result of Liu. When (4.37) holds for all $\mathbf{i} \in \partial\Upsilon$, the Hausdorff dimension of the set $K_{\mathbb{E}\log\rho} := \{\mathbf{i} \in \partial\Upsilon \text{ with regularity } \mathbb{E}\log\rho\}$ is $\log\mu$. This follows from the classical result of the Hausdorff dimension of the boundary $\partial\Upsilon$ of a Galton-Watson tree to be $\log\mu$. See for instance [55, 81].

Chapter 5

Further work

In this thesis, we have studied two models of (multi)fractal signals with underlying branching structure. The first one, called Galton-Watson IFS, proposes a generalization of the construction of a random IFS acting over the space $\mathbb{L}_p$ of p-integrable functions. The second one, called MEBP processes, is constructed using the concept of crossing tree of a process. We now discuss open questions about those two processes.

Galton-Watson IFS. Chapter 2 presents the theoretical foundations for the existence and uniqueness of a fixed point when considering an IFS with random operators and random construction tree. Simulations also indicate a possible multifractal structure of these processes. It would be interesting to validate this in theory. One could start by considering deterministic operators after conditioning on the number of offspring, and give a simple form for these operators, keeping a random tree structure. Some work on the spectrum of fixed point of IFS has been done before; we could refer to those results as a starting point. See for example [65].

The problem of estimating the parameters of the model could also be contemplated. As far as I am aware, there are no results concerning estimation techniques of the parameters of a deterministic IFS acting over the space of random functions. In other words, given one or more realisations of a signal, can we derive a method to estimate the functions ϕ, ϱ and the distribution of $Z_\emptyset$ and then fit a Galton-Watson IFS? This inverse problem is far from being trivial, even when the tree is deterministic. One could first give a particular form of random operators, assuming an M-ary tree, and estimate the first and second moments of the random parameter. This problem was suggested in [36], but a much deeper analysis is needed to generalize the results to arbitrary random operators. As a starting point, one could refer to existing techniques in image compression where attempts are made to represent a target image as the attractor of an IFS, within a degree of accuracy [17, 46, 64, 112].

MEBP processes Chapters 3 and 4 present a new class of multifractal processes, called MEBP, generalizing the construction of the CEBP process.

It is known that CEBP processes are discrete scale invariant (Theorem 6) and monofractal (Theorem 10). It would be interesting to relate the Hölder exponent of the CEBP process to the box dimension of the sample path. Indeed, one can expect

the graph in $\mathbb{R}^2$ of a CEBP process to have box dimension $2-\log 2/\log \mu$. To see this, let X^m be the random walk on $2^m\mathbb{Z}$, defined by $X^m(k) = X(T_k^m)$ for $k = 1, 2, \ldots$ where X is a CEBP process for some offspring and orientation distributions and T_k^m is the k-th level m hitting time. Then one can cover X^m with $Z_m(\lceil \mu/2 \rceil)^m$ boxes of size $\mu^{-m} \times \mu^{-m}$. Since Z_m/μ^m converges to some fixed limit as $m \to \infty$, we have

$$\frac{\log(Z_m(\lceil \mu/2 \rceil)^m)}{\log \mu^m} \to 2 - \frac{\log 2}{\log \mu},$$

where $\log 2/\log \mu$ is the Hölder exponent of the CEBP process (Theorem 10). To prove this result rigorously, one needs to consider coverings of the process X itself instead of coverings of X^m. The presence of the double limit here (box size and path) is problematic and does not let us conclude immediately that the box dimension of the graph of X is indeed $2 - \log 2/\log \mu$.

Although we have derived an upper bound for the Hausdorff spectrum of the time change, it would be interesting to derive theoretically the exact spectrum. From the simulations, it is believed that the upper bound obtained is in fact a tight bound, if not the exact spectrum. Also, a generalization of the multifractal study for general MEBP is needed. Only the subordinated Brownian motion has been studied (Section 4.3.4).

Applications of MEBP processes may include modelling financial time series. If P_t is the price at time t of a risky asset, it is commonly accepted that depending on the time interval over which one take the returns, the log-returns $\log P_t - \log P_{t-1}$ show heavier tails than a Gaussian distribution, possess a strong non-linear dependence and a high volatility and intermittency [56]. The marginal distribution of returns can be fitted very well with Student t-distributions with ν degrees of freedom, typically such that $3 \leqslant \nu \leqslant 5$ [58, 59]. Figure 5.1 presents the marginal distribution of an MEBP process with $2\,\text{geometric}_1(0.5)$ offspring distribution and i.i.d. Gamma weights with parameters $\alpha = 3.5$ and $\beta = 1$. We have fitted a Student-t distribution with 4 degrees of freedom. The Q-Q plot shows a very good fit. Furthermore, increments of this MEBP process have very little linear correlation (bottom left image), as it should be for financial time series, since it is known that the autocorrelations of the log returns die away quickly. Moreover, there is correlation between the absolute value of the increments (bottom right image), which is another accepted property of the log-returns.

To fit data with an MEBP process, one needs to estimate the offspring distribution of the crossing tree, the distribution of orientations and finally the weight distribution.

The estimation of the offspring distribution is quite straightforward. One can start by calculating the crossing tree of a given signal, and then computing the relative frequency of each family size. Also, one may want to estimate the orientation distribution. For diffusion processes, one expects i.i.d. orientations, independent of the family size [70]. This can be tested in practice by considering a sequence of N excursions. If each of them is 'up-down' or 'down-up' with probability $1/2$, then the

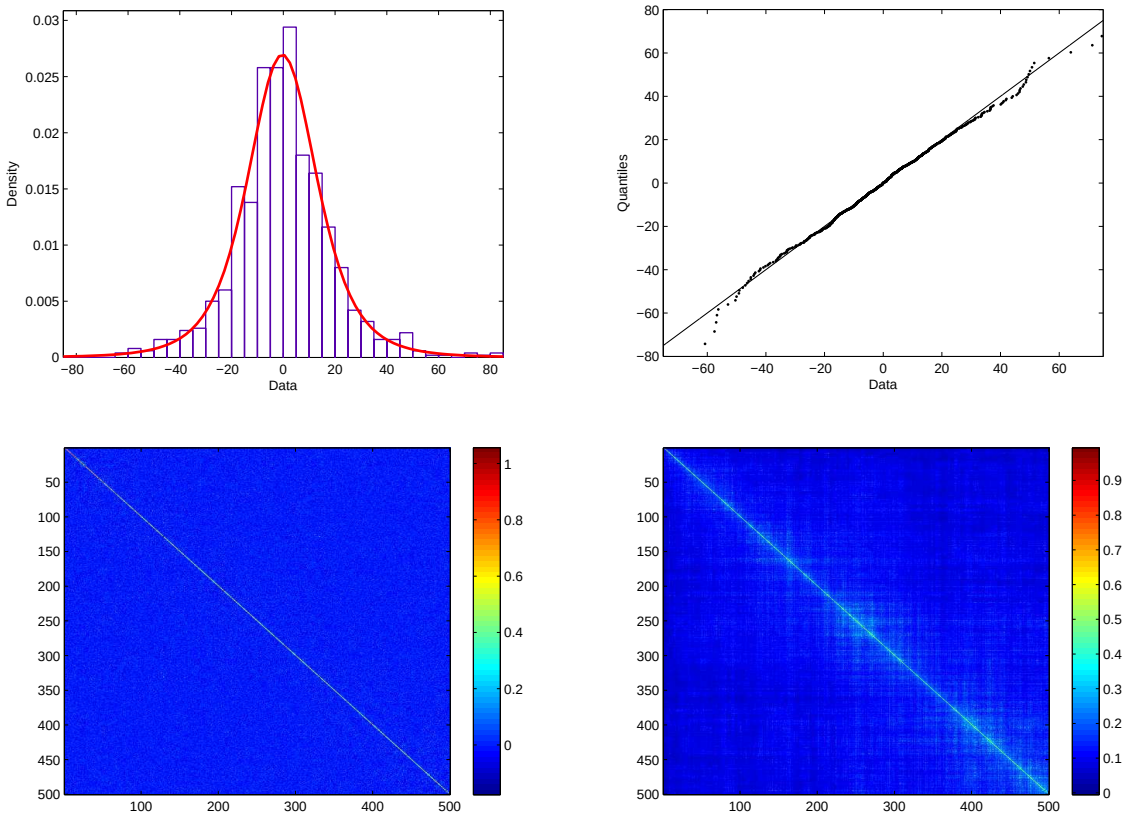


Figure 5.1: Marginal distribution of an MEBP process with $2\,\text{geometric}_1(0.5)$ offspring distribution and i.i.d. gamma weights with parameters $\alpha = 3.5$ and $\beta = 1$ with a fitted t distribution, Q-Q plot and correlation matrix of returns and absolute returns. Estimated mean, variance and degrees of freedom are respectively -0.35, 13.91 and 4.01. The correlation matrices are estimated using 1000 realisations of the process, each of length 500.

sum of N observations should have a binomial distribution with parameters N and $1/2$. In [70], Jones and Rolls developed methods to test whether or not a signal is a time changed Brownian motion, based on the previous observations.

Estimating the weight distribution from one realisation of a signal is still an open problem. One could start by considering the case of i.i.d. weights. Consider a finite crossing tree with n generations starting from a level 0 crossing. By taking the logarithm of each weight $\rho_{\mathbf{i}}\mathcal{W}_{\mathbf{i}}$ on the n-th generation, one transforms products of i.i.d. random variables along a line of descent to a sum

$$\log \rho_{\mathbf{i}}\mathcal{W}_{\mathbf{i}} = \log \rho_{i_1}(\emptyset) + \ldots + \log \rho_{i_n}(\mathbf{i}|_{n-1}) + \log \mathcal{W}_{\mathbf{i}}.$$

Thus, it is possible to express the problem as follows

$$y = Ax + \epsilon$$

where y are the observed values $\log \rho_{\mathbf{i}}\mathcal{W}_{\mathbf{i}}$, A is a sparse matrix with entries 0 and 1, x is a column vector containing the log of the weights and ϵ a column vector containing $\log \mathcal{W}_{\mathbf{i}}$. If N denotes the total number of weights on the first n generations of the finite crossing tree and M is the cardinality of generation n, then y is an $M \times 1$ vector, A is an $M \times N$ matrix, x is an $N \times 1$ vector and ϵ a $M \times 1$ vector. Since the crossing tree of a process is known, the matrix A is known. However, the rank of A is not full and one cannot directly invert this system to recover the original weights. Different methods can be investigated, such as a ridge regression.

One can also give a parametric distribution to the weights. Considering lognormal weights on the crossing tree, weights are i.i.d. normal after a log transform. In a first step, one could ignore the vector of errors ϵ and consider the system $y = Ax$. Given this distribution, one could derive the likelihood of the joint distribution of y and derive maximum likelihood estimators of the parameters of the normal distribution. When this problem is solved, one could add the vector of errors ϵ, consider other weight distributions and contemplate a non-parametric approach.

Appendix A.
The Legendre Transform

In this appendix, we address a few questions concerning the Legendre-Fenchel (LF) transform, in the context of multifractal analysis. Let $\tau(q)$ be any function. Define its LF transform $\tau^*(h)$ by

$$\tau^*(h) = \inf_{q\in\mathbb{R}}(qh - \tau(q)).$$

The LF transform is a generalisation of the Legendre transform, defined by

$$\tau_L^*(h) = hq^* - \tau(q^*), \tag{5.1}$$

where q^* is given by solving $\tau'(q) = h$. This transform is only defined for differentiable functions. We will see later in this appendix that the LF transform reduces to the Legendre transform for concave and differentiable functions.

We address the following questions: why does the LF transform plays a central role in the theory of multifractals? How do we derive the LF transform of a given function? What are the properties of this transform? Why does the negative (respectively positive) part of a partition function typically correspond to the increasing (resp. decreasing) part of the Hausdorff spectrum after LF transform, for many cascade processes?

LF transform in multifractal analysis. To give mathematical justification of how the LF transform appears in multifractal analysis, consider a measure μ, whose fluctuations are described by its local Hölder exponent at any $x \in \mathbb{R}^n$, given by

$$\dim_{\text{loc}}\mu(x) = \lim_{r\to 0}\frac{\log\mu(B(x,r))}{\log r},$$

where $B(x,r)$ is the open ball centered at x with radius r. The Hausdorff spectrum $D(h)$ of μ is a global description of its local fluctuations, given by the Hausdorff dimension of the set of points with Hölder exponent h. The complexity in estimating $D(h)$ directly in practice has lead to the discovery of alternative methods, giving birth to the multifractal formalism, which provides methods to recover the multifractal spectrum via the LF transform of some partition function. For example, the spectrum of multiplicative cascades is usually related to a so-called partition function $\tau(q)$, whose LF transform is typically equal to $D(h)$. Let us now give mathematical justification to the previous claim. To do so, write

$$N_r(h) = \#\left\{\text{r-mesh cubes } A \text{ with } \mu(A) \geqslant r^h\right\}.$$

The coarse multifractal spectrum of μ is given by [44]

$$f_c(h) := \lim_{\epsilon \to 0} \lim_{r \to 0} \frac{\log^+ \left(N_r(h+\epsilon) - N_r(h-\epsilon) \right)}{-\log r},$$

where $\log^+ x = \max(0, \log x)$. It is a famous result that f_c provides an upper bound for $D(h)$ (Lemma 11.1 in [44]). However, due to the presence of a double limit, f_c is not much simpler to estimate than $D(h)$. In fact, the coarse spectrum is related to the sum of the q-th moments of μ. Define

$$M_r(q) = \sum \mu(A)^q,$$

where the sum is taken over all r-mesh cubes A such that $\mu(A) > 0$. For all $h \geqslant 0$, it follows from the definition of $N_r(h)$ that for all $q \geqslant 0$,

$$M_r(q) = \sum \mu(A)^q \geqslant \sum (r^h)^q = r^{qh} N_r(h).$$

If $q < 0$, we have

$$M_r(q) \leqslant r^{qh} \# \left\{ \text{r-mesh cubes } A \text{ with } 0 < \mu(A) < r^h \right\}. \tag{5.2}$$

We are interested in the power law behaviour of $M_r(q)$, that we identify with

$$\tau(q) := \lim_{r \to 0} \frac{\log M_r(q)}{\log r}$$

provided the limit exists. If not, one can use lim inf and lim sup in the definition of τ. The following result can be found in [44], Lemma 11.2.

Let $q \geqslant 0$ and $\epsilon > 0$. Then, from the definition of $f_c(h)$ and r small enough,

$$\begin{aligned} M_r(q) &\geqslant r^{q(h+\epsilon)} N_r(h+\epsilon) \\ &\geqslant r^{q(h+\epsilon)} r^{-f_c(h)+\epsilon}. \end{aligned}$$

From the definition of τ,

$$f_c(h) \leqslant qh - \tau(q)$$

by sending ϵ to 0. It can be shown that this inequality also holds for $q < 0$, using relation (5.2). Since it holds for all $q \in \mathbb{R}$, we have

$$f_c(h) \leqslant \inf_{q \in \mathbb{R}} (qh - \tau(q)) := \tau^*(h).$$

The LF transform can also appear by considering a Large Deviation approach to the problem, which enables a statistical description of the distribution of the Hölder exponents. The Gärtner-Ellis theorem, central in the theory of large deviations, relates probabilistic quantities through the LF transform (Theorem 13). This transformation is however not limited to the study of multifractal measures, but plays an important role in physics, notably in thermodynamics.

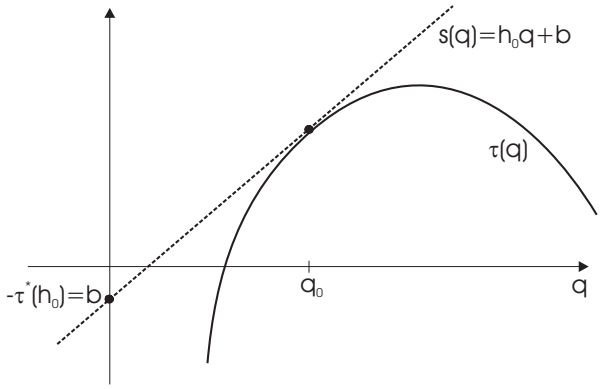


Figure 5.2: Illustration of the Legendre transform.

A brief study of the LF transform. The results of this section are based mainly on the Appendix A of [105]. Let $\tau(q)$ be any function. We first make the following assumption about τ:

i- τ is concave at some q_0.
This assumption is convenient when studying the LF transform since it enables us to calculate τ^* at a special point. τ is concave at q_0, so there exists a linear function $s(q) = h_0 q + b$ such that $\tau(q) \leqslant s(q)$ for all $q \in \mathbb{R}$, with equality at $q = q_0$. h_0 may not be unique. Clearly

$$qh_0 - \tau(q) \geqslant qh_0 - s(q)$$

with equality at $q = q_0$. It follows,

$$\begin{aligned} \inf_q (h_0 q - \tau(q)) &= q_0 h_0 - s(q_0) \\ &= -b \\ &= \tau^*(h_0). \end{aligned}$$

Therefore, $\tau^*(h_0) = -b$ and $\tau^*(h_0) = q_0 h_0 - \tau(h_0)$; the LF transform of τ at q_0 is the opposite of the intercept of s with the ordinate axis, as illustrated Figure 5.2.

ii- In addition, assume τ is differentiable at q_0.
Then h_0 is unique and the slope h_0 is given by the derivative of τ evaluated at q_0:

$$\frac{d}{dq}(qh_0 - \tau(q))|_{q=q_0} = h_0 - \tau'(q_0) = 0,$$

which implies $h_0 = f'(q_0)$. We have indeed a maximum

$$\frac{d^2}{dq^2}(qh_0 - \tau(q))|_{q=q_0} = -\tau''(q_0) > 0,$$

since we assume τ concave at q_0. Then invert $h_0 = \tau'(q_0)$ to express q_0 as a function of h_0: $q_0(h_0)$. It follows that

$$\tau^*(h_0) = q_0(h_0)h_0 - \tau(q_0(h_0)).$$

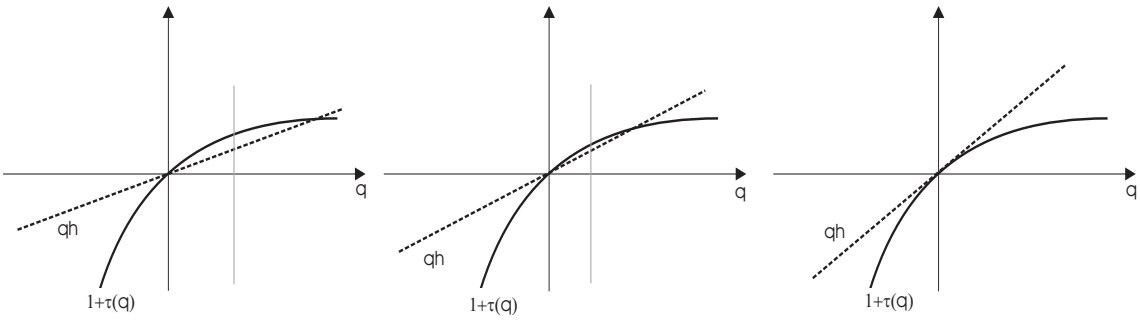
For concave differentiable functions, the LF transform therefore reduces to the Legendre transform of τ (equation(5.1)). Some other important properties of the LF transform are summarized below. More details and proofs can be found in [106].

- The LF transform of τ is always concave.
- $(\tau^*)^*(h) = \tau(h)$ if and only if τ is concave at h. We say that the LF transform is involutive.
- If τ is not concave, then $(\tau^*)^*$ is the smallest concave function satisfying $\tau \leqslant (\tau^*)^*$. $(\tau^*)^*$ is then usually called concave hull of τ.

LF transform for multiplicative cascades. We now study the LF transform from the paradigm of multiplicative cascades. Typically, the Legendre-Fenchel transfor of $\tau(q)$ provides an upper bound for the multifractal spectrum

$$D(h) \leqslant \inf_q(qh - \tau(q))) := \tau^*(h)$$

We have presented in the introduction of this thesis different expressions for $\tau(q)$, by looking at binary, c-ary, deterministic and random cascades (equations (1.11), (1.12) and (1.13)). They all share the property $\tau(0) = -1$. In the remainder of this section, we consider $\tilde{\tau}(q) = 1 + \tau(q)$ instead of $\tau(q)$ so that $\tilde{\tau}(0) = 0$. We just keep in mind that we need to add one when going back to the Hausdorff spectrum. The typical shape of $\tilde{\tau}$ is plotted below (compare with Figure 1.3, for a deterministic binomial cascade). The LF transform compares $\tilde{\tau}(q)$ with the linear function qh and then returns the infimum of the difference between qh and $\tilde{\tau}(q)$ with respect to q. The figure below depicts a typical $\tilde{\tau}(q)$ for multiplicative cascades, together with qh, for different values of h:



The left figure is obtained for very small values of h, so that $\inf(qh - \tilde{\tau}(q))$, whose position is indicated by a grey line, is smaller than -1, corresponds to $\tilde{\tau}^*(h) < 0$, the LF transform of $\tilde{\tau}$. The middle figure corresponds to small h, but not too small, so that the value of the infimum lies between -1 and 0, which yields $\tilde{\tau}^*(h) \in [0, 1]$. Finally, the right plot is obtained for a critical value h^* for which the infimum is 0, corresponding to the maximum value of the Hausdorff spectrum $\tilde{\tau}^*(h^*) = 1$. In summary, $\tilde{\tau}^*(h)$ increases with h, where the value of the infimum is obtained for q positive.

Likewise, for values of h slightly larger than h^*, the infimum becomes negative once again but remains betweens -1 and 0, corresponding to the positive decreasing part of $\tilde{\tau}^*(h)$. For h large enough, the infimum is smaller than -1 and corresponds to negative values of $\tilde{\tau}^*(h)$.

To sum up, τ for q positive typically corresponds to the increasing part of the Hausdorff spectrum and its negative part to the decreasing part of $\tilde{\tau}^*(h)$.

Bibliography

[1] ABRY, P., BARANIUK, R., FLANDRIN, P., RIEDI R. and VEITCH, D. The Multiscale Nature of Network Trafic: Discovery, Analysis and Modelling. *Signal Processing Magazine, IEEE*, 19(3):28–46, 2002.

[2] ABRY, P., FLANDRIN, P., TAQQU, M.S. and VEITCH, D. Wavelets for the analysis, estimation and synthesis of scaling data. *Self Similar Network Trafic Analysis and Performance Evaluation*, K. Park and W. Willinger, Eds., Wiley, 2000.

[3] ABRY, P., JAFFARD, S. and LASHERMES, B. Revisiting Scaling, Multifractal, and Multiplicative Cascades with the Wavelet Leader Lens. *SPIE*, 2004.

[4] ADLER, R. *The Geometry of Random Fields*. John Wiley & Sons, New York, 1981.

[5] ANDERSON, T.W. and DARLING, D.A. A test of goodness-of-fit. *J. Amer. Statist. Assoc.*, 49:765–769, 1954.

[6] ARBEITER, M.A. Random recursive constructions of self-similar fractal measures. the noncompact case. *Prob. Th. Rel. Fields*, 88:497–520, 1991.

[7] ARNEODO, A., ARGOUL, F., BACRY, E., ELEZGARAY, J. and MUZY, J.F. *Ondelettes, multifractales et turbulences. De l'ADN aux croissances cristallines*. Diderot, Paris, 1995, 1995.

[8] ARNEODO, A., BACRY, E. and MUZY, J. The thermodynamics of fractals revisited with wavelets. *Physica A*, 213:232–275, 1995.

[9] ATHREYA, K.B. A note on a functional equation arising in Galton-Watson processes. *J. Appl. Probab.*, 8:589–598, 1971.

[10] ATHREYA, K.B. and NEY, P.E. *Branching Processes*. Springer-Verlag, 1972.

[11] BACHELIER, L. Théorie de la spéculation. *Ann. Sci. de l'Ecole Norm. Sup.*, (17):21–86, 1900.

[12] BACRY, E., MUZY, J. and ARNEODO, A. Singularity spectrum of fractal signals from wavelet analysis: exact results. *Journal of Statistical Physics*, 70(314), 1993.

[13] BARLOW, M.T. and PERKINS, E.A. Brownian Motion on Sierpinski Gasket. *Probab. Th. Rel. Fields*, 79:543–623, 1988.

[14] BARNSLEY, M.F. *Fractals everywhere.* Academic Press, 1988.

[15] BARNSLEY, M.F. and DEMKO, S. Iterated function systems and the global construction of fractals. *Proc. Roy. Soc. London*, A399:243–275, 1985.

[16] BARNSLEY, M.F., ELTON, J.H. and HARDIN, D.P. Recurrent iterated function systems. *Constr. Approx.*, B5:3–31, 1989.

[17] BARNSLEY, M.F., ERVIN, V., HARDIN, D. and LANCASTER, J. Solution of an inverse problem for fractals and other sets. *Proc. of the National Academy of Science*, 83:1975–1977, 1986.

[18] BARNSLEY, M.F., SAUPE, D. and VRSCAY, E.R. *Fractals in Multimedia.* The IMA volumes in Mathematics and its applications. Vol 132, 2001.

[19] BARRAL, J. and MANDELBROT, B.B. Multiplicative products of cylindrical pulses. *Probab. Theory Relat. Fields*, 124:409–430, 2002.

[20] BERAN, J. *Statistics for Long-Memory Processes.* Chapman and Hall, New York, 1994.

[21] BIENAYME, I.J. De la loi de multiplication et de la durée des familles. *Soc. Philomath.Paris Extraits*, Ser. 5, 37-9, 1845.

[22] BIGGINS, J.D. Growth Rates in the Branching Random Walk. *Z. Wahrsch. verw. Gebeite*, 48:17–34, 1979.

[23] BIGGINS, J.D. How fast does a General Branching Random Walk spread? *Classical and Modern Branching processes,(K.B. Athreya, P. Jagers, eds.). IMA Volumes in Mathematics and its Applications (1996)*, 84:19–40, Springer-Verlag, New York, 1996.

[24] BIGGINS, J.D. and BINGHAM, N.H. Large Deviations in the Supercritical Branching Process. *Adv. Appl. Prob.*, 25:757–772, 1993.

[25] BIGGINS, J.D. and KAPRIANOU, A.E. Measure change in multitype branching. *Ann. Appl. Probab.*, 36:544–581, 2004.

[26] BIGGINS, J.D. and KYPRIANOU, A.E. Branching Random Walk: Senata-Heyde norming. *Ann. Prob.*, 25:337–360, 1997.

[27] BINGHAM , N.H. and DONEY, R.A. Asymptotic properties of supercritical branching processes I: The Galton Watson process. *Adv. Appl. Prob.*, 6:711–731, 1974.

[28] BROWN, R. A brief account of microscopical observations made in the months of june, july and august, 1827, on the particles contained in the pollen of plants; and on the general existence of active molecules in organic and inorganic bodies. *Philosophical Magazine N. S.*, (4):161–173, 1828.

[29] BUCKLEW, J.A. *Large deviations techiniques in decision, simulation and estimation*. John Wiley & Sons, New York, 1990.

[30] BURD, G.A. and WAYMIRE, E.C. Independent random cascades on galton-watson trees. *Proc. Amer. Math. Soc.*, 128(9):2753–2761, 2000.

[31] BURRILL, C.W. *Measure, Integration and Probability*. Mc Graw-Hill, 1972.

[32] CHAINAIS, P., RIEDI R. and ABRY, P. On non scale invariant infinitely divisible cascades. *IEEE Trans. on Info. Theory*, 51(3), 2005.

[33] CRAMER, H. Sur un nouveau théorème limite de la théorie des probabilités. *Actualités Scientifiques et Industrielles*, (736):5–23, 1938.

[34] D'AGOSTINO, R.B. and STEPHENS, M.A. *Goodness-of-fit techniques*. New York: M. Dekker, 1986.

[35] DAOUDI, K. *Lois d'échelle, fractales et ondelettes*, volume 2, chapter Généralisations des systèmes de fonctions itérées: analyse de la régularité locale et modélisation multifractale des signaux, pages 129–163. Hermès Lavoisier, 2002.

[36] DECROUEZ, G., AMBLARD, P-O. and BROSSIER, J-M. A wavelet analysis of random iterated function systems. *7th Conference on Mathematics in Signal Processing*, 2006.

[37] DECROUEZ, G., AMBLARD, P-O., BROSSIER, J-M. and JONES, O.D. Galton watson fractal signals. *IEEE ICASSP, Hawaii, USA*, III:1157–1160, 2007.

[38] DORAN, M. *Conversations with Cezanne*. University California Press, 2001.

[39] EINSTEIN, A. On the movement of small particles suspended in a sationary liquid demanded by the molecular-kinetic theory of heat. *Ann. Physik*, 17, 1905.

[40] ELLIS, R.S. Large deviations for a general class of random vectors. *Ann of Prob*, 12(1):1–12, 1984.

[41] ELLIS, R.S. *Entropy, Large deviations and Statistical Mechanics*. Springer-Verlag, New York, 1985.

[42] ELTON, J. An ergodic theorem for iterated maps. *Ergodic Theory Dynamical Systems*, 7:481–488, 1987.

[43] FALCONER, K. Random fractals. *Math. Proc. Camb. Phil. Soc.*, 100:559–582, 1986.

[44] FALCONER, K.J. *Techniques in Fractal Geometry*. John Wiley, 1997.

[45] FALCONER, K.J. *Fractal Geometry: Mathematical Foundations and Applications*. John Wiley, 2003.

[46] FORTE, B. and VRSCAY, E.R. Solving the inverse problem for function and image approximation using iterated function systems. *Dynamics of Continuous, Discrete and Impulsive Systems*, 1(2):177–231, 1995.

[47] GALTON, F. and WATSON, H.W. On the probability of the extinction of families. *J. Anthrop. Soc. London (Royal Anthropol. Inst. G. B. Ireland)*, 4:138–144, 1875.

[48] GRAF, S. Statistically self-similar fractals. *Prob. Th. Rel. Fields*, 74:357–392, 1987.

[49] GRAF, S., MAULDIN, R.D. and WILLIAMS, S.C. The exact Hausdorff dimension in random recursive constructions. *Mem. Amer. Math. Soc.*, 71(381), 1988.

[50] GROSSMAN, A. and MORLET, J. Decomposition of hardy functions into square integrable wavelets of constant shape. *SIAM J. Math. Anal.*, 15(4):723–736, 1984.

[51] GÄRTNER, J. On large deviations from the invariant measure. *Theor. Prob. Applic.*, 22(1):24–39, 1977.

[52] HAMBLY, B.M. On constant tail behaviour for the lmiting random variable in a supercritical branching process. *J. Appl. Prob.*, 32(1):267–273, 1995.

[53] HAMBLY, B.M. and JONES, O.D. Thick and Thin Points for Random Recursive Fractals. *Adv. Appl. Prob.*, 35:251–277, 2003.

[54] HARRIS, T.E. *The Theory of Branching Processes*. Springer, Berlin, 1963.

[55] HAWKES, J. Trees Generated by a Simple Branching Process. *J. London Math. Soc.*, 24(2):373–384, 1981.

[56] HEYDE, C.C. A risky asset model with strong dependence through fractal activity time. *J. Appl. Prob.*, 36:1234–1239, 1999.

[57] HEYDE, C.C. and SENATA, E. Studies in the History of Probability and Statistics. XXXI. The Simple Branching Process, a Turning Point Test and a Fundamental Inequality: A Historical Note on I.J. Bienaymé. *Biometrika*, 59(3):680–683, 1972.

[58] HURST, S.R. and PLATEN, E. The marginal distribution of returns and volatility. *L_1-Statistical Procedures and Related Topics (IMS Lecture Notes Monogr. Ser. 31), ed. Y. Dodge, IMS, Hayward, CA*, pages 301–314, 1997.

[59] HURST, S.R., PLATEN, E. and RACHEV, S.R. Subordinated Markov Models: a comparison. *Finan. Eng. Japanese Markets*, 4, 1997.

[60] HUTCHINSON, J.E. Fractals and self-similarity. *Indiana Univ. Math. J. 30*, pages 713–747, 1981.

[61] HUTCHINSON, J.E. and RÜSCHENDORFF, L. Random fractal measures via the contraction method. *Indiana Univ. Math. J.*, 47(2):471–487, 1998.

[62] HUTCHINSON, J.E. and RÜSCHENDORFF, L. Random Fractals and Probility Metrics. *Adv. in Appl. Probab.*, 32:925–947, 2000.

[63] HUTCHINSON, J.E. and RÜSCHENDORFF,L. *Fractal geometry and stochastics*, chapter Self-Similar fractals and selfsimilar random fractals, pages 109–123. ed. C. Bandt, S. Graf, M. Zähle, Birkhauser, 2000.

[64] JACQUIN, A. Image coding based on a fractal theory of iterated contractive image transformations. *IEEE Trans. Image Proc.*, 1:18–30, 1992.

[65] JAFFARD, S. Multifractal formalism for functions, part ii: Self-similar functions. *SIAM J. of Math. Anal.*, 28(4):971–998, 1997.

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

[67] JAGERS, P. *Branching Processes with Biological Applications*. London: Wiley, 1975.

[68] JONES, O.D. Homepage of Owen Dafydd Jones, Accessed 01/10/08, www.ms.unimelb.edu.au/∼odj.

[69] JONES, O.D. Fast, efficient on-line simulation of self-similar processes. *Thinking in Patterns: Fractals and Related Phenomena in Nature*, pages 165–175, M.M. Novak Ed. World Scientific 2004.

[70] JONES, O.D. and ROLLS, D.A. A characterization of and hypothesis test for continuous local martingales. *To appear*.

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

[72] JONES, O.D. and SHEN, Y. 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*, pages 219–234, 2005.

[73] KESTEN, H. and STIGUM, B.P. A Limit Theorem for Multidimentional Galton Watson Processes. *Ann Math Statist*, 37(3):1211–1233, 1966.

[74] KIMMEL, M. and AXELROD, D. *Branching Processes in Biology.* Interdisciplinary Applied Mathematics, Springer, 2002.

[75] KOLMOGOROV, A.N. Sulla determinazione empirica di una legge di distibuziane. *Giorna. Ist. Attuari*, 4:83–91, 1933.

[76] KURTZ, T., LYONS, R., PEMANTLE, R. and PERES, Y. A conceptual proof of the Kesten-Stigum theorem for multi-type branching processes. *Classical and modern branching processes. IMA Vol. Math. Appl., Springer New York*, 84:181–185, 1997.

[77] LAMPERTI, J. Semi-stable stochastic processes. *Trans. of the Am. Math. Society*, 104, 1962.

[78] LIU, Q. The growth of an entire characteristic function and the tail probability of the limit of a tree martingale. *In Trees, Progress in Probability*, 40:51–80, 1996. Birkhäuser: Verlag Basel, Eds.B.Chauvin, S.Cohen, A.Rouault.

[79] LIU, Q. On Generalized Multiplicative Cascades. *Stochastic Processes and Their Applications*, 86:263–286, 2000.

[80] LIU, Q. Local Dimensions of the Branching Measure on a Galton-Watson Tree. *Ann. Inst. H. Poincaré, Probabilités et Statistiques*, 37(2):195–222, 2001.

[81] LYONS, R., PEMANTLE, R. and PERES, Y. Ergodic theory on Galton-Watson trees, Speed of random walk and dimension of harmonic measure. *Ergodic Theory Dynamical Systems*, 15:593–619, 1995.

[82] MALLAT, S. A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Transactions on Patern Analysis and Machine Intelligence.*, 11(7):674–693, July 1989.

[83] MALLAT, S. and HWANG, W.L. Singularity detection and processing with wavelets. *IEEE Trans. in Info. Theory*, 38(2), March 1992.

[84] MANDELBROT, B.B. Intermittent turbulence in self-similar cascades: divergence of high moments and dimension of the carrier. *J. of Fluids Mechanics*, 62(2):331–358, 1974.

[85] MANDELBROT, B.B. *The Fractal Geometry of Nature*. W. H. Freeman and Company, San Francisco, 1982.

[86] MANDELBROT, B.B. *Fractals and Scaling in Finance: Discontinuity, Concentration, Risk*. Springer-Verlag, New York, 1997.

[87] MANDELBROT, B.B. A multifractal walk down wall street. *Scientific American*, 280:70–73, 1999.

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

[89] MANDELBROT, B.B. and VAN NESS, J.W. Fractional Brownian motions, Fractional noises and applications. *SIAM review*, 10(4), 1968.

[90] MAULDIN, R.D. and WILLIAMS, S.C. Random recursive constructions : asymptotic geometric and topological properties. *Trans. Amer. Math. Soc.*, 295:325–346, 1986.

[91] MEYER, Y. Orthonormal wavelets. In Wavelets. J.M. Combes, A. Grossmann, and Ph. Tchamitchian Eds. New York. Springer-Verlag. pages 21–37, 1989.

[92] MOLCHAN, G.M. Scaling exponents and multifractal dimensions for independent random cascades. *Comm. Math. Phys.*, 179:681, 1996.

[93] MORAN, P.A.P. Additive functions of intervals and Hausdorff measure. *Proc. Camb. Phil. Soc.*, 42:15–23, 1946.

[94] MÖRTERS, P. and SHIEH, N-R. On the Multifractal Spectrum of the Branching Measure on a Galton-Watson Tree. *J. Appl. Prob*, 41:1223–1229, 2004.

[95] NEVEU, J. Arbres et processus de galton-watson. *Annales de l'Institut Henri Poincaré. Probabilités et Statistiques*, 22(2):199–207, 1986.

[96] OKIKIOLU, G.O. *Aspects of the Theory of Bounded Integral Operators in L^p spaces*. Academin Press Inc. (London), 1971.

[97] OLSEN, R. Random geometrically graph directed self-similar multifractals. Pitman Research Notes. 307, Longman (1994).

[98] OSSIANDER, M. and WAYMIRE, E.C. Statistical estimation for muliplicative cascades. *The Annals of Statistics*, 28(6):1533–1560, 2000.

[99] PAKES, A.G. Extreme Order Statistics on Galton-Watson Trees. *Metrika*, 47:95–117, 1998.

[100] PEYRIÈRE, J. Calculs de dimensions de Hausdorff. *Duke Mathematical Journal.*, 44(3):591–601, 1977.

[101] RIEDI, R.H. An improved multifractal formalism and self-similar measures. *Journal of Mathematical Analysis and Applications*, 189:462–490, 1995.

[102] RIEDI, R.H. *Multifractal processes*, pages 625–716. In Theory And Applications Of Long-Range Dependence, P. Doukhan, G. Oppenheim, and M. S. Taqqu, editors, 2003.

[103] RIEDI, R.H. and MANDELBROT, B.B. Inversion Formula for Continuous Multifractals. *Advences in Applied Mathematics*, 19:332–354, 1997.

[104] RIEDI, R.H. and MANDELBROT, B.B. Exceptions to the multifractal formalism for discontinuous measures. *Mathematics Proceedings Cambr. Phil. Society*, 123:133–157, 1998.

[105] RIEDI, R.H., CROUSE, M.S., RIBEIRO, V.J. and BARANIUK, R.G. A multifractal wavelet model with applications to network traffic. *IEEE Transactions on Information Theory*, 45(3):992–1018, April 1999.

[106] ROCKAFELLAR, R.T. *Convex Analysis.* Princeton University Press, Princeton, 1970.

[107] ROGERS, L.C.G. and WILLIAMS, D. *Diffusions, Markov Processes, and Martingales. Volume One: Foundations. Second Edition.* Wiley, 1994.

[108] SALA, N. Fractal Geometry in the Arts: An Overview Across the Different Cultures. *Thinking in Patterns: Fractals and Related Phenomena in Nature*, pages 177–188, 2004.

[109] SENATA, E. Functional Equations and the Galton-Watson Process. *Adv. in Applied Probability.*, 1(1):1–42, 1969.

[110] SHIEH, N-R. and TAYLOR, S.J. Multifractal Spectra of Branching Measure on a Galton-Watson Tree. *J. Appl. Prob.*, 39:100–111, 2002.

[111] SHIRYAYEV, A.N. *Probability.* Springer Verlag, New York, 1984.

[112] VRSCAY, E.R. Moment and colage methods for the inverse problem of fractal construction with iterated function systems. *Proc. of the First IFIP Conference on Fractals, Amsterdam*, 1990.

[113] WENDT, H. Contributions à l'analyse multifractale des coefficients d'ondelettes dominants et du bootstrap: Images, performances d'estimation, nombre de moments nuls et structure de dépendance. Intervalles de confiance et tests d'hypothèse. *PhD thesis*, 2008.

[114] WENDT, H. and ABRY, P. Multifractality Tests using Bootstrapped Wavelet Leaders. *IEEE Trans. on Signal Processing*, 55(10):4811–4820, 2007.

[115] WHITT, W. *Stochastic-Process Limits. An Introduction to Stochastic-Process Limits and Their Application to Queues.* Springer, 2002.

[116] WOOD, A.T.A. and CHAN, G. Simulation of Stationary Gaussian Processes in $[0,1]^d$. *Journal of Computational and Graphical Statistics*, 3(4):409–432, 1994.

[117] YATES, A., CHAN, C., STRID, J., MOON, S., CALLARD, R., GEORGE, A. and STARK, J. Reconstruction of cell population dynamics using CFSE. *BMC Bioinformatics.*, 8:196, 2007.

Résumé étendu en français

Dans le cadre de ma cotutelle internationale de thèse entre le Gipsa-Lab (Grenoble) et le Département de Mathématiques et de Statistiques de l'Université de Melbourne (Australie), je présente un résumé étendu de mes résultats de thèse en français, en accord avec les conventions de cotutelle de thèse signées par les deux établissements. Une introduction plus détaillée ainsi que les preuves des résultats énoncés ici peuvent être trouvées dans les chapitres précédents, en anglais. Mes apports personnels sont surlignés en gras tout au long de ce résumé. Les numéros des théorèmes, corollaires, lemmes et hypothèses donnés dans ce résumé renvoient aux résultats énoncés dans les chapitres précédents. Ils ne suivent donc pas un nécessairement ordre croissant ici.

Les objets fractals, considérés comme des "monstres mathématiques" au milieu du XIXème siècle, n'obtiennent un statut à part entière que depuis les travaux de Benoît Mandelbrot dans les années 70. Depuis, la géométrie fractale est reconnue dans de nombreux domaines en sciences, comme en témoigne le nombre exponentiel de publications depuis ces 20 dernières années. En traitement du signal, l'invariance d'échelle est une propriété qui a été largement observée en pratique, dans des domaines aussi divers que l'étude de la turbulence développée, du télétrafic informatique ou des signaux biomédicaux. Les signaux présentant une invariance d'échelle sont extrêmement irréguliers : une partie du signal possède la même information statistique que le signal d'origine, à un facteur de renormalisation près. L'absence d'échelle caractéristique rend leur analyse impossible à l'aide de quantités classiques en traitement du signal (périodes, fréquences, ...). Cependant, il est possible d'analyser leur régularité locale à $t = t_0$ à l'aide de l'exposant de Hölder $H(t_0)$. Cet exposant compare le signal d'origine $X(t)$ à des fonctions polynomiales. $X(t)$ appartient à l'ensemble $C_{t_0}^{h(t_0)}$ s'il existe un polynôme P_{t_0} de degré au plus égal à la partie entière de $h(t_0)$ tel que

$$|X(t) - P_{t_0}(t)| \leqslant K|t - t_0|^{h(t_0)}$$

dans un voisinage de t_0. La plus grande valeur H de $h(t_0)$ telle que $X \in C_{t_0}^{h(t_0)}$ est l'exposant de Hölder de X à $t = t_0$ [102]. Un signal qui possède un unique coefficient $h(t)$ pour tout t est qualifié de monofractal, par opposition aux signaux multifractals,

pour lesquels h varie de manière erratique avec l'échelle des temps. Dans le cas des signaux multifractals, il est impossible d'estimer en pratique la valeur de l'exposant local de Hölder en chaque point du signal (caractère discret des signaux, précision finie des données, ...). On s'intéresse alors au spectre multifractal D, qui à un exposant local donné H_0 détermine la dimension de Hausdorff $D_{\mathcal{H}}$ de l'ensemble des points possédant comme régularité H_0

$$D(H_0) = D_{\mathcal{H}}\{t \mid h(t) = H_0\}.$$

Pour les processus monofractals, $D(h)$ se réduit à un seul point $h = H_0$, $D(H_0) = 1$ et par convention, $D(h) = -\infty$ pour $h \neq H_0$. Différentes techniques s'intéressent aux procédures d'estimation du spectre multifractal. Citons par exemple la méthode du Maximum du Module de la Transformée en Ondelettes (MMTO) [7, 12] ou la technique des coefficients dominants [66].

Dans cette thèse, je m'intéresse à l'élaboration de deux nouveaux modèles pour la génération de signaux (multi)fractals. Leur point commun est leur structure de branchement sous jacente. Un processus à branchements est par définition un ensemble de particules qui vivent pendant un temps aléatoire et qui peuvent donner naissance à un nombre aléatoire de fils jusqu'au moment de leur mort. Le plus simple des processus à branchements est le processus de Galton-Watson, que l'on décrit de la manière suivante. Considérons un unique ancêtre vivant pendant exactement une unité de temps et donnant naissance à un nombre aléatoire Z_1 de fils au moment de sa mort. Soit p la distribution de cette variable aléatoire. Chaque fils de la première génération se comporte alors comme la particule initiale, et indépendamment les uns des autres. Ils vivent exactement une unité de temps et donnent naissance à un nombre aléatoire de fils (formant la deuxième génération) au moment de leur mort, selon la distribution p. Et ainsi de suite. Ce processus peut être décrit mathématiquement à l'aide d'un index de temps discret, donnant la taille de la population Z_n à $n = 0, 1, 2, \ldots$. Les processus à branchement sont de bons modèles mathématiques pour l'étude de la démographie des populations et possèdent de nombreuses applications en biologie [67, 74, 117]. L'application qui nous intéresse ici concerne la génération de signaux fractals à l'aide de processus de Galton-Watson. En 1986, Falconer, Mauldin et Williams se sont penchés sur l'existence d'ensembles compacts fractals obtenus à l'aide d'une structure aléatoire de Galton-Watson [43, 90]. Ici nous proposons deux nouveaux modèles. Le premier est une généralisation de la construction des Systèmes de Fonctions Itérés (IFS). Le deuxième concerne l'ensemble des signaux dont l'arbre de branchement ('crossing tree') est un processus de Galton-Watson. Une construction de tels signaux est proposée dans cette thèse. Nous les nommons Multifractal Embedded Branching Processes (MEBP). Dans les deux sous-sections qui suivent, j'introduis les notions mathématiques nécessaires à la compréhension de mes résultats de thèse et je présente un résumé de mes travaux. J'invite le lecteur à se réferrer aux chapitres précédents en anglais pour les preuves de mes résultats.

Systèmes de Fonctions Itérés de Galton-Watson.

Les Systèmes de Fonctions Itérés (IFS) procurent un moyen simple de générer des ensembles fractals. La terminologie 'IFS' a été introduite par Barnsley et Demko [15] mais c'est Hutchinson en 1981 qui a le premier démontré l'existence d'ensembles et de mesures déterministes auto-similaires en utilisant des opérateurs contractants [60]. Le formalisme, d'abord introduit pour la génération d'ensembles compacts et de mesures, a par la suite été adapté au cas des signaux. Nous présentons Figure 5.3 quatre ensembles fractals obtenus à l'aide de cette procédure.

Soit $L_p(\mathbb{X})$ la classe des signaux p-intégrables $\mathbb{X} \to \mathbb{R}$, $1 < p < \infty$, où $\mathbb{X}$ est un sous ensemble compact de $\mathbb{R}$. $||\cdot||_p$ dénote la norme classique dans $L_p(\mathbb{X})$: $||f||_p = (\int |f|^p d\mu)^{1/p}$ où μ est la mesure de Lebesgue. On munit $L_p(\mathbb{X})$ d'une métrique $d_p(f,g) = ||f-g||_p$ où f et g sont dans L_p. Le théorème de Riesz-Fisher assure que (L_p, d_p) est complet [96]. On considère dans la suite $\mathbb{X} = [0,1]$.

Un IFS consiste en l'application récursive d'un opérateur T. Soit une fonction initiale f_0, on dénote par $T^n f_0$ le n-ième itéré de T sur f_0. Pour une classe d'opérateurs T, l'IFS converge dans $L_p(\mathbb{X})$ vers une fonction f^*

$$T^n f_0 \to f^* \text{ lorsque } n \to +\infty. \tag{5.3}$$

f^* est l'unique fonction satisfaisant la relation $f = Tf$. On dit que f^* satisfait la loi d'échelle T [63]. f^* est communément appelé point fixe ou attracteur de T. On suppose généralement que T se décompose en un ensemble de M opérateurs non-linéaires plus simples $\phi_i : \mathbb{R} \times \mathbb{X} \to \mathbb{R}$ pour $1 \leqslant i \leqslant M$. Chaque ϕ_i déforme le signal d'origine et le place dans un sous-intervalle $\mathbb{X}_i = \varrho_i(\mathbb{X})$ de $\mathbb{X}$

$$(Tf)(x) = \sum_{j=1}^{M} \phi_j[f(\varrho_j^{-1}(x)), \varrho_j^{-1}(x))]\mathbf{1}_{\varrho_j(\mathbb{X})}(x) \tag{5.4}$$

où $\{\varrho_i(\mathbb{X})\}_{i=1}^M$ partitionne $\mathbb{X}$. $\mathbf{1}_{\varrho_i(\mathbb{X})}$ est la fonction indicatrice de l'interval $\varrho_i(\mathbb{X})$. ϕ_i sont des fonctions à deux variables. L'arbre de construction sous jacent est un arbre déterministe M-aire: l'image de f par T est construite à l'aide de M copies de f. La construction de $T^2 f$, image de f par le deuxième itéré de T, requiert M copies de Tf, soit M^2 copies de f, et ainsi de suite. Au n-ième itéré de T on associe alors un arbre déterministe M-aire contenant n niveaux de profondeur.

Des conditions sous lesquelles l'IFS converge sont données explicitement dans [63] pour des fonctions $\phi_i : \mathbb{R} \to \mathbb{R}$ à une variable. Le résultat se généralise facilement pour des fonctions $\phi_i : \mathbb{R} \times \mathbb{X} \to \mathbb{R}$ à deux variables, comme ci-dessous. Ce résultat est prouvé plus loin dans un cadre plus général (Théorème 4, Chapitre 2).

THÉORÈME 2. *Soient ϱ_i des fonctions contractantes avec facteur de contraction $r_i < 1$ pour $i = 1, \ldots, M$, et ϕ_i Lipschitz en leur première variable, avec constante de Lipschitz s_i. Si pour un p, $\lambda_p = \sum_{i=1}^{M} r_i s_i^p < 1$ et $\sum_{i=1}^{M} r_i \int |\phi_i(0,x)|^p dx < \infty$, alors T possède un unique point fixe dans $L_p(\mathbb{X})$.*

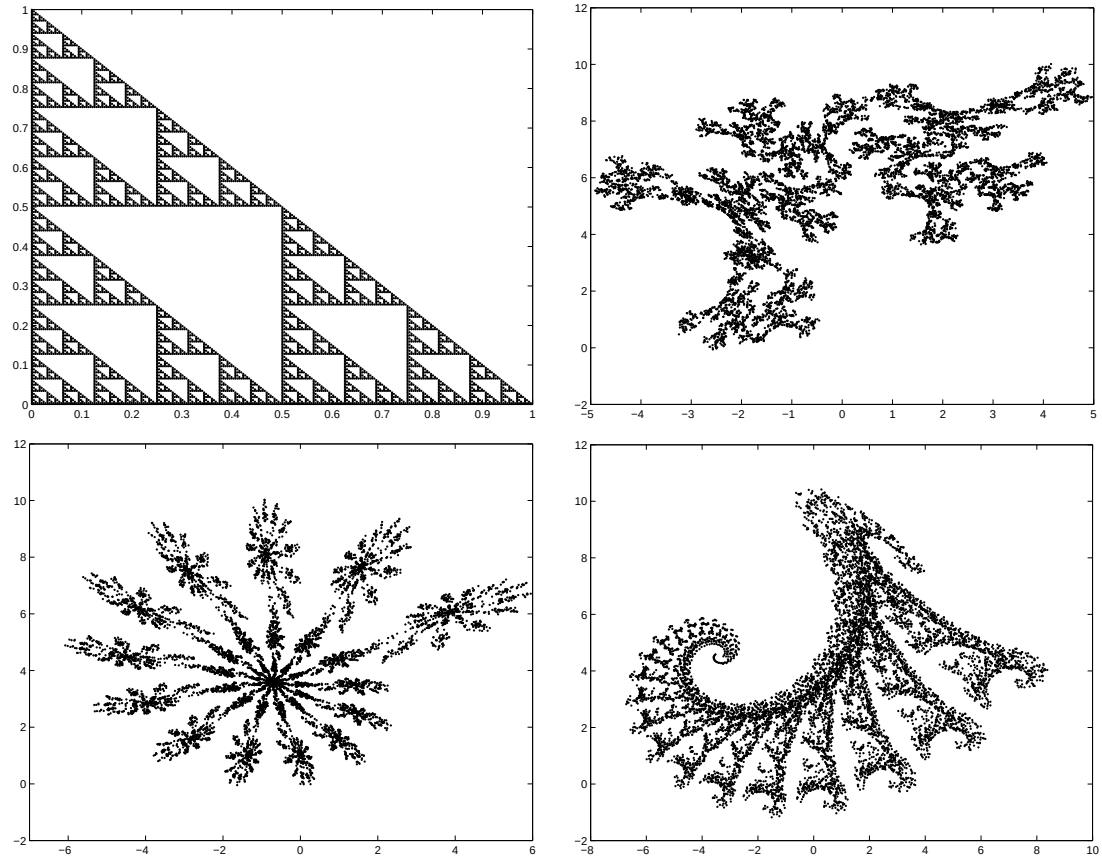


Figure 5.3: Ensembles fractals.

Nous proposons de généraliser ce résultat en introduisant de l'aléa dans la définition de l'opérateur. Comme nous venons de le voir, le point fixe généré possède un arbre de construction déterministe M-aire. En reconduisant une étude semblable à celle de Hutchinson et Ruschendorff, **nous donnons des conditions sous lesquelles il existe un unique point fixe lorsque l'arbre de construction est aléatoire, de type Galton-Watson. Nous autorisons également une structure aléatoire pour T.** Il faut alors définir de nouveaux espaces de travail.

Soit $(\Sigma, \mathcal{F}, P)$ un espace probabilisé. On munit L_p d'une σ-algèbre $\mathcal{L}_p$ ([107], définition 25.2). Un processus aléatoire p-intégrable est une variable aléatoire $f : \Sigma \to L_p(\mathbb{X})$. On définit

$$\mathbb{L}_p = \{f : \Sigma \to L_p(\mathbb{X}), f \text{ mesurable} \mid \mathbb{E}\big[||f||_p^p\big] < +\infty\}$$

où $\mathbb{E}$ dénote l'espérance mathématique sous P. $f(x) : \Sigma \to \mathbb{R}$ est la variable aléatoire obtenue en évaluant f en x. On montre alors que $||f||_p^* = \mathbb{E}^{\frac{1}{p}}\big[||f||_p^p\big]$ est une norme sur $\mathbb{L}_p$ et que la métrique définie par $d_p^*(f,g) = ||f-g||_p^*$ rend l'espace $(\mathbb{L}_p, d_p^*)$ complet. Muni de cet espace métrique complet, on définit alors un opérateur aléatoire T de la manière suivante

$$(Tf)(x) = \sum_{j=1}^{Z} \phi_j[f^{(j)}(\varrho_j^{-1}(x)), \varrho_j^{-1}(x))]\mathbf{1}_{\varrho_j(\mathbb{X})}(x) \tag{5.5}$$

où $(Z, \phi_1, \varrho_1, \ldots, \phi_Z, \varrho_Z)$ est une variable aléatoire et $f^{(j)}$ sont des copies i.i.d. de f. Les ϱ_j partitionnent de manière aléatoire $\mathbb{X}$ en Z sous-intervalles. Le facteur de contraction de ϱ_j est la variable aléatoire r_j. ϕ_j sont des fonctions de deux variables, Lipschitz en leur première variable, avec coefficient aléatoire de Lipschitz s_j. Z est distribué selon le vecteur de probabilités $\mathbf{q} = (q_1, q_2, \ldots)$ où q_i est la probabilité associée à l'événement $\{Z = i\}$. Sauf mention du contraire (Théorème 5), on autorise un nombre infini de fils. L'arbre de construction sous-jacent possède alors un nombre aléatoire de fils (Z) à chaque noeud, Z étant indépendant d'un noeud à l'autre de l'arbre. Il est donc de type Galton-Watson. On se propose de déterminer des conditions sous lesquelles il existe une unique fonction f^* satisfaisant $f^* = Tf^*$ dans $\mathbb{L}_p$ pour cette définition de T. Pour ce faire, il est indispensable de se munir de l'espace probabilisé des arbres étendus de Galton-Watson $(\Sigma, \mathcal{F}, P) = (K, \mathcal{K}, \kappa)$, dont la construction est détaillée plus loin dans le manuscrit. Un élément $k \in K$ est une réalisation d'un arbre de Galton-Watson, dont chaque branche est munie d'un couple d'opérateur réalisés $\{\varrho_j, \phi_j\}$. **Le théorème qui suit est le résultat principal de l'étude menée au Chapitre 2 de cette thèse.**

THÉORÈME 4. *Soit $(K, \mathcal{K}, \kappa)$ l'espace des arbres étendus de Galton-Watson. On définit $\mathbb{L}_p$ avec $(\Sigma, \mathcal{F}, P) = (K, \mathcal{K}, \kappa)$. On suppose $\mathbb{E}\sum_{j=1}^{Z} r_j \int |\phi_j(0,x)|^p dx < +\infty$ pour un $1 < p < +\infty$ et $\lambda_p = \mathbb{E}\sum_{j=1}^{Z} r_j s_j^p < 1$, où $\mathbb{E}$ dénote l'espérance sous κ. Alors il existe une unique fonction f^* satisfaisant $f^* = Tf^*$ dans $\mathbb{L}_p$. En outre, pour tout*

$f_0 \in \mathbb{L}_p(\mathbb{X})$,

$$d_p^*(T^n f_0, f^*) \leqslant \frac{\lambda_p^{n/p}}{1-\lambda_p^{1/p}} d_p^*(f_0, Tf_0) \tag{5.6}$$

qui tend vers 0 *lorsque* $n \to +\infty$. *De plus, la distribution de* f^* *est l'unique distribution satisfaisant* $f^* \stackrel{d}{=} Tf^*$, *où* $\stackrel{d}{=}$ *dénote l'égalité des distributions.*

Nous illustrons ce théorème à l'aide de la Figure 5.4, où l'on présente une réalisation du point fixe d'un IFS de Galton-Watson. Nous représentons également une estimation du signal moyen obtenu à l'aide de 100 réalisations indépendentes du point fixe.

Nous proposons ensuite de déterminer quelques propriétés du point fixe. On commence par énoncer le résultat concernant la continuité des réalisations du point fixe.

PROPOSITION 1. $\mathbb{X} = [a, b]$. *Soit* α *l'unique point fixe aléatoire de* $\phi_1(., a)$ *et* β *celui de* $\phi_Z(., b)$: $\phi_1(\alpha, a) = \alpha$ *et* $\phi_Z(\beta, b) = \beta$. *On suppose que* α *et* β *sont les même pour toutes les réalisations de* ϕ_1 *et* ϕ_Z. *Si* $\phi_i(\beta, b) = \phi_{i+1}(\alpha, a)$ *p.s. pour tout* $i \in \{1, \ldots, Z-1\}$ *et si les opérateurs considérés sont continus, alors les traces de* f^* *sont continues,* $f^*(a) = \alpha$ *et* $f^*(b) = \beta$ *p.s.*

Nous étudions dans un deuxième temps le comportement des moments du point fixe de l'IFS en fonction du vecteur de probabilités q. On considère le modèle présenté dans [37], où l'arbre est aléatoire et les opérateurs déterministes. On considère l'ensemble des opérateurs déterministes

$$\{\{\phi_{k,1}, \ldots, \phi_{k,k}, \varrho_{k,1}, \ldots, \varrho_{k,k}\}\}_{k=1,2,\ldots}.$$

Pour $Z = j$, on applique $\{\phi_{j,1}, \ldots, \phi_{j,j}, \varrho_{j,1}, \ldots, \varrho_{j,j}\}$. $\phi_{k,j}$ et $\varrho_{k,j}$ peuvent avoir différentes expressions pour différentes valeurs de k, $j = 1, \ldots, k$. L'opérateur T devient:

$$(Tf)(x) = \sum_{j=1}^{Z} \phi_{Z,j}[f^{(j)}(\varrho_{Z,j}^{-1}(x)), \varrho_{Z,j}^{-1}(x))]\mathbf{1}_{\varrho_{Z,j}(\mathbb{X})}(x). \tag{5.7}$$

Le facteur de Lipschitz de $\phi_{Z,j}$ est $s_{\mathbb{Z},j}$ et le facteur de contraction de $\varrho_{Z,j}$ est $r_{Z,j}$. Puisque les opérateurs attachés aux branches de l'arbre sont identiques pour un nombre de fils donné, à une réalisation de l'arbre correspond une et une seule réalisation du point fixe.

THÉORÈME 5. *On suppose vérifiées les hypothèses du Théorème 4. Soit* $f^* \in \mathbb{L}_p$ *le point fixe de l'IFS de Galton-Watson possédant un nombre borné de fils et des opérateurs déterministes de la forme* $\phi(u, v) = su + \zeta(v)$, *où* $0 \leqslant s < 1$ *et* ζ *est une fonction non-linéaire. On suppose que* $\lambda_r = \mathbb{E}\sum_{j=1}^{Z} r_{Z,j} s_{Z,j}^r < 1$ *pour* $r = 1, \ldots, p$. *Alors le* r*-ième moment de* f^* *varie continuement avec le vecteur de probabilités* $\mathbf{q}$,

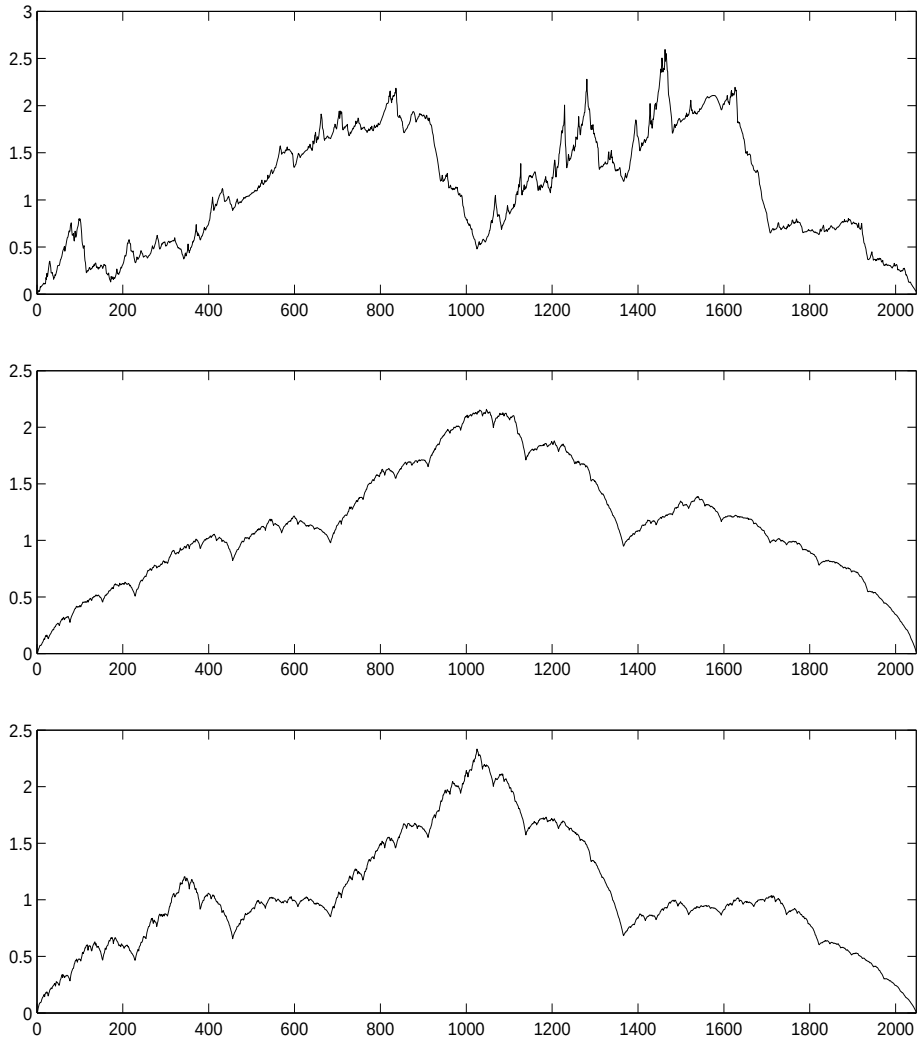


Figure 5.4: Une réalisation de l'attracteur d'un IFS de Galton-Watson (haut) avec une estimation de sa moyenne (milieu). La figure du bas est la moyenne du point fixe d'un autre IFS de Galton-Watson. Les paramètres de ces IFS sont décrits dans le texte en anglais.

pour $r = 1, \ldots, p$.

Enfin, **nous testons de manière empirique le caractère multifractal des points fixes des IFS de Galton-Watson à l'aide de la technique des coefficients dominants.** Pour un IFS donné, le spectre obtenu est non trivial, indiquant une structure riche du point fixe. Les simulations proposées indiquent l'existence d'une classe d'IFS pour lesquels le point fixe semble multifractal et motive une étude future afin de déterminer le spectre théorique de ces signaux.

Multifractal Embedded Branching Processes (MEBP).

Dans la deuxième partie de cette thèse (Chapitres 3 et 4), **nous proposons la construction d'un nouveau modèle de signaux multifractals, que nous appellons Multifractal Embedded Branching Processes (MEBP)**. Les processus MEBP sont construits à l'aide de leur arbre de branchement, de type Galton-Watson, dont la construction est donnée plus loin. Pour tout processus de branchement adéquat, il existe une famille de processus monofractals possédant une invariance d'échelle discrète pour lesquels c'est l'arbre de branchement. On identifie l'un d'entre eux comme étant le processus Canonical Embedded Branching Process (CEBP). Le processus MEBP est alors construit à partir du processus CEBP à l'aide d'un changement de temps multifractal, construit à l'aide d'une cascade multiplicative définie sur l'arbre de branchement. **On prouve que le processus CEBP est monofractal et on obtient une borne supérieure pour le spectre multifractal d'un processus MEBP particulier. Aussi, on donne un algorithme de simulation efficace pour la génération de processus MEBP.** Les mouvements Browniens en temps multifractals introduits par Mandelbrot [86, 87] sont un cas particulier des processus MEBP, laissant imaginer de potentielles applications en finance. Nous expliquons maintenant la construction du processus CEBP, du processus MEBP et donnons les principaux résultats dont les preuves sont données dans les chapitres 3 et 4.

Soit $X : \mathbb{R}^+ \to \mathbb{R}$ un processus continu, avec $X(0) = 0$. Pour $n \in \mathbb{Z}$ on définit les temps de passage de niveau n (correspondant à l'échelle 2^n), dénotés T_k^n, par $T_0^n = 0$ et

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

Soit C_k^n le k-ième passage de niveau n. On entend par passage une partie du processus X de T_{k-1}^n à T_k^n avec comme information supplémentaire le temps du début de passage T_{k-1}^n et la position du processus $X(T_{k-1}^n)$ en début de passage. Un passage est de type haut si $X(T_k^n) = X(T_{k-1}^n)+2^n$ ou bas si $X(T_k^n) = X(T_{k-1}^n)-2^n$. En allant d'une échelle grossière à une échelle plus fine, on décompose chaque C_k^n en une suite de passages de niveau $n-1$. En associant chaque passage à un noeud d'un arbre et aux sous-passages les fils d'un noeud, on définit alors l'arbre de branchement du processus X, comme illustré Figure 5.5.

On représente le nombre de sous-passages à l'échelle 2^{n-1} remplaçant le k-ième passage de niveau n à l'aide de la variable aléatoire Z_k^n. Les Z_k^n sous-passages consistent de $(Z_k^n - 2)/2$ excursions suivies d'un passage direct. Une excursion est une paire haut-bas ou bas-haut; un passage direct est une paire haut-haut ou bas-bas, comme illustré Figure 5.6: le premier passage de niveau n en pointillés est remplacé par 6 sous passages de niveau $n-1$, contenant 2 excursions haut-bas et d'un passage direct haut-haut. Le deuxième passage de taille 2^n est lui remplacé par 4 sous passages, contenant une excursion bas-haut et un passage direct haut-haut.

Un processus continu X est appelé Embedded Branching Process (EBP) si les variables aléatoires Z_k^n sont indépendantes et identiquement distribuées.

On adopte les notations suivantes pour la définition de l'arbre de branchement. Soit $\emptyset$ la racine de l'arbre, représentant un unique passage de niveau 0. On dénote

Figure 5.5: Construction de l'arbre de branchement.

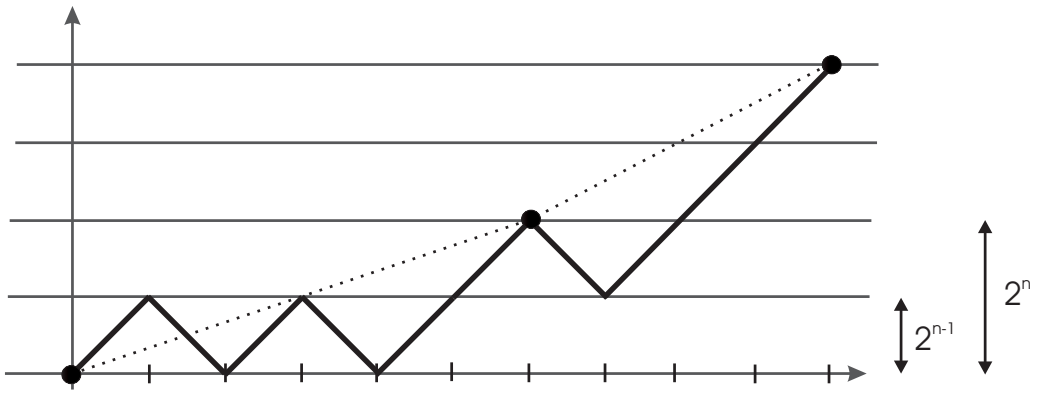


Figure 5.6: Construction du processus CEBP.

la première génération par i, $1 \leqslant i \leqslant Z_\emptyset$, où $Z_\emptyset$ représente le nombre de fils de $\emptyset$. La deuxième génération est alors dénotée par ij, $1 \leqslant j \leqslant Z_i$, et ainsi de suite. $Z_\emptyset$ est distribué selon $P(Z_\emptyset = j) = p(j)$. Plus généralement, un nœud est un élément de $U = \cup_{n \geqslant 0} \mathbb{N}^{*n}$ et une branche un couple $(\mathbf{u}, \mathbf{u}j)$ où $\mathbf{u} \in U$ et $j \in \mathbb{N}^*$. La longueur d'un nœud $\mathbf{i} = i_1 \dots i_n$ est $|\mathbf{i}| = n$. Si $|\mathbf{i}| > n$, $\mathbf{i}|_n$ est la restriction de $\mathbf{i}$ à ses n premiers termes. Par convention, $|\emptyset| = 0$ et $\mathbf{i}|_0 = \emptyset$.

Un arbre Υ est un ensemble de nœud, c'est à dire un sous-ensemble de U, tel que

- $\emptyset \in \Upsilon$
- Si un nœud $\mathbf{i}$ appartient à l'arbre, alors chaque restriction $\mathbf{i}|_k$, $k \leqslant |\mathbf{i}|$ appartient également à l'arbre.
- Si $\mathbf{u} \in \Upsilon$, alors $\mathbf{u}j \in \Upsilon$ pour $j = 1, \dots, Z_\mathbf{u}$ et $\mathbf{u}j \notin \Upsilon$ for $j > Z_\mathbf{u}$, où $Z_\mathbf{u}$ est le nombre de fils de $\mathbf{u}$.

Soit $F(s) = \sum_{j=0}^{\infty} p(j) s^j$ la fonction génératrice de $Z_\emptyset$, définie pour s complexe tel que $|s| \leqslant 1$. Soit Υ_n^{GW} la n-ième génération de l'arbre, c'est à dire l'ensemble des nœuds de longueur n et Z_n^{GW} son cardinal et $\mu = \sum_x xp(x)$ la taille moyenne de la première génération. Alors $\mu^{-n} Z_n^{GW}$ est une martingale non-négative et converge presque sûrement vers une limite $W_\emptyset$ [10]. Soit $\Lambda(s) = \mathbb{E}(e^{-sW_\emptyset})$ la transformée de Laplace de W définie pour s complexe tel que $\mathrm{Re}(s) \geqslant 0$. ψ satisfait l'équation fonctionnelle de Poincaré [109]

$$\Lambda(\mu s) = F(\Lambda(s)). \tag{5.8}$$

On considère ensuite $\Upsilon_\mathbf{i} = \{\mathbf{j} \in \Upsilon \mid |\mathbf{j}| \geqslant |\mathbf{i}| \text{ et } \mathbf{j}|_{|\mathbf{i}|} = \mathbf{i}\}$. La frontière de l'arbre est donnée par $\partial\Upsilon = \{\mathbf{i} \in \mathbb{N}^\mathbb{N} \mid \forall n \geqslant 0,\ \mathbf{i}|_n \in \Upsilon\}$.

Soit $(\Omega, \mathcal{F}, \mathbb{P})$ l'espace des arbres aléatoires marqués, que l'on peut construire selon la même procédure que les arbres étendus de Galton-Watson $(K, \mathcal{K}, \kappa)$ de la section précédente. Un arbre marqué est un arbre donc chaque branche est munie d'une variable aléatoire. Pour un processus EBP donné, soit $p(x) = \mathbb{P}(Z_k^n = x)$ la distribution du nombre de ses fils. Si p est tel que $p(2) < 1$ et $\sum_x x \log(x) p(x) < \infty$ on dit que la distribution est régulière. On remarque que Z_k^n prend ses valeurs dans

$2\mathbb{N}^*$ puisque les sous-passages viennent par paires. Soit α_k^n un vecteur de taille Z_k^n dont les composantes définissent les types des Z_k^n sous-passages de C_k^n. Chacunes des $(Z_k^n - 2)$ entrées viennent par paires, chaque paire étant 'haut-bas' ou 'bas-haut'. Les deux dernières composantes sont soit 'haut-haut' ou 'bas-bas'. On définit alors la distribution $p_{c|z}$ des orientations de la manière suivante: $p_{c|z} = \mathbb{P}(\alpha_k^n = \cdot \mid Z_k^n = z)$.

THÉORÈME 6. *Pour toute distribution régulière p sur $2\mathbb{N}^*$ il existe un unique processus EBP continu X sur $[0, T_1^0]$ tel que*

- *Les excursions sont distribuées selon $p_{c|z}$.*
- *La longueur du passage de niveau n est distribuée comme $\mu^{-n} W_\emptyset$ où la transformée de Laplace de $W_\emptyset$ satisfait (5.8)*

Alors on appelle l'unique processus X le processus EBP Canonique (CEBP). Soit $\mu = \sum_x xp(x)$ la taille moyenne de la première génération. Soit $H = \log 2 / \log \mu$, alors pour tout $a = \mu^n$, $n \in \mathbb{Z}$ et $t \in [0, T_1^0]$,

$$X(t) \stackrel{fdd}{=} a^{-H} X(at) \text{ pour les distributions de dimension finie.} \tag{5.9}$$

Nous illustrons Figure 5.7 trois réalisations de processus CEBP. La distribution du nombre de fils est géométrique avec paramètres 0.3, 0.5 et 0.7 et les excursions sont i.i.d. et de type 'haut-bas' et 'bas-haut' avec probabilité 1/2.

Par ailleurs, **on propose l'extension de la construction du processus CEBP à tout interval compact de** $\mathbb{R}$. Pour ce faire, on considère la suite $\{X^{(n)}\}$ de processus CEBP et on prouve la convergence en distribution de $\{X^{(n)}\}$ vers un processus limite X. Soit $X^{(0)}$ un processus CEBP construit comme dans Théorème 6. Soit $\mathcal{C}_T$ l'espace des fonctions continues à support compact $[0, T]$ dans $\mathbb{R}^+$. On définit la suite de processus CEBP $\{X^{(n)}\}$ de la manière suivante

$$X^{(n)}(t) = 2^n X^{(0)}(\mu^{-n} t).$$

On montre alors que

COROLLAIRE 2. *$X^{(n)}$ converge en distribution vers un processus limite X dans $\mathcal{C}_T$.*

Il est clair d'après la relation (5.9) que les processus CEBP possèdent une invariance d'échelle discrète. **On prouve par ailleurs que ce sont des processus monofractals,** comme l'indique le résultat suivant.

HYPOTHÈSE 4. *Pour tout $p > 0$, $\mathbb{E} Z_\emptyset^p < \infty$.*

THÉORÈME 10. *Soit X un processus CEBP avec distribution régulière p et taille moyenne μ. Suppose que le nombre de fils est borné, c'est-à-dire qu'il existe M tel que $p(x) = 0$ pour tout $x \geqslant M$. On suppose vraie l'Hypothèse 4. Alors X est monofractal: avec probabilité 1, pour tout t,*

$$h(t) = \frac{\log 2}{\log \mu}. \tag{5.10}$$

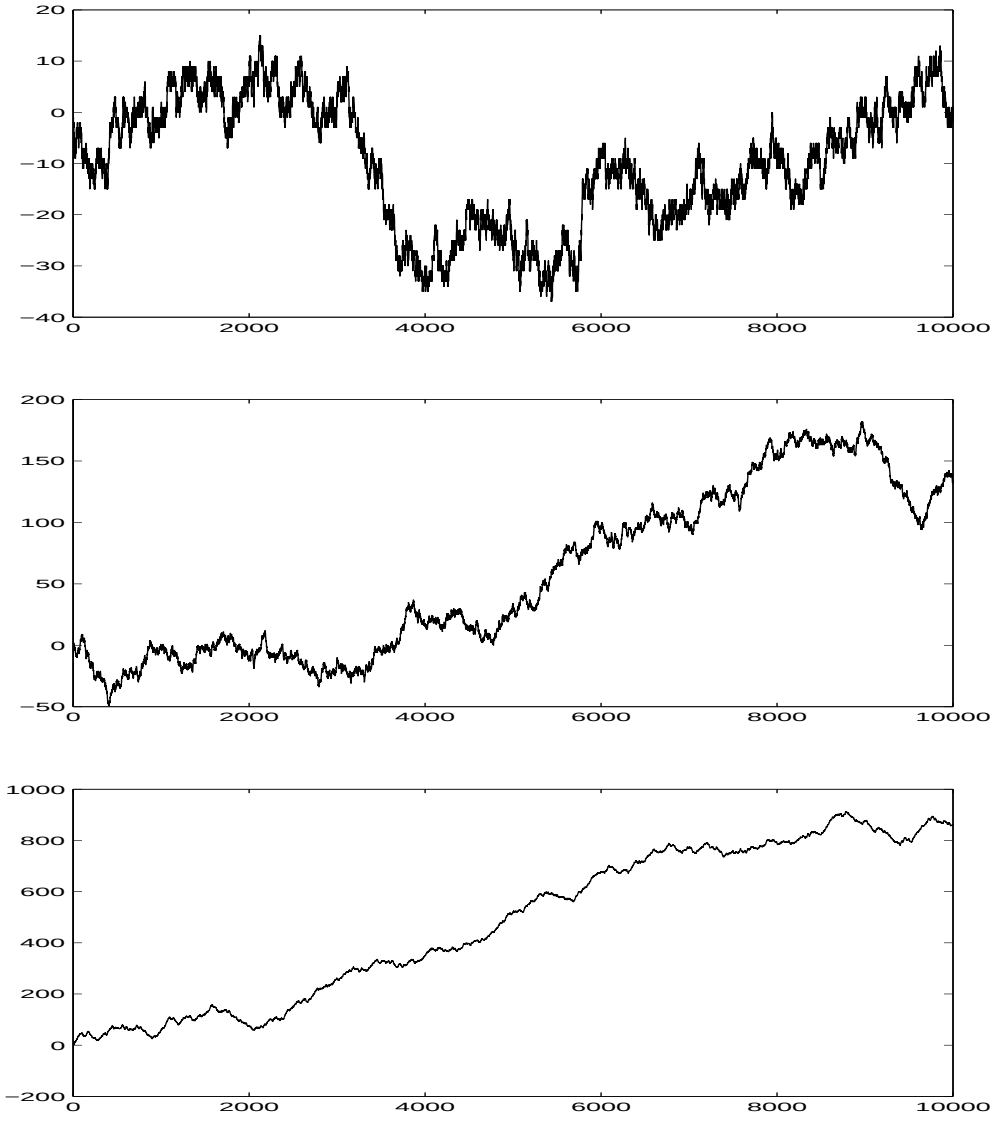


Figure 5.7: De haut en bas. Réalisations de processus CEBP. La distribution du nombre de fils est géométrique avec paramètres 0.3, 0.5 et 0.7.

On est maintenant en position d'introduire les processus MEBP, définis comme des processus CEBP en temps multifractal. Le changement de temps est défini à partir d'un processus en cascade sur l'arbre de branchement. A chaque branche $(\mathbf{i}, \mathbf{i}j)$ de l'arbre, on associe une variable aléatoire (ou poids) $\rho_j(\mathbf{i})$. Les $\rho_1(\mathbf{i}), \ldots, \rho_{Z_\mathbf{i}}(\mathbf{i})$ peuvent être dépendantes et dépendre de $Z_\mathbf{i}$, mais sont indépendantes des autres nœuds de l'arbre. Le poids associé au nœud $\mathbf{i}$ est alors

$$\rho_\mathbf{i} = \prod_{k=1}^{|\mathbf{i}|} \rho_{i_k}(\mathbf{i}|_{k-1}).$$

$\rho_\mathbf{i}$ est le produit de tous les poids le long d'une ligne de descente, de la racine jusqu'à $\mathbf{i}$. $\rho_\mathbf{i}$ nous sert à définir les temps de passage du nouveau processus. En effet, on utilise ces poids pour définir une mesure, ν, sur la frontière $\partial\Upsilon$ de l'arbre de branchement. A partir de cette mesure (à support discret), on construit une deuxième mesure, ζ, dont le support est un intervalle de $\mathbb{R}$. On définit alors le changement de temps $\mathcal{M}$ par $\mathcal{M}(t) = \zeta([0,t))$. Le processus MEBP est alors donné par $Y = X \circ \mathcal{M}^{-1}$, où X est le processus CEBP. Les processus X et Y possèdent donc la même structure, mais ont des temps de passage différents. Il existe des restrictions sur les poids associés aux branches de l'arbre afin que le changement de temps soit continu. **Nous pouvons maintenant énoncer le premier résultat du Chapitre 3 sur l'existence et la continuité des processus MEBP.**

Hypothèse 1.

$$\rho_i > 0, \quad \mathbb{E}\sum_{i=1}^{Z_\emptyset} \rho_i(\emptyset) = 1, \quad 0 > \mathbb{E}\sum_{i=1}^{Z_\emptyset} \rho_i(\emptyset)\log\rho_i(\emptyset) > -\infty$$
$$\text{et } \mathbb{E}\sum_{i=1}^{Z_\emptyset} \rho_i(\emptyset)\log\sum_{i=1}^{Z_\emptyset} \rho_i(\emptyset) < \infty.$$

Hypothèse 2. *Il existe $\epsilon > 0$ tel que $\mathbb{E}Z_\emptyset^{1+\epsilon} < \infty$*

Théorème 8. *Soit p une distribution régulière sur $2\mathbb{N}^*$ et X le processus CEBP associé. Si les hypothèses 1 et 2 sont vérifiées, alors on peut construire un processus croissant $\mathcal{M}$ tel que $\mathcal{M}^{-1}$ soit continu, et ainsi définir pour tout $t \in \mathbb{R}$*

$$Y = X \circ \mathcal{M}^{-1}.$$

Le processus Y est appelé processus MEBP.

Nous présentons Figure 5.8 une réalisation d'un processus CEBP et une réalisation du processus MEBP, obtenu à l'aide d'un changement de temps multifractal défini sur l'arbre de branchement.

Hypothèse 3. *On suppose que*

$$\rho_i \in (0,1],\ \mathbb{P}(\sum_{i=1}^{Z_\emptyset} \rho_i(\emptyset) = 1) < 1 \ \textit{et}$$
$$\textit{pour tout } p > 1,\ \mathbb{E}\Big[\Big(\sum_{i=1}^{Z_\emptyset} \rho_i(\emptyset)\Big)^p\Big] < \infty.$$

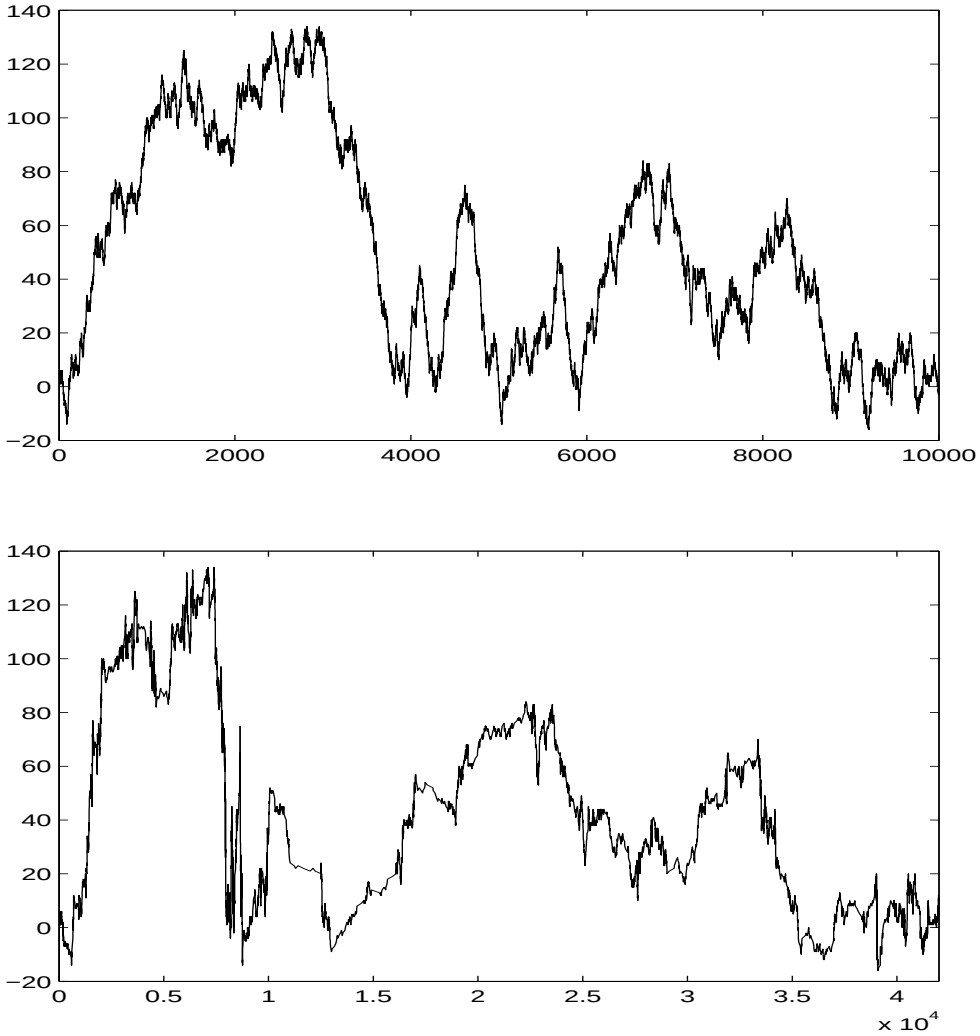


Figure 5.8: Figure du haut: processus EBP. La distribution du nombre de fils est géométrique. Figure du bas: processus MEBP obtenu à l'aide d'un changement de temps multifractal du processus CEBP. Le changement de temps est défini sur l'arbre de branchement, à l'aide de poids distribués selon une loi Gamma.

COROLLAIRE 3. *Sous les hypothèses 1, 2 et 3, Y et $\mathcal{M}$ sont continus.*

Définir le changement de temps à l'aide de l'arbre de branchement nous permet de simuler efficacement les processus MEBP. **Nous proposons un algorithme de simulation dans la troisième section du Chapitre 3.** Pour ce faire, nous utilisons une représentation de Markov de dimension finie du processus. En considérant une restriction de l'espace d'état, nous représentons $X(n)$ à l'aide d'un vecteur de taille $O(\log n)$, ce qui nous permet de simuler $X(n+1)$ à partir de $X(n)$ en $O(\log n)$ opérations.

Dans la dernière section du Chapitre 3, on propose d'imiter un mouvement Brownien fractionnaire à l'aide d'un processus MEBP.

Soit $\mathbf{i} \in \Upsilon_{n-1}$ un noeud de la génération $n-1$ de l'arbre de branchement et $\mathbf{i}k$ son k-ième fils. Soit $D^n_{\mathbf{i}k}$ la durée du passage associé avec le noeud $\mathbf{i}k$. On rappelle que $Z_{\mathbf{i}}$ correspond au nombre de fils du noeud $\mathbf{i}$. On considère alors la durée moyenne des sous-passages pour un fBm après conditionnement sur la taille du passage parent:

$$f(z) := \alpha \mathbb{E}(D^n_{\mathbf{i}k} \mid Z_{\mathbf{i}} = z) \tag{5.11}$$

où α est tel que $\sum z f(z) \mathbb{P}(Z = z) = 1$. Le rôle du facteur de renormalisation α est expliqué après. On estime alors $f(z)$, noté $\hat{f}_{\text{fBm}}(z)$, pour un fBm avec exposant $H = 0.7$. $\hat{f}_{\text{fBm}}(z)$ est obtenu en moyennant 80 estimations de $f(z)$. On se sert alors de ce graphe pour déterminer la valeur des poids à attacher à l'arbre de branchement du processus MEBP:

$$\rho \mid (Z = z) = \hat{f}_{\text{fBm}}(z).$$

Les poids sont déterministes conditionnellement à Z. Avec cette définition, le facteur de renomalisation α nous assure que $\mathbb{E} \sum_{i=1}^{Z} \rho_i = 1$, hypothèse fondamentale pour l'existence des processus MEBP (Hypothèse 1). On montre alors que la durée moyenne des sous passages du processus MEBP est

$$\hat{f}_{\text{fBm}}(z) \left[\mathbb{E}\rho\right]^n . \tag{5.12}$$

Les deux équations (5.12) et (5.11) sont alors identiques.

La Figure 5.9 présente l'estimation de $f(z)$ pour le fBm et pour le processus MEBP associé. Les deux courbes suivent la même tendance, surtout pour les petites valeurs de z. La figure du milieu présente la distribution marginale du processus MEBP ainsi défini, que l'on compare à une loi normale. Nous avons également réalisé deux tests statistiques (Kolmogorov-Smirnov et Anderson-Darling) afin de décider si les données proviennent d'une loi normale. Les résultats obtenus indiquent que oui, avec un niveau de confiance de 95%. Enfin nous présentons sur la figure du bas la structure de corrélation du processus MEBP en échelle logarithmique avec celle théorique d'un fBm. On observe une légère déviation par rapport à la pente théorique $2H - 2$. Un autre test, validant la structure monofractale du processus MEBP ici défini a été réalisé. Nous ne le présentons pas dans ce résumé.

En conclusion, l'ensemble des tests réalisés indiquent une bonne approximation d'un fBm par un processus MEBP.

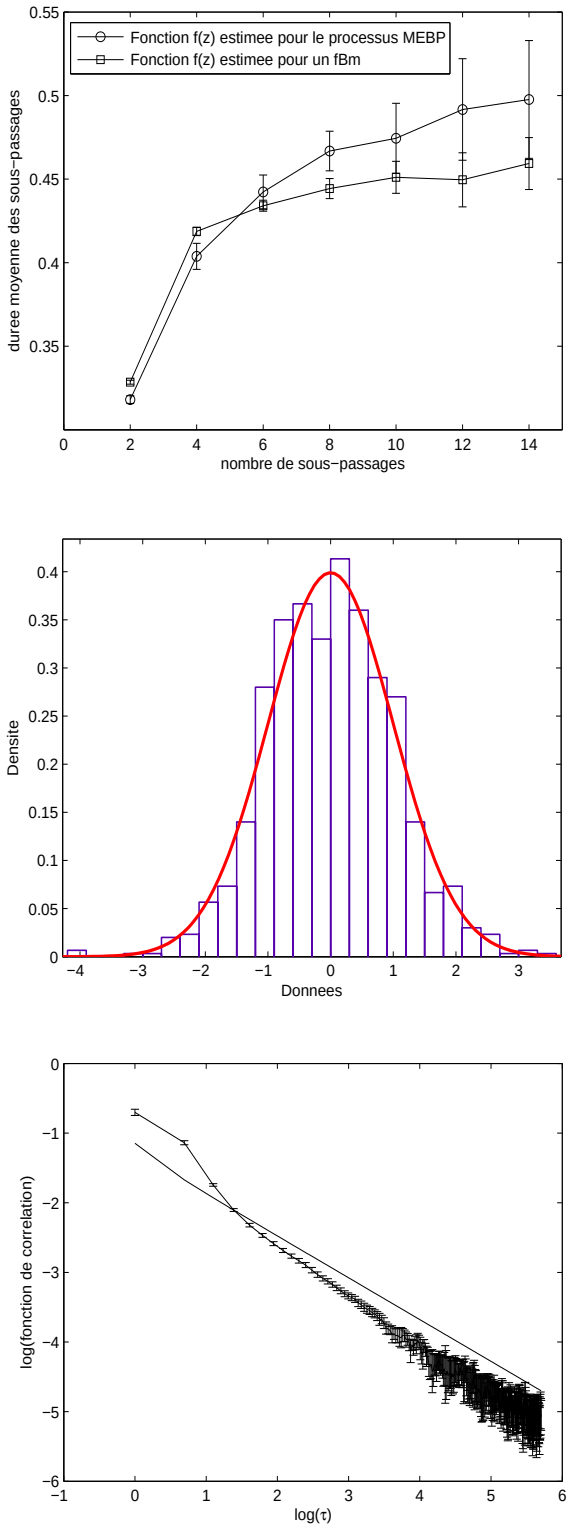
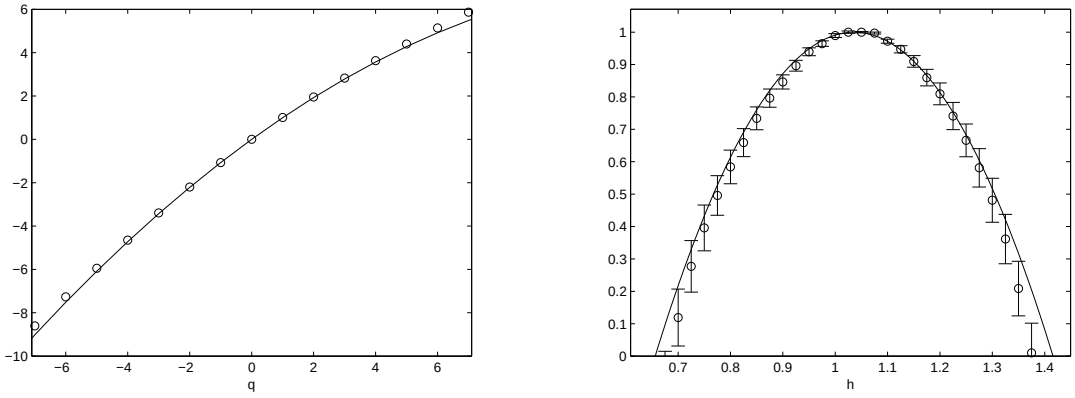


Figure 5.9: La figure du haut présente $\hat{f}_{\mathrm{fBm}}(z)$ et $\hat{f}_{\mathrm{MEBP}}(z)$, estimations de $f(z)$ pour un fBm et pour le processus MEBP associé. Sur la figure du milieu, nous présentons la distribution marginale du processus MEBP, que l'on compare à une loi normale. La figure du bas présente la fonction de corrélation du processus MEBP avec la corrélation théorique d'un fBm avec $H = 0.7$.



(a) (b)

Figure 5.10: **Estimation du spectre de $\mathcal{M}$.** Figure (a): Fonction de partition ('o') estimée à l'aide de la méthode des coefficients dominants. La courbe en trait plein représente $1 + \tilde{T}(q)$. Figure (b): Transformée de Legendre-Fenchel de la fonction de partition. Des intervalles de confiance à 95% sont également représentés.

Dans le Chapitre 4, on s'intéresse au spectre multifractal de $\mathcal{M}$ et de Y. **En suivant une méthode développée par Riedi [102] pour les cascades définies sur des arbres déterministes, nous obtenons une borne supérieure pour le spectre de $\mathcal{M}$ sous certaines conditions.**

Corollaire 6. *Soit p une distribution régulière. On suppose que le nombre de fils à chaque nœud est borné, que les poids ρ sont i.i.d. et que les hypothèses 1, 2, 3 et 4 sont vérifiées. Soit $\tilde{T}(q) = -1 - \log_\mu \mathbb{E}\rho^q$. Alors, le spectre multifractal $D(a)$ de $\mathcal{M}$ est borné par*

$$D(a) \leqslant \inf_{q \in \mathbb{R}} (qa - \tilde{T}(q)). \tag{5.13}$$

Nous avons effectué des simulations pour estimer le spectre de $\mathcal{M}$ à l'aide de la méthode des coefficients dominants. Les résultats obtenus, présentés sur la Figure 5.10, supportent la théorie et semblent indiquer que la borne supérieure proposée est en fait le spectre lui-même, sous l'hypothèse que le spectre de $\mathcal{M}$ est concave. En effet, la fonction de partition $\zeta(q)$ estimée à l'aide des coefficients dominants se superpose avec la courbe théorique $1 + \tilde{T}(q)$. Sous l'hypothèse que le spectre est concave, la transformée de Legendre-Fenchel de $\zeta(q)$ (et, par la même occasion, de $\tilde{T}(q)$) donne le spectre de Hausdorff du processus [66].

Nous n'avons pas poursuivi l'étude théorique afin de démontrer cette égalité, qui reviendrait à déterminer une borne inférieure, problème théorique reconnu comme étant difficile en général.

On utilise ensuite le résultat du Théorème 6 pour l'obtention d'une borne supérieure du spectre de Y dans un cas particulier. En effet, lorsque la distribution p est géométrique avec paramètre 0.5, le processus CEBP est réduit à un mouvement Brownien B classique. Dans ce cas, il est possible d'établir une relation

entre le spectre de $\mathcal{M}$ et celui de Y, comme l'indique le théorème suivant.

THÉORÈME 20. *Soit $Y = B \circ \mathcal{M}^{-1}$ le processus MEBP obtenu à l'aide d'un changement de temps du processus CEBP B, qui se réduit à un mouvement Brownien. On suppose les hypothèses 1, 2 et 3 vérifiées. Soient $D_H^{\mathcal{M}}$ et D_H^Y le spectre multifractal de Y et $\mathcal{M}$ respectivement. Alors*

$$D_H^Y(a) = 2aD_H^{\mathcal{M}}(1/2a).$$

La relation précédente nous permet d'obtenir une borne supérieure pour le spectre de Y. **Nous avons également estimé le spectre de Y à l'aide des coefficients dominants.** Les estimations obtenues pour Y vont dans le même sens que celles obtenues pour $\mathcal{M}$ et nous laissent penser une fois de plus à une égalité entre les spectres.

Dans la dernière section du Chapitre 4, nous nous intéressons au spectre de la mesure de branchement ν, dont le support est la bordure de l'arbre $\partial\Upsilon$. Soient $\mathbf{i}$ et $\mathbf{j} \in \partial\Upsilon$. On se doit dans un premier temps de redéfinir la notion d'exposant local. Dans ce but, on équipe la bordure de l'arbre avec la métrique

$$d(\mathbf{i}, \mathbf{j}) = e^{-n(\mathbf{i},\mathbf{j})} \tag{5.14}$$

où $n(\mathbf{i}, \mathbf{j})$ est tel que $\mathbf{i}|_k = \mathbf{j}|_k$ pour tout $k \leqslant n(\mathbf{i}, \mathbf{j})$ et $\mathbf{i}|_k \neq \mathbf{j}|_k$ sinon. Étant donné d, la régularité locale de Hölder de la mesure ν à un point x est définie comme

$$\lim_{r \to 0} \frac{\log \nu[B(x, r)]}{\log r}$$

où $B(x, r)$ est la boule ouverte centrée en x et de rayon r. **On obtient 2 résultats concernant ν. Dans le premier cas, ν est multifractale et on donne la valeur de l'exposant qui concentre toute la masse ν. Dans le deuxième cas, on donne des conditions plus restrictives sur les poids associés aux branches de l'arbre afin que la mesure ν soit monofractale.**

HYPOTHÈSE 7

$$\mathbb{E}\Big[\Big(\sum_{i=1}^{Z_1} \rho_i\Big)\Big(\log\Big(\sum_{i=1}^{Z_1} \rho_i\Big)\Big)^2\Big] < \infty.$$

THÉORÈME 24 *Lorsque les Hypothèses 1 et 7 sont vérifiées, alors avec probabilité 1,*

$$\lim_{n \to \infty} \frac{\nu[B(\mathbf{i}|_n, r)]}{n} = \mathbb{E} \log \rho$$

pour ν-presque tout $\mathbf{i} \in \partial\Upsilon$.

THÉORÈME 25 *Sous les conditions des Hypothèses 1 et 3,*

$$\lim_{n \to \infty} \frac{\nu[B(\mathbf{i}|_n, r)]}{n} = \mathbb{E} \log \rho \tag{5.15}$$

pour tout $\mathbf{i} \in \partial\Upsilon$*, avec probabilité* 1*.*

Ces résultats sont à mettre en relation avec le cas de la mesure de branchement obtenue avec des poids constants égaux à $1/\mu$. Cette mesure a beaucoup été étudiée [55, 79, 80, 94, 110] et les résultats obtenus s'accordent avec ceux de la littérature en remplaçant $\mathbb{E}\log\rho$ par $-\log\mu$.

GENERATION DE SIGNAUX MULTIFRACTALS POSSEDANT UNE STRUCTURE DE BRANCHEMENT SOUS-JACENTE

Résumé

La géométrie fractale, développée par Mandelbrot dans les années 70, a connu un essor considérable ces 20 dernières années. Dans cette thèse, je m'intéresse à la génération de signaux dits fractals et multifractals. J'étudie en particulier 2 modèles, dont leur point commun est leur structure d'arbre de branchement sous jacente. Le premier modèle est une généralisation des Systèmes de Fonctions Itérés ou IFS, introduits par Hutchinson dans les années 80. Les IFS constituent un moyen simple et efficace pour produire des ensembles et des processus fractals en itérant un nombre fixed d'opérateurs. L'idée est d'autoriser un nombre aléatoire d'opérateurs aléatoires à chaque itération de l'algorithme. Nous donnons des conditions simples et faciles à vérifier sous lesquelles l'IFS admet un point fixe. Quelques propriétés du point fixe sont également étudiées. Le deuxième modèle, que nous appellons Multifractal Embedded Branching Process (MEBP), s'obtient à l'aide d'un changement de temps multifractal d'un processus à invariance d'échelle discrète, le processus EBP Canonique (CEBP). Nous donnons un algorithm efficace de simulation "on-line" de ces processus, permettant de générer $X(n+1)$ à partir de $X(n)$ en $O(\log n)$ opérations. Nous obtenons également un borne supérieure pour le spectre multifractal du changement de temps et confirmons les résultats théoriques à l'aide de simulations. Les mouvements Browniens en temps multifractal sont des cas particuliers des processus MEBP, ce qui suggère une application potentielle des processus MEBP en finance. Enfin, nous proposons d'imiter un mouvement Brownien fractionnaire à l'aide d'un processus MEBP.

Mots Clés

Invariance d'échelle, Systèmes de Fonctions Itérés, Formalisme multifractal, Spectre de Hausdorff, Arbre de branchement, Arbre de Galton-Watson, mouvement Brownien fractionnaire.

GENERATION OF MULTIFRACTAL SIGNALS WITH UNDERLYING BRANCHING STRUCTURE

Summary

Fractal geometry, pioneered by Mandelbrot in the 70s, has been recognized in many areas of science. The novelty of this thesis is the generation of fractal and multifractal processes with underlying construction tree. I study two models in particular. The first one is a generalisation of Iterated Function Systems (IFS), introduced by Hutchinson in the early 80s. IFS are an efficient tool to generate fractal sets and functions, by iterating a given set of operators. The idea here is to allow a random number of random operators at each iteration of the algorithm. We derive simple conditions under which the IFS possesses a fixed point. A few properties of the fixed point are also investigated. The second model, called Multifractal Embedded Branching Process (MEBP), is obtained via a multifractal time change of a discrete self-similar process, the Canonical EBP (CEBP). We give an efficient simulation online algorithm which generates $X(n+1)$ from $X(n)$ in $O(\log n)$ steps. We also derive an upper bound of the multifractal spectrum of the time change and we confirm the theoretical results with simulations. Subordinated Brownian motions are particular cases of MEBP processes, which suggests a potential application of MEBP in finance. Finally, we propose to imitate a fractional Brownian motion with an MEBP.

Key words

Self-similarity, Iterated Function Systems, Multifractal formalism, Hausdorff spectrum, Crossing tree, Galton-Watson tree, fractional Brownian motion.