

# Galton-Watson Iterated Function Systems

Geoffrey Decrouez<sup>†‡</sup>, Pierre-Olivier Amblard<sup>†</sup>,  
Jean-Marc Brossier<sup>†</sup> and Owen Jones<sup>‡</sup>

<sup>†</sup> GIPSA-lab/DIS (UMR CNRS 5216, INPG) BP 46, 38402 Saint-Martin-d'Hères  
cedex, France

<sup>‡</sup> Department of Mathematics and Statistics, The University of Melbourne, Parkville,  
Australia

E-mail: bidou.amblard@gipsa-lab.inpg.fr

**Abstract.** Iterated Function Systems (IFS) are interesting parametric models for generating fractal sets and functions. The general idea is to compress, deform and translate a given set or function with a collection of operators and to iterate the procedure. Under weak assumptions, IFS possess a unique fixed point which is in general fractal. IFS were introduced in a deterministic context, then were generalized to the random setting on abstract spaces in the early 80's. Their adaptation to random signals was carried out by Hutchinson and Rüschemdorff [9] by considering random operators. This study extends their model with not only random operators but also a random underlying construction tree. We show that the corresponding IFS converges under certain hypothesis to a unique fractal fixed point. Properties of the fixed point are also described.

Submitted to: *J. Phys. A: Math. Theor.*

PACS numbers: PACS-key : 02.50.-r Probability theory, stochastic processes, and statistics, 05.45.Df Fractals



**Figure 1.** Shadow of a shape drawn in the sand of a Queensland beach at dusk. Credit G. Decrouez.

## 1. Introduction

Signals presenting scale invariance have been widely studied during the past 20 years. The name scale invariance refers to signals not presenting any characteristic scale, each scale playing a similar role. Applications of such signals are wide and range from biology [1] to finance [2], from network traffic [3] to turbulence [4]. Iterated Function Systems (IFS) have received interest in image compression and decompression, where attempts are made to solve the IFS inverse problem: identifying the parameters of an IFS whose attractor is a target image [5]. IFS can also be used for fractal interpolation [6, 7]. The classical IFS considered in the literature are deterministic IFS, where the object (a set, a measure, a function) is transformed by means of deterministic operators. The formalism was first introduced on abstract sets, then adapted to produce fractal measures and functions. Signals obtained from this procedure can be multifractals [8]. Generalization to the random setting has previously been carried out by Hutchinson and Rüschemdorff [9], where only IFS operators are random. This study extends their model by allowing more randomness. As an example of a natural shape that may be modelled as the attractor of a random IFS, figure 1 shows the shadow of a shape drawn in the sand of a Queensland beach.

IFS were first introduced over the space of compact subsets of  $\mathbb{R}^2$ , usually denoted  $\mathcal{H}(\mathbb{R}^2)$ . This space is particularly interesting when dealing with black and white pictures.  $\mathcal{H}(\mathbb{R}^2)$  is generally endowed with the Hausdorff metric  $d_{\mathcal{H}}$ . For  $A$  and  $B$  in  $\mathcal{H}(\mathbb{R}^2)$ ,

$$d_{\mathcal{H}}(A, B) = \max[d(A, B), d(B, A)] \quad (1)$$

where  $d(A, B) = \max[d(x, B), x \in A]$  and  $d(x, A) = \min[d(x, y), y \in A]$  for any  $x \in \mathbb{R}^2$ .

Let  $\omega : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  be a contractive application with contraction ratio  $s$ . Then

$\omega : \mathcal{H}(\mathbb{R}^2) \rightarrow \mathcal{H}(\mathbb{R}^2)$  defined by

$$\omega(B) = \{\omega(x) | x \in B\} \quad \forall B \in \mathcal{H}(\mathbb{R}^2) \quad (2)$$

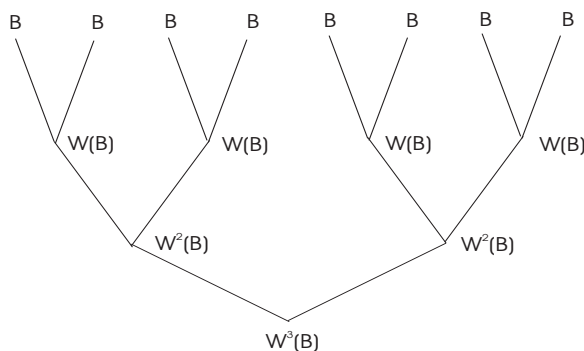
is contractive in the metric space  $(\mathcal{H}(\mathbb{R}^2), d_{\mathcal{H}})$  with contraction constant  $s$ . Now consider a set of  $M$  contractive maps  $\omega_n$  with contraction ratios  $s_n$ ,  $n = 1, \dots, M$ . Then  $W : \mathcal{H}(\mathbb{R}^2) \rightarrow \mathcal{H}(\mathbb{R}^2)$

$$W(B) = \bigcup_{i=1}^M \omega_i(B) \quad (3)$$

is contractive in  $(\mathcal{H}(\mathbb{R}^2), d_{\mathcal{H}})$  with contraction factor  $s = \max_n s_n$  [10]. In other words, the operator  $W$  starts with an initial set  $B$  and compute its image by taking the union of contracted and translated copies of the original set  $B$ . By completeness of the metric space  $(\mathcal{H}(\mathbb{R}^2), d_{\mathcal{H}})$ , it follows from the Banach fixed point theorem that the IFS possesses a unique fixed point  $B^*$ , which satisfies  $W(B^*) = B^*$ . Many well known fractal sets such as the Sierpinski gasket are obtained from this procedure. Such attractors can also produce images close to shapes found in nature, such as the famous example of fern leaves [10]. One can associate a tree with this construction, which is deterministic in the present setting. For IFS with  $M$  maps, the underlying construction tree is an  $M$ -ary tree, where each node possesses exactly  $M$  offspring, as illustrated in figure 2, for  $M = 2$ .

This formalism was adapted to produce fractal measures and signals, first in a deterministic setting. Random IFS were introduced in the 80's on abstract mathematical sets [11, 12]. More recently, Hutchinson and Rüschemdorff have randomized the construction for signals, where operators are randomized, but the deterministic tree structure is retained. In this study we further randomize this model, allowing a random construction tree, or random branching process. Applications of branching processes, which started with the study of demography of populations [13], has been applied to many areas of science and provides good models biology [14]. A good review on branching processes can be found in [15]. Random cascades and measures defined on the boundary of random trees have also been widely studied. See for example the works of Peyrière [16], Hawkes [17], Burd and Waymire [18], Liu [19, 20], Mörters and Shieh [21]. Their work differs from the present study as they consider measures with discrete support. We are interested in this paper in models defined over compact intervals. The novelty of this study is to introduce new models for generating fractal signals using branching processes.

The paper is organized as follows. We first introduce the complete metric space of random signals, where the fixed point of the random IFS lies. We then build the probability space of extended Galton-Watson trees, presented in section 3. Extended trees are Galton-Watson trees whose branches are endowed with a random operator. In section 4, we derive precise conditions under which the IFS possesses a unique fixed point and illustrate the type of signals one can obtain with this new model. In the last section we study various properties of the fixed point, such as conditions for it to be



**Figure 2.** Underlying binary tree associated with a deterministic IFS ( $M = 2$ ).  $W^i(B)$  for  $i = 2, 3$  represents the  $i$ -th iterate of  $W$ . Under mild conditions, convergence to a fixed point occurs when the tree considered is infinite. The fixed point is then at the bottom of the tree (root).

continuous. Furthermore, it is shown in [10] that the fractal attractor of a deterministic IFS continuously depends on the parameters of the IFS. We extend this result and show here, in a special case, that the moments of the fixed point continuously depend on the probability vector of the random variable giving the number of maps used at each iteration of the algorithm. Finally, we give empirical results on the multifractal behaviour on the fixed point.

## 2. Iterated Function Systems on functions

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

### 2.1. Deterministic IFS

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

An IFS consists of recursively applying an operator  $T$  with certain properties. 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  $f^*$

$$T^n f_0 \rightarrow f^* \text{ as } n \rightarrow +\infty \quad (4)$$

in  $L_p(\mathbb{X})$ .  $f^*$  is the unique function satisfying  $f = Tf$ . 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  $\phi_i : \mathbb{R} \times \mathbb{X} \rightarrow \mathbb{R}$  for  $1 \leq i \leq M$ . Each  $\phi_i$  deforms the original signal and maps it to a subinterval  $\mathbb{X}_i = \varrho_i(\mathbb{X})$  of  $\mathbb{X}$ . Specifically,

$$(Tf)(x) = \sum_{i=1}^M \phi_i[f(\varrho_i^{-1}(x)), \varrho_i^{-1}(x)] \mathbf{1}_{\varrho_i(\mathbb{X})}(x) \quad (5)$$

where  $\{\varrho_i(\mathbb{X})\}_{i=1}^M$  partitions  $\mathbb{X}$ .  $\mathbf{1}_{\varrho_i(\mathbb{X})}$  is the indicator function of the interval  $\varrho_i(\mathbb{X})$ . In (5),  $\phi_i$  are functions of two variables. The second variable is however optional and one can define the operator  $T$  with  $\phi_i : \mathbb{R} \rightarrow \mathbb{R}$ . The underlying construction tree is an  $M$ -ary deterministic tree. Conditions of convergence of the IFS are derived explicitly in [9] for non-linear functions  $\phi_i : \mathbb{R} \rightarrow \mathbb{R}$ . The result can be easily generalized to functions  $\phi_i : \mathbb{R} \times \mathbb{X} \rightarrow \mathbb{R}$ , as above.

**Theorem 2.1** *Suppose that  $\varrho_i$  are strict contractions with contraction factors  $r_i < 1$  for  $i = 1, \dots, M$ , and that  $\phi_i$  are Lipschitz in their first variable with Lipschitz constants  $S_i$ , i.e.  $\forall (u_1, u_2, v) \in \mathbb{R}^2 \times \mathbb{X} \quad |\phi_i[u_1, v] - \phi_i[u_2, v]| \leq S_i |u_1 - u_2|$ . 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 4.1, so the proof is not given here. The conditions for convergence are quite weak. The second condition only requires that  $\phi_i$  must be  $p$  integrable with respect to their second variable.

Figure 3 presents attractors of two different IFS, one continuous and one discontinuous. Conditions for continuity are derived in section 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 3. One way to break this pattern is to add randomness to the construction. Therefore section 2.3 defines random IFS with random operators and a random construction tree.

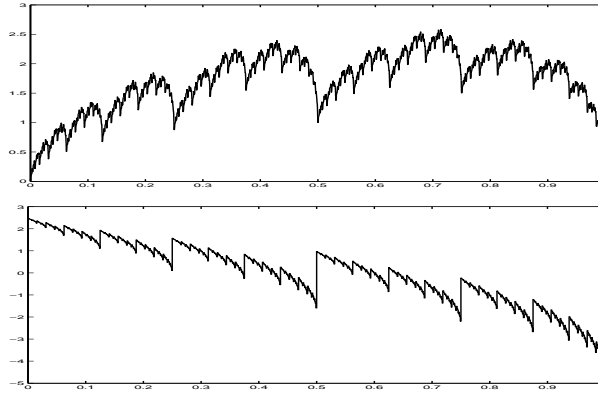
## 2.2. $\mathbb{L}_p$ spaces

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, then a  $p$ -integrable random process is a random variable  $f : \Sigma \rightarrow L_p(\mathbb{X})$ . Define

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

where  $\mathbb{E}$  denotes expectation under  $P$ . We denote by  $f_\sigma$  a realization of the random process  $f \in \mathbb{L}_p$ , where  $\sigma \in \Sigma$ , which will be useful in the proof of Theorem 4.1.  $f(x) : \Sigma \rightarrow \mathbb{R}$  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

$$\|f\|_p^* = \mathbb{E}^{\frac{1}{p}}[\|f\|_p^p] \quad (6)$$



**Figure 3.** 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 bottom 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.

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

**Lemma 2.2**  $\|\cdot\|_p^*$  is a norm on  $\mathbb{L}_p$ ,  $p \geq 1$ .

*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, as applying Fubini by swapping integral and expectation does not lead to the desired result, as illustrated below

$$\|f + g\|_p^{*p} = \mathbb{E} \int |f + g|^p \leq 2^p \int \mathbb{E}(|f|^p + |g|^p) = 2^p \|f\|_p^{*p} + 2^p \|g\|_p^{*p}$$

since for any reals  $x$  and  $y$  one have  $|x + y|^p \leq 2^p(|x|^p + |y|^p)$ ,  $p \geq 1$ .

Integrals defined in this proof are with respect to the Lebesgue measure. We drop the notation  $d\mu$  for simplicity. 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  $ab \leq \frac{a^p}{p} + \frac{b^q}{q}$ . This inequality can be derived using the concavity of  $\log$ . This gives  $\mathbb{E} \int |\bar{f}\bar{g}| \leq \mathbb{E}[\int \frac{1}{p}|\bar{f}|^p + \int \frac{1}{q}|\bar{g}|^q]$  where we define  $\bar{f} = \frac{f}{\|f\|_p^*}$  and  $\bar{g} = \frac{g}{\|g\|_q^*}$ , from which

$$\|fg\|_1^* = \mathbb{E} \int |fg| \leq \|f\|_p^* \|g\|_q^* \quad (7)$$

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  $\mathbb{E} \int |f + g|^{p-1}(|f| + |g|) = \| |f + g|^{p-1} |f| \|_1^* + \| |f + g|^{p-1} |g| \|_1^*$ . Thus, using the previous Hölder's inequality:

$$\|f + g\|_p^{*p} \leq \| |f + g|^{p-1} \|_q^* [\|f\|_p^* + \|g\|_p^*] \quad (8)$$

Since  $pq = p + q$ ,

$$\|f + g\|_q^{*p-1} = \mathbb{E}_q^{\frac{1}{q}} \left( \int |f + g|^p \right) = \|f + g\|_p^{*p-1} \quad (9)$$

Hence:

$$\|f + g\|_p^{*p} \leq \|f + g\|_p^{*p-1} [\|f\|_p^* + \|g\|_p^*] \quad (10)$$

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 2.2 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 [29] to show that  $(\mathbb{L}_p, d_p^*)$  is a Banach space.

### 2.3. Galton-Watson IFS

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

$$(Tf)(x) = \sum_{j=1}^{\nu} \phi_j[f^{(j)}(\varrho_j^{-1}(x)), \varrho_j^{-1}(x)] \mathbf{1}_{\varrho_j(\mathbb{X})}(x) \quad (11)$$

where  $(\nu, \phi_1, \varrho_1, \dots, \phi_\nu, \varrho_\nu)$  is random and  $f^{(j)}$  are i.i.d. copies of  $f$ . The  $\varrho_j$  are affine maps and randomly partition  $\mathbb{X}$  into  $\nu$  subintervals. The contraction factor of  $\varrho_j$  is the random variable  $r_j$ , such that  $0 < r_j < 1$  almost surely.  $\phi_j$  are functions of two variables, Lipschitz in their first variable, with random Lipschitz factor  $S_j$ .  $\nu$  is distributed according to a probability vector  $\mathbf{q} = (q_1, q_2, \dots)$ . The underlying construction tree has therefore a random number of offspring at each node. Assuming that, in this construction, the random variable  $\nu$  is independent and identically distributed from one node to another, the construction tree is of Galton-Watson type [22], hence the name of the IFS.

## 3. Space of extended trees

An ad-hoc structure of  $\Sigma$  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 realization of a Galton-Watson tree whose branches are equipped with realizations of the IFS operators.

**Ionescu-Tulcea [24].** 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  $\kappa_2 : A_1 \times \mathcal{A}_2 \rightarrow [0, 1]$  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  $\kappa_2$  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  $\kappa_3 : A_1 \times A_2 \times \mathcal{A}_3 \rightarrow [0, 1]$  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

$$(\kappa_2 \otimes \kappa_3)(a_1, E) = \int \int \mathbf{1}_E(b, c) \kappa_2(a_1, db) \kappa_3(a_1, b, dc) \quad (12)$$

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  $(A_i, \mathcal{A}_i)$ ,  $i = 1, \dots, n$  by defining a joint probability on the product space  $\prod_{i=1}^n A_i$  from  $n$  probability kernels  $\kappa_i$ . The result of Ionescu-Tulcea is then the following [24]. Let  $\kappa_1$  be a probability measure on  $(A_1, \mathcal{A}_1)$  and for all  $n \geq 2$ ,  $\kappa_n : \left( \prod_{i=1}^{n-1} A_k \right) \times \mathcal{A}_n \rightarrow [0, 1]$  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 (12).

**Daniell-Kolmogorov [30].** 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  $A_1, A_2, \dots$  be a sequence of measurable spaces and  $\mu_n$  a measure on the product space  $A_1 \times \dots \times A_n$ . We say that the sequence of probability measures  $\mu_n$  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  $\mu_n$  forms a projective family, then there exists a measure  $\mu$  on  $\prod_{i=1}^{\infty} A_i$  such that  $\mu_n$  is equal to the projection of  $\mu$  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

$$\delta = (\nu, \phi_1, \varrho_1, \dots, \phi_\nu, \varrho_\nu)$$

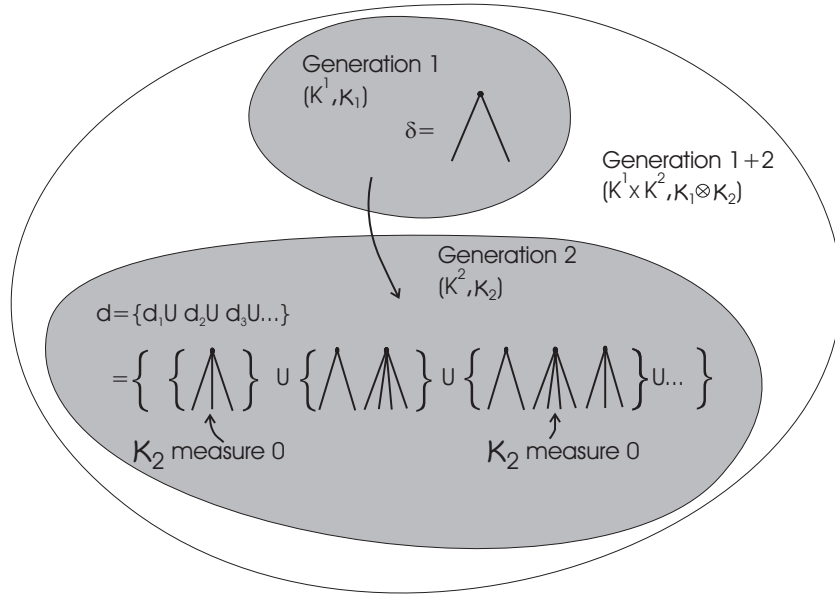
An element of this space carries information about the node to which it is attached: it contains the number of children of the node (random variable  $\nu$ ) and the operators attached to each of its branches. The probability measure  $\kappa_1 = P$  lets us build the sample space for first generation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

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

The  $\sigma$ -algebra associated with  $K^j$  is

$$\mathcal{D}_j = \sigma \left( \bigcup_{k \geq 1} \mathcal{D}^k \right) = \{ d_1 \cup d_2 \cup \dots \mid d_i \in \mathcal{D}^i \} \quad (13)$$

where  $\mathcal{D}^k = \mathcal{D} \times \dots \times \mathcal{D}$   $k$  times. In (13), 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  $j \geq 2$ . The  $\sigma$ -algebra attached to each  $K^j$  is therefore the same. Also, we need to consider the smallest  $\sigma$ -algebra spanned by the union of  $\mathcal{D}^k$  since the union of  $\sigma$ -algebras is not in general a  $\sigma$ -algebra.



**Figure 4.** Spaces  $K^1$ ,  $K^2$  and  $K^1 \times K^2$  with their respective probability measures  $\kappa_1$ ,  $\kappa_2$  and  $\kappa_1 \otimes \kappa_2$ .  $\delta \in K^1$  has 2 children. Conditionally on  $\delta$ , only  $d_2 \in K^2$  represented here has non-zero measure as it is the only element composed of 2 families.  $\kappa_2$  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.

The construction of  $\kappa_2$  supposes we know the first generation and in particular its size  $Z_1$ . For  $d = d_1 \cup d_2 \cup \dots \in \mathcal{D}_2$ , with  $d_j = E_1^j \times \dots \times E_j^j \in \mathcal{D}^j$ , we define

$$\kappa_2(d|Z_1) = \prod_{i=1}^{Z_1} P(E_i^{Z_1}) \quad (14)$$

Sets  $d_i$  for  $i \neq Z_1$  therefore receive a zero measure. This is illustrated on figure 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  $\kappa_2$  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  $\sigma$ -algebra and  $\kappa$  the probability distribution over this space.

**Definition 3.1**  $(K, \mathcal{K}, \kappa)$  is the probability space of extended Galton-Watson trees.

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  $\nu_\emptyset$  be the number of branches rooted at  $\emptyset$ . Then each node coming from the root is denoted by  $i$ , for  $i = 1, \dots, \nu_\emptyset$ . The

second generation of the tree is denoted  $ij$  for  $1 \leq j \leq \nu_i$ . More generally, a node is an element of  $U = \bigcup_{n \geq 0} \mathbb{N}^{*n}$  and a branch is a couple of nodes  $(u, uj)$  where  $u \in U$  and  $j$  is a strictly positive integer. Lastly, we consider  $T_u(k)$  the subtree of  $k \in K$  rooted at  $u$ :  $T_u(k) = \{v \mid v \in U \text{ and } uv \in k\}$ . By construction, the random variables  $T_i$  are independent and identically distributed (equation (14)).

To be consistent with the fact that the fixed point lies at the root of its construction tree, we write  $\nu_0$  for  $\nu$  in (11) for the remainder of this paper.

#### 4. Existence and Uniqueness of a fixed point

This section makes precise the conditions under which the Galton-Watson IFS defined in equation (19) possesses a unique fixed point.

**Theorem 4.1** *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  $\mathbb{E}_\kappa \sum_{j=1}^{\nu_0} r_j \int |\phi_j(0, x)|^p dx < +\infty$  for some  $1 < p < +\infty$ , where  $r_j$  is the contractive factor of  $\varrho_j$  with  $0 < r_j < 1$  almost surely, each  $\phi_j(\cdot, \cdot)$  is a.s. Lipschitz in its first variable, with Lipschitz constant  $S_j$  and  $\lambda_p = \mathbb{E} \sum_{j=1}^{\nu_0} r_j S_j^p < 1$ , where  $\mathbb{E}$  denotes the expectation under  $\kappa$ , there exists a unique function  $f^*$  which satisfies  $f^* = Tf^*$  in  $\mathbb{L}_p$ . Moreover, for all  $f_0 \in \mathbb{L}_p(\mathbb{X})$ ,*

$$d_p^*(T^n f_0, f^*) \leq \frac{\lambda_p^{n/p}}{1 - \lambda_p^{1/p}} d_p^*(f_0, Tf_0) \quad (15)$$

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

*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, write  $f_k$  for the realization of  $f$  at point  $k \in K$ , then define  $f_k^{(j)}$  by  $f_k^{(j)} := f_{T_j(k)}$ . Since the random variables  $T_j$  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 (11) 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

$$\mathbb{E} \int_{\mathbb{X}} |(Tf)(x)|^p dx = \mathbb{E} \sum_{j=1}^{\nu_0} \int_{\varrho_j(\mathbb{X})} |\phi_j[f^{(j)}(\varrho_j^{-1}(x)), \varrho_j^{-1}(x)]|^p dx$$

Since the  $\varrho_j$  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  $r_j$ .  $\mathbb{E} \int_{\mathbb{X}} |(Tf)(x)|^p dx$  is therefore bounded above by:

$$\mathbb{E} \sum_{j=1}^{\nu_0} r_j \mathbb{E} \left[ \int_{\mathbb{X}} |\phi_j[f^{(j)}(y), y]|^p dy \middle| \phi_j \right] \quad (16)$$

In (16), we have also used the law of total probability:  $\mathbb{E}(\cdot) = \mathbb{E}[\mathbb{E}(\cdot | \{\nu_0, \{\phi_j, \varrho_j\}\})]$  where the second expectation is conditioned on the IFS parameters. Terms depending on  $\nu$  and  $\varrho_j$  can be put outside the second expectation, leaving us with a term which only depends on  $\phi_j$ , hence (16). Note that the term in the inner expectation does not depend any more on the  $\varrho_i$ '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  $|x+y|^p \leq 2^p(|x|^p + |y|^p)$  it follows that  $\mathbb{E} \int_{\mathbb{X}} |(Tf)(x)|^p dx$  is bounded by:

$$2^p \mathbb{E} \sum_{j=1}^{\nu_0} r_j d_p^{*p}(\phi_j[f_\kappa^{(j)}, Id], \phi_j[0, Id]) + 2^p \mathbb{E} \sum_{j=1}^{\nu_0} r_j d_p^{*p}(\phi_j[0, Id], 0) \quad (17)$$

Using the Lipschitz property of the  $\phi_j$ , the first term of (17) is smaller than

$$2^p \mathbb{E} \sum_{j=1}^{\nu_0} r_j S_j^p d_p^{*p}(f_\kappa^{(j)}, 0)$$

which is bounded since  $f \in \mathbb{L}_p$ . The second term of (17) is proportional to  $\mathbb{E} \sum_{j=1}^{\nu_0} r_j \int |\phi_j(0, x)|^p dx$  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 4.1. 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

$$\begin{aligned} d_p^{*p}(Tf, Tg) &= \mathbb{E} \int \left| (Tf)(x) - (Tg)(x) \right|^p dx \\ &\leq \mathbb{E} \sum_{j=1}^{\nu_0} r_j \mathbb{E}^* \left[ \int_{\mathbb{X}} \left| \phi_j[f^{(j)}(y), y] - \phi_j[g^{(j)}(y), y] \right|^p dy \right] \end{aligned}$$

where  $\mathbb{E}^* = \mathbb{E}[\cdot | \{\nu_0, \{\phi_i, \varrho_i\}\}]$ . Lastly, we use the Lipschitz property of the non linear random maps  $\phi_j$  to conclude that

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

where the definition of  $\lambda_p$  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,

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

which leads to

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

Now using the triangle inequality:

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

so that:

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

which concludes the proof of the theorem. ■

To illustrate, we present in figure 5 a realization 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

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

where

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

It follows that

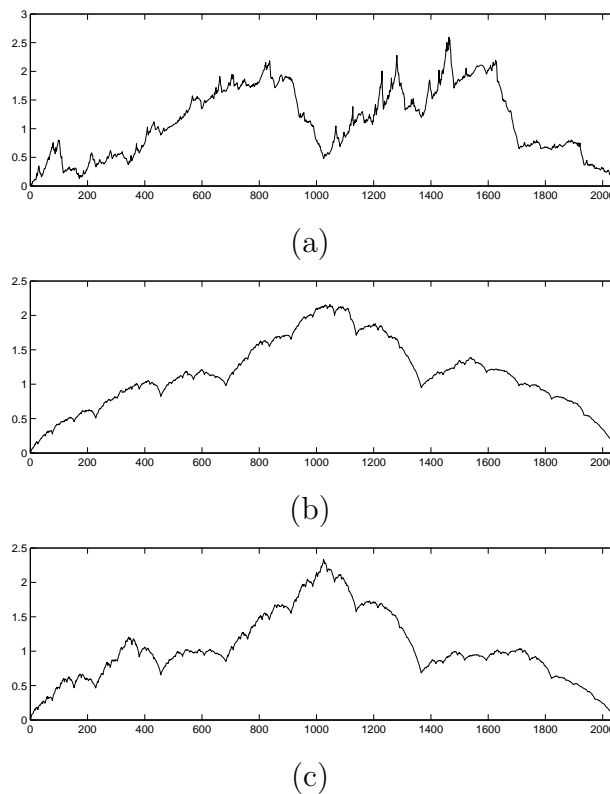
$$\sum_{n \geq 1} P(d_p^p(T^n f, f^*) > \epsilon) < \infty$$

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

$f^*$  is the unique fixed point for which  $f^* = T f^*$  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  $T f^0$ . The following result can be proven in the same way as Hutchinson and Rüschemdorff [9].

**Corollary 4.2** *The distribution of  $f^*$  is the unique distribution which satisfies  $f^* \stackrel{d}{=} T f^*$ , 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 lead 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.



**Figure 5.** A realization of the fixed point (a) and its mean (b).  $\phi_j$  are decomposed as follows  $\phi_j(x, t) = s_j x + X\zeta_j(t)$  for  $j = 1, \dots, \nu$  where  $X$  is normally distributed with mean 1 and variance 0.25. When  $\nu = 1$ ,  $s_1 = 0.6$  and  $\zeta_1(t) = t(1 - t)$ . For  $\nu = 2$  we define  $s_1 = 0.6$ ,  $s_2 = 0.7$ ,  $\zeta_1(t) = t^3$ ,  $\zeta_2(t) = 1 - t^2$  and for  $\nu = 3$  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  $\zeta_3(t) = 0.5(1 - t^2)$ .  $\nu$  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.

## 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 figure 5 where a small change in  $\mathbf{q}$  induces 'small' variations in the mean of the fixed point.

### 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 5.1**  $\mathbb{X} = [a, b]$ . Let  $\alpha$  be the unique random fixed point of  $\phi_1(\cdot, a)$  and  $\beta$  the unique random fixed point of  $\phi_{\nu_0}(\cdot, b)$ :  $\phi_1(\alpha, a) = \alpha$  and  $\phi_{\nu_0}(\beta, b) = \beta$ . Assume that  $\alpha$  and  $\beta$  are the same for all possible realisations of  $\phi_1$  and  $\phi_{\nu_0}$ . If  $\phi_i(\beta, b) = \phi_{i+1}(\alpha, a)$

a.s. for all  $i \in \{1, \dots, \nu_0 - 1\}$  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  $\phi_1$  and  $\phi_{\nu_0}$ :

$$\begin{aligned} f^*(a) &= \phi_1[f^*(\varrho_1^{-1}(a)), \varrho_1^{-1}(a)] = \phi_1[f^*(a), a] \\ f^*(b) &= \phi_{\nu_0}[f^*(\varrho_{\nu_0}^{-1}(b)), \varrho_{\nu_0}^{-1}(b)] = \phi_{\nu_0}[f^*(b), b] \end{aligned}$$

Those equalities have to remain true whatever  $\nu_0$  is, which is realized under the assumption of proposition 5.1.

Let  $\varrho_i[a, b] = [a_{i-1}, a_i]$  for  $i \in \{1, \dots, \nu_0\}$  and  $a_0 = a$ ,  $a_{\nu_0} = b$  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 [9]. 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  $\varrho_{i+1}[a, b]$ :

$$f^*(a_i) = \phi_i[f^*(\varrho_i^{-1}(a_i)), \varrho_i^{-1}(a_i)] = \phi_i[f^*(b), b] = \phi_i[\beta, b] \quad (18)$$

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  $\phi_{\nu_0, i}$  and  $\phi_{\nu_0, i+1}$  at the random point  $a_i$  will result in an everywhere discontinuous fixed point. A realization of a continuous fixed point is represented figure 5.

## 5.2. Continuous dependency w.r.t. $\mathbf{q}$

The continuity of the moments of the fixed point with respect to  $\mathbf{q}$  is suggested in figure 5. This observation is related to the one made by Barnsley in [10] 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 [23], with deterministic maps and a random tree.

Consider the set of deterministic maps

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

Given  $\nu_0 = j$ , then apply  $\{\phi_{j,1}, \dots, \phi_{j,j}, \varrho_{j,1}, \dots, \varrho_{j,j}\}$ .  $\phi_{k,j}$  and  $\varrho_{k,j}$  may differ for different values of  $k$ ,  $j = 1, \dots, k$ . The operator  $T$  becomes:

$$(Tf)(x) = \sum_{j=1}^{\nu_0} \phi_{\nu_0, j}[f^{(j)}(\varrho_{\nu_0, j}^{-1}(x)), \varrho_{\nu_0, j}^{-1}(x)] \mathbf{1}_{\varrho_{\nu_0, j}(\mathbb{X})}(x) \quad (19)$$

Lipschitz factor of  $\phi_{\nu_0, j}$  is  $S_{\nu_0, j}$  and the contraction factor of  $\varrho_{\nu_0, j}$  is denoted by  $r_{\nu_0, j}$ . Since the operators attached to the branches of the tree are the same for a given number

of offsprings, to one realization of the tree is associated one and only one realization of the fixed point.

**Theorem 5.2** *Suppose conditions of theorem 4.1 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  $0 \leq s < 1$  and  $\zeta$  is a nonlinear function. Suppose that  $\lambda_r = \mathbb{E} \sum_{j=1}^{\nu} r_j S_j^r < 1$  for  $r = 1, \dots, p$ . Then the  $r$ -th moment of  $f^*$  continuously varies with respect to the probability generating vector  $\mathbf{q}$ , for  $r = 1, \dots, p$ .*

*Proof:* We prove the theorem by recurrence on  $r$ , for  $r = 1, \dots, p$ . 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  $\Upsilon$  be the space of probability vectors

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

where  $\mathbb{N}^* := \{1, 2, \dots\}$ . 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 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  $\psi : \Upsilon \rightarrow L_1$  which associates with each probability vector the mean of the fixed point of the Galton-Watson IFS. Let  $\mathbf{p} \in \Upsilon$ , we want to show that for all  $\epsilon > 0$ , there exists  $\eta > 0$  such that

$$\forall \mathbf{q} \in \Upsilon \quad l(\mathbf{p}, \mathbf{q}) \leq \eta \Rightarrow d_1(\mathbb{E}_{\mathbf{p}} f_{\mathbf{p}}^*, \mathbb{E}_{\mathbf{q}} f_{\mathbf{q}}^*) \leq \epsilon \quad (20)$$

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

$$\begin{aligned} d_1(\mathbb{E}_{\mathbf{p}} f_{\mathbf{p}}^*, \mathbb{E}_{\mathbf{q}} f_{\mathbf{q}}^*) &= \int |\mathbb{E}_{\mathbf{p}} \sum_{j=1}^{\nu_1} \phi_{\nu_1, j} [f_{\mathbf{p}}^* \circ \varrho_{\nu_1, j}^{-1}, \varrho_{\nu_1, j}^{-1}] \mathbf{1}_{\varrho_{\nu_1, j}(\mathbb{X})} \\ &\quad - \mathbb{E}_{\mathbf{q}} \sum_{j=1}^{\nu_2} \phi_{\nu_2, j} [f_{\mathbf{q}}^* \circ \varrho_{\nu_2, j}^{-1}, \varrho_{\nu_2, j}^{-1}] \mathbf{1}_{\varrho_{\nu_2, j}(\mathbb{X})}| \end{aligned}$$

where  $\kappa_{\mathbf{p}}(\nu_1 = k) = p_k$  and  $\kappa_{\mathbf{q}}(\nu_2 = k) = q_k$ . We omit the variable  $x$  in the integrand to keep the notation clear. By conditioning with respect to  $\nu_1$  and  $\nu_2$ , the right hand

side becomes:

$$\int \left| \sum_{i \geq 1} p_i \sum_{j=1}^i \mathbb{E}_{\mathbf{p}} \phi_{i,j} [f_{\mathbf{p}}^* \circ \varrho_{i,j}^{-1}, \varrho_{i,j}^{-1}] \mathbf{1}_{\varrho_{i,j}(\mathbb{X})} - \sum_{i \geq 1} q_i \sum_{j=1}^i \mathbb{E}_{\mathbf{q}} \phi_{i,j} [f_{\mathbf{q}}^* \circ \varrho_{i,j}^{-1}, \varrho_{i,j}^{-1}] \mathbf{1}_{\varrho_{i,j}(\mathbb{X})} \right|$$

The sums can be taken outside the integral. By setting  $y = \varrho_{i,j}^{-1}(x)$  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:

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

To go further, a particular form for the  $\phi_{i,j}$  is required. We consider the case when they are pure contractions in their first variable and are non linear in their second variable:  $\phi_{i,j}(u, v) = s_{i,j}u + \zeta_{i,j}(v)$ . 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

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

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  $l(\mathbf{p}, \mathbf{q}) \leq \eta$ . We have

$$\begin{aligned} d_1(\mathbb{E}_{\mathbf{p}} f_{\mathbf{p}}^*, \mathbb{E}_{\mathbf{q}} f_{\mathbf{q}}^*) &\leq \eta \sum_{i,j} r_{i,j} s_{i,j} \int |\mathbb{E}_{\mathbf{p}} f_{\mathbf{p}}^*| \\ &\quad + \sum_{i,j} q_i r_{i,j} s_{i,j} \int |\mathbb{E}_{\mathbf{p}} f_{\mathbf{p}}^* - \mathbb{E}_{\mathbf{q}} f_{\mathbf{q}}^*| + \eta \sum_{i,j} r_{i,j} \int |\zeta_{i,j}(y)| dy \end{aligned}$$

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  $\lambda_1(\mathbf{q}) \leq 1 - \varepsilon_p$  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

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

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

**Second step.** We now show that the previous result holds for the  $r$ -th moment of the fixed point, as long as  $r \leq p$ . 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  $d_1(\mathbb{E}_{\mathbf{p}}f_{\mathbf{p}}^{*2}, \mathbb{E}_{\mathbf{q}}f_{\mathbf{q}}^{*2})$  by

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

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

$$\begin{aligned} d_1(\mathbb{E}_{\mathbf{p}}f_{\mathbf{p}}^{*2}, \mathbb{E}_{\mathbf{q}}f_{\mathbf{q}}^{*2}) \leq & \sum_{i,j} r_{i,j} \left[ \int s_{i,j}^2 |p_i \mathbb{E}_{\mathbf{p}} f_{\mathbf{p}}^{*2} - q_i \mathbb{E}_{\mathbf{q}} f_{\mathbf{q}}^{*2}| \right. \\ & \left. + 2s_{i,j} \zeta_{i,j}(y) |p_i \mathbb{E}_{\mathbf{p}} f_{\mathbf{p}}^* - q_i \mathbb{E}_{\mathbf{q}} f_{\mathbf{q}}^*| + \zeta_{i,j}^2(y) dy \right] \end{aligned}$$

Since for all  $r \leq p$ ,  $f \in \mathbb{L}_p(\mathbb{X}) \Rightarrow f \in \mathbb{L}_r(\mathbb{X})$ ,  $f^*$  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  $\eta$  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  $\eta$ . The continuity follows.

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

### 5.3. Test for multifractality.

Fractal processes are by construction highly irregular. Information about the local fluctuations of a process  $X(t)$  is made precise with the definition of the Hölder exponent at a specific time  $t = t_0$ . The process  $X$  is said to belong to  $C_{t_0}^h$  if there is a polynomial  $P_{t_0}$  such that

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

in a neighborhood of  $t_0$ . The largest value  $H$  of  $h$  such that  $X \in C_{t_0}^h$  is the Hölder exponent of  $X$  at  $t = t_0$  [25]. Monofractal processes have a constant local Hölder exponent along their sample paths. In opposition, multifractals possess a richer structure. Their Hölder exponent behaves erratically with time: each interval of positive length exhibit a full range of different exponents. In practice, it is not possible to estimate the evolution of the Hölder exponent with time. Instead, multifractal processes are described by their Hausdorff spectrum  $D(h)$ , which gives the size (Hausdorff dimension) of sets with a given exponent  $h$ . The spectrum cannot be observed precisely in practice and alternative methods for its estimation have been proposed, giving birth to the multifractal formalism.  $D(h)$  is usually estimated via the Legendre transform of a partition function  $\zeta(q)$ , obtained as a power law behaviour of multiresolution quantities. Jaffard, Lahermes and Abry have proposed an estimator of  $\zeta(q)$  using wavelet leaders

[26]. The Legendre transform of  $\zeta(q)$  provides in general an upper bound of the Hausdorff spectrum of  $X$

$$D(h) \leq \inf_{q \neq 0} (1 + qh - \zeta(q))$$

Consider the polynomial expansion of  $\zeta(q) = c_1q + c_2q^2/2 + c_3q^3/6 + \dots$ . When  $c_p = 0$  for all  $p \geq 2$ ,  $\zeta(q)$  is linear with  $q$  and the Hausdorff spectrum degenerates to a single point: the process is monofractal. Any departure from a linear behaviour is characteristic of multifractals. Wendt and Abry have designed a test to decide whether  $c_p = 0$  or not [27]. In particular, the case  $p = 2$  permits to conclude between a mono and multifractal process. The null hypothesis  $c_p = c_{p,0}$  ( $H_0$ ) is tested versus the two sided alternative  $c_p \neq c_{p,0}$ . The test statistic is therefore

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

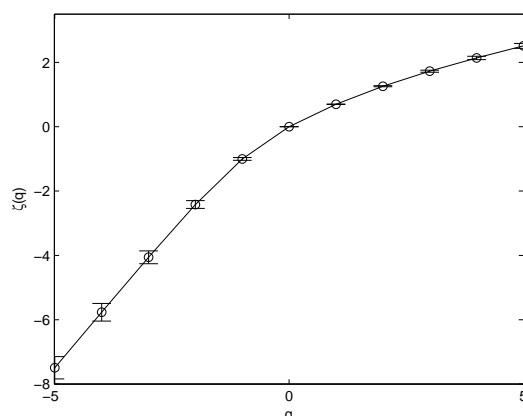
The distribution of the test statistic  $T$  under the null hypothesis is unknown in general and is estimated using non parametric bootstrap techniques. From the empirical distribution, one can design an acceptance region  $\mathcal{T}_{1-\alpha} = [t_{\alpha/2}, t_{1-\alpha/2}]$  where  $t_\alpha$  is the  $\alpha$  quantile of the null distribution. In the present study, we set  $\alpha = 0.1$  and  $c_{p,0} = 0$ . We refer the reader to [27] for further details about the test statistics.

Consider the Galton-Watson IFS presented in the figure 5 with  $\nu$  taking values 1, 2 or 3 with probabilities 0.2, 0.3 and 0.5. We simulate 100 realizations of the fixed point, each of length  $2^{12}$  and estimate the partition function using wavelet leaders. The wavelets coefficients are computed with Daubechies wavelets with 2 vanishing moments and the scale of analysis ranges from  $j_1 = 3$  to  $j_2 = 12$ . The partition function presented figure 6 is obtained by averaging 100 estimations of  $\zeta(q)$ . It clearly appears non-linear, which suggests the multifractal behaviour of the fixed point. Also, in the multifractality test previously described,  $H_0$  is rejected 99% of time for  $p = 2$ , confirming non-linear shape of  $\zeta(q)$ . Those observations indicate the existence of a class of Galton-Watson fixed points which are multifractals. This result motivates a further study in which one could derive conditions on the IFS parameters in order to obtain multifractal fixed points.

## 6. Conclusion

This study extends the model proposed by Hutchinson and Rüschemdorff by adding more randomness into it, considering random operators on a random underlying construction tree. The strict self-similarity observed in attractors of deterministic IFS is a drawback when modeling natural signals. Adding randomness gives visually more interesting models and it would be interesting to consider parameter estimation problems on Galton-Watson IFS. To start, one could give operators a particular form and estimate the offspring distribution of the random tree.

Further theoretical properties of the fixed point need also to be carried out, starting with its multifractal analysis. The cascade nature of IFS lets us imagine a non



**Figure 6.** Estimation of the partition function  $\zeta(q)$  of the fixed point of a Galton-Watson IFS using wavelet leaders. The parameters of the IFS are given in the figure caption 5, with  $\nu$  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.

trivial multifractal spectrum of the fixed point, intuition confirmed by the simulations performed in the previous section. Existing results for deterministic IFS on functions [8] and for random IFS on measures [28] motivate this study.

## References

- [1] A.Bunde and S.Havlin (eds.), *Fractals in Science*, Berlin, New-York: Springer-Verlag, 1994
- [2] B.B.Mandelbrot, *Gaussian Self-Affinity and Fractals*, New-York: Springer-Verlag, 2002
- [3] P.Abry, R.Baraniuk, P.Flandrin, R.Riedi and D.Veitch, *The Multiscale Nature of Network Traffic: Discovery, Analysis and Modelling*, IEEE Signal Processing Magazine, 19(3): pp.28–46, 2002
- [4] A.Arneodo, F.Argoul, E.Bacry, J.Elezgaray and J.F.Muzy, *Ondelettes, multifractales et turbulences. De l'ADN aux croissances cristallines*, Diderot, Paris, 1995.
- [5] M.F.Barnsley and L.P.Hurd, *Fractal Image Compression*, AK Peters Limited, 1993
- [6] D.S.Mazel and H.Hayes, *Fractal modeling of time-series data*, in proc. 23rd Asilomar Conf. Signals, Syst. Comput, 1989
- [7] D.S.Mazel and H.Hayes, *Using Iterated Function Systems to Model Discrete Sequences*, IEEE trans. on signal processing, vol.40, no7, pp.1724–1734, 1992
- [8] S.Jaffard, *Multifractal formalism for functions, part 1 and 2*, SIAM J. of Math. Anal. 28(4): pp.944–998, 1986
- [9] J.E.Hutchinson and L.Rüschendorf *Fractal Geometry and Stochastics*, Chapter Self-Similar fractals and selfsimilar random fractals pp.109–123, Ed. C. Bandt, S.Graf, M.Zhhe, Birkhauser Addison-Wesley, 2000
- [10] M.F.Barnsley, *Fractals Everywhere*, Academic Press, 1988
- [11] K.Falconer, *Random fractals*, Math. Proc. Camb. Phil. Soc., vol.100, pp.559–582, 1986
- [12] R.D.Mauldin and S.C.Williams, *Random recursive constructions: asymptotic geometric and topological properties*, Trans. Amer. Math. Soc., vol.295, pp.325-346 1986
- [13] F. Galton and H.W. Watson, *On the probability of the extinction of families*, J. Anthropol. Soc. London (Royal Anthropol. Inst. G.B. Ireland), vol.4, pp.138-144 1875

- [14] A. Yates, C. Chan, J. Strid, S. Moon, R. Callard, A. George and J. Stark, *Reconstruction of cell population dynamics using CFSE*, BMC Bioinformatics., vol.8, pp.196 2007
- [15] K.B Athreya and P.E. Ney, *Branching Processes*, BSpringer-Verlag, 1972
- [16] J. Peyrière, *Calculs de dimensions de Hausdorff*, Duke Math J., 44, pp.591–601 1977
- [17] J. Hawkes, *Trees Generated by a Simple Branching Process*, J. London Math. Soc., 24(2), pp.373–384 1981
- [18] G.A. Burd and E.C. Waymire, *Independent Random Cascades on Galton-Watson Trees*, Proc. Amer. Math. Soc., 128(9), pp.2753–2761 2000
- [19] Q. Liu, *On Generalized Multiplicative Cascades*, Stochastic Processes and Their Applications, 86, pp.263–286 2000
- [20] Q. Liu, *Local Dimensions of the Branching Measure on a Galton-Watson Tree*, Ann. Inst. H. Poincaré, Probabilités et Statistiques, 37(2), pp.195–222 2001
- [21] P. Mörters and N-R. Shieh, *On the Multifractal Spectrum of the Branching Measure on a Galton-Watson Tree*, J. Appl. Prob, 41, pp.1223–1229 2004
- [22] J.Neveu, *Arbres et processus de Galton-Watson*, Annales de l’Institut Henri Poincaré. Probabilités et Statistiques, vol.22, no.2, pp.199–207, 1986
- [23] G.Decrouez, P-O.Amblard, J-M.Brossier and O.Jones, *Galton Watson fractal signals*, In Proc. IEEE ICASSP, Hawaii, USA, III:1157–1160, 2007
- [24] A.N.Shiryayev, *Probability*, Springer Verlag, New York, 1984
- [25] R.H. Riedi, *Multifractal processes*, In Theory And Applications Of Long-Range Dependence, P. Doukhan, G. Oppenheim, and M. S. Taqqu, editors. pp.625–716, 2003
- [26] S.Jaffard, B.Lashermes and P.Abry, *Wavelet Leaders in Multifractal Analysis*, Applied and Numerical Harmonic Analysis, Birkhuser Verlag Basel/Switzerland, pp.219–264, 2006
- [27] H. Wendt and P. Abry, *Multifractality Tests using Bootstrapped Wavelet Leaders*, IEEE Trans. on Signal Processing. 55(10): pp.4811–4820, 2007
- [28] N.Patzschke and U.Zähle, *Self-similar random measures. IV. The recursive construction model of Falconer, Graf and Mauldin and Williams.*, Math. Nachr., 149–pp.285–302, 1990
- [29] G.O. Okikiolu, *Aspects of the Theory of Bounded Integral Operators in  $L^p$  spaces.*, Academic Press Inc. (London), 1971
- [30] C.W. Burrill, *Measure, Integration and Probability.*, Mc Graw-Hill, 1972