Asymptotic behaviour of random growing trees
SYNOPSIS

Anna Rudas

December 2012

Contents

1 Introduction

1.1 Model class and context

In my PhD thesis I investigate the asymptotic properties of a random tree growth model which generalizes the basic concept of preferential attachment.

In this family of tree growth models, the tree stems from a root in the beginning, and vertices are added one at a time, the new vertex always attaching to exactly one already existing vertex. The rule by which the new vertex chooses its "parent", is dependent on the degree distribution apparent in the tree at the time the vertex is born. This dependence on the degree structure is characterised by a *weight function* $w : \mathbb{N} \to \mathbb{R}_+$, which function is the parameter of the model.

The models can be either in discrete time, when a vertex is born in every second, or in continuous time, then birth times are random. For the problems we discuss, these two versions are equivalent and can be translated into each other. The classical models and results of the area use the discrete time setting. However, for the proofs we give, the continuous-time version is much more natural and convenient, so this is what we will use.

One of the famous models, a realization of preferential attachment, is the Barabási - Albert graph [3], where the random choice of the parent for the new vertex is made using probabilities exactly proportional to the degree of the existing vertices. The tree case of this model corresponds to the the special case of the model considered in this dissertation, namely, when w is chosen to be linear. The Barabási - Albert graph reproduces certain phenomena observed in real-world networks, the power-law decay of the degree sequence, for example. This was proved in a mathematically precise way in Bollobás et al. [8] and, independently, in Móri [32]. In these papers, the techniques strongly depend on martingales that are apparent in the system only in the linear case.

The concept of preferential attachment generally means that the weight function w is an increasing function. In the family of models that we are interested in, this is not necessarily true. General weight functions are considered in the papers of Krapivsky and Redner [25] and [26], where $w(k) \sim k^\gamma$, and non-rigorous results are obtained, showing the different behaviour for $\gamma > 1$ and $\gamma \leq 1$. In the first region a single dominant vertex appears which is linked to almost every other vertex, the others having only finite degree. The weight functions we consider in the present dissertation are such that the model does not "blow up" this way, our class includes the second regime $\gamma \leq 1$ mentioned above.

In the work of Dereich and Mörters [11], the authors take a closer look at the temporal evolution of the degrees of individual vertices, in the same sublinear preferential attachment case as we do. This paper refers to our work [41]. Certain random recursive trees and random plane-oriented trees similar to our setting have also been studied before in Smythe and Mahmoud [44].

Population growth models, studied excessively in the theory of branching processes (see e.g. Jagers [22]), are intimately related to our model. This connection is the basis for many of the proofs in the present dissertation.

Tree growth processes based on fragmentation processes are closely related to our investigation of the *global properties* of the model. Limiting objects called "random real trees" and "continuum random trees" were introduced, to which the evolving trees converge, after an appropriate rescaling of the *distances* on the tree. Much of the structure of these limiting objects is understood, see e.g. Haas, Miermont et al. [19, 20, 21].

Our concept of the limiting measure μ in Section 4 is different from these. It is a measure on the set of leaves of the infinite complete tree (with each vertex having exactly K children), which is a metric space, but the metric structure is trivial: it is not a result of any spatial scaling, and it carries no information about the tree growth process. On the other hand, the weights given by μ are a result of an appropriate rescaling of the tree size, where size means *cardinality.* In short, we are really interested in the asymptotic weight distribution, and not the asymptotic metric structure. This asymptotic weight distribution is also studied in the Physics literature, see e.g. Berestycki [5], where a quantity analogous to the local dimension is calculated for a continuous time fragmentation process.

Similarly, in the limiting continuous trees obtained in Haas, Miermont et al. [19, 20, 21] by a spatial rescaling of the evolving tree, the metric structure is of main interest, and the Hausdorff dimension and Hausdorff measure of *sets* are the natural questions to ask, see Duquesne and Le Gall [15, 16]. However, in our model it is not the set, but the measure which captures the long-term structure of the tree well, and of which the dimension is interesting.

The continuous time version of our tree growth process can also be translated into a branching random walk, with time turning into displacement. Then the asymptotic growth can be described analogously, see the Biggins theorem in [7] or Lyons [30]. However, with that point of view, the natural questions about the limiting structure are quite different.

1.2 Structure of the thesis

The thesis is divided into three main sections in which the definition of the model, and the analysis of two families of properties of the random tree growth model are provided. The rest of this document is also divided into three sections, accordingly.

In Section 3 we analyse *local properties* of the tree: we focus on the neighborhood of the *typical* vertex (e.g. sampled uniformly randomly after a long time) of the random tree. This section is based on [41], joint work with Benedek Valkó and Bálint Tóth, and [40], joint work with Bálint Tóth.

The topic of Section 4 is the analysis of the *global properties* of the model, these capture phenomena observable by looking at the whole tree in the limit (e.g. asymptotic speed of tree growth in the continuous time setting, and the "limiting success level" of a fixed vertex in the limit). This section is based on [42], joint work with Imre Péter Tóth.

In this document we begin both Section 3 and 4 with a short résumé of the context, continue with the description of the scope of the analyses, and we precisely state the results of the PhD thesis. We also hint at the methods of our proofs.

2 Terminology, notation and the model

2.1 Vertices, individuals, trees

We consider rooted ordered trees, which are also called family trees or rooted planar trees in the literature.

In order to refer to these trees it is convenient to use genealogical phrasing. The tree is thus regarded as the coding of the evolution of a population stemming from one individual, the root of the tree, whose "children" form the "first generation", these are the vertices connected directly to the root. In general, the edges of the tree represent parent-child relations, the parent always being the one closer to the root. The birth order between brothers is also taken into account, this is represented by the tree being an ordered tree (planar tree).

Let us fix a subset of positive integers, $\mathbb{I}$, and let us label the vertices of a rooted ordered tree using the elements of

$$\mathcal{N} := \bigcup_{n=0}^{\infty} \mathbb{I}^n, \quad \text{where} \quad \mathbb{I}^0 := \{\emptyset\} . \tag{1}$$

We will consider slightly different cases of the model in Sections 3 and 4, and define $\mathbb{I}$ in the two sections accordingly, as follows.

- Throughout Section 3 we choose $\mathbb{I} = \mathbb{Z}^+$, this corresponds to the fact that any vertex can have any number of children.
- In Section 4, we will fix a positive integer $K \in \mathbb{N},; K \geq 2$, and choose $\mathbb{I} := \{1, 2, \dots, K\}$. This means that we restrict the weight function in such a way that the vertices can have at most K number of children.

In our notation $\emptyset$ denotes the root of the tree, its children are labelled with the elements of $\mathbb{I}$, and in general the children of $x = (x_1, x_2, \dots, x_k) \in \mathcal{N}$ are labelled by $(x_1, x_2, \dots, x_k, 1), (x_1, x_2, \dots, x_k, 2), \dots$. Thus if a vertex has the label $x = (x_1, x_2, \dots, x_k) \in \mathcal{N}$ then this means that it is the x_k^{th} child of its parent, which is the x_{k-1}^{th} child of its own parent and so on.

We will identify a rooted ordered tree with the set of labels of its vertices, since this already contains the necessary information about the edges. It is clear that a $G \subset \mathcal{N}$ may represent a rooted ordered tree if and only if $\emptyset \in G$ and for each $(x_1, x_2, \dots, x_k) \in G$ we have $(x_1, x_2, \dots, x_{k-1}) \in G$ as well as $(x_1, x_2, \dots, x_k - 1) \in G$, if $x_k > 1$.

The set of finite rooted ordered trees will be denoted by $\mathcal{G}$. We think about $G \in \mathcal{G}$ as an oriented tree with edges pointing from parents to children. The *degree* of a vertex $x \in G$ is the number of its children in G, so this terminology differs a little bit from the usual:

$$\deg(x, G) := \max \{n \in \mathbb{I} : xn \in G\}.$$

The n^{th} *generation* of $G \in \mathcal{G}$ is

$$G_{[n]} := \{x \in G : |x| = n\}, \quad n \geq 0,$$

where $|x| = n$ iff $x \in \mathbb{I}^n$.

The n^{th} *ancestor of* $x = (x_1, x_2, \ldots, x_k) \in \mathcal{N}$ with $k \geq n$ is $x^n = (x_1, x_2, \ldots, x_{k-n})$ if $k > n$ and $x^n = \emptyset$ if $k = n$. In Section 4 we will also use the notation $p(x) = x^1$ for the parent of x.

The *subtree* rooted at a vertex $x \in G$ is:

$$G_{\downarrow x} := \{y : xy \in G\}, \tag{2}$$

this is just the progeny of x viewed as a rooted ordered tree. Also, (again with a slight abuse of notations) for an $x = (x_1, x_2, \ldots, x_n) \in \mathcal{N}$ with $|x| = n \geq k$ we use the notation $x_{\downarrow k} = (x_{n-k+1}, x_{n-k+2}, \ldots, x_n)$. This would be the new label given to $x \in G$ in the subtree $G_{\downarrow x^k}$.

2.2 The random tree model

As the parameter of the random tree model, we fix a weight function $w : \mathbb{N} \to \mathbb{R}_+$.

For the definition of the discrete time model, we do not need any further restrictions on w. In the continuous time case, we impose certain restrictions on w, see condition (M) later in this Section, these are needed for the model definition, and also for our results in Section 3. In Section 4, we will require $w(k) = 0, \quad k \geq K$, which will on one hand make sure that each vertex can have at most K children, and on the other hand, it automatically implies condition (M).

Discrete time model

Given the weight function $w : \mathbb{N} \to \mathbb{R}_+$, let us define the following discrete time Markov chain Υ^d on the countable state space $\mathcal{G}$, with initial state $\Upsilon^d(0) = \{\emptyset\}$. If for $n \geq 0$ we have $\Upsilon^d(n) = G$, then for a vertex $x \in G$ let $k := \deg(x, G) + 1$. Using this notation, let the transition probabilities be

$$\mathbf{P}(\Upsilon^d(n+1) = G \cup \{xk\}) = \frac{w(\deg(x, G))}{\sum_{y \in G} w(\deg(y, G))}.$$

In other words, at each time step a new vertex appears, and attaches to exactly one already existing vertex. If the tree at the appropriate time is G, then the probability of choosing vertex x in the tree G is proportional to $w(\deg(x, G))$.

Continuous time model

Given the weight function $w : \mathbb{N} \to \mathbb{R}_+$, let $X(t)$ be a Markovian pure birth process with $X(0) = 0$ and birth rates

$$\mathbf{P}\big(X(t + \mathrm{d}t) = k + 1 \,\big|\, X(t) = k\big) = w(k)\,\mathrm{d}t + o(\mathrm{d}t).$$

Let $\rho : [0, \infty) \mapsto (0, \infty]$ be the density of the point process corresponding to the pure birth process $X(t)$, namely let

$$\rho(t) = \lim_{\varepsilon \to 0} \varepsilon^{-1} \mathbf{P}((t, t+\varepsilon) \text{ contains a point from } X) \ .$$

Let $\widehat{\rho} : (0, \infty) \to (0, \infty]$ the (formal) Laplace transform of ρ:

$$\widehat{\rho}(\lambda) := \int_0^\infty e^{-\lambda t} \rho(t) \, \mathrm{d}t = \sum_{n=0}^\infty \prod_{i=0}^{n-1} \frac{w(i)}{\lambda + w(i)}.$$

Let

$$\underline{\lambda} := \inf\{\lambda > 0 : \widehat{\rho}(\lambda) < \infty\}.$$

Throughout the thesis we impose the following condition on the weight function w:

$$\lim_{\lambda \searrow \underline{\lambda}} \widehat{\rho}(\lambda) > 1. \tag{M}$$

We are now ready to define our randomly growing tree $\Upsilon(t)$ which will be a continuous time, time-homogeneous Markov chain on the countable state space $\mathcal{G}$, with initial state $\Upsilon(0) = \{\emptyset\}$.

The jump rates are the following: if for a $t \geq 0$ we have $\Upsilon(t) = G$ then the process may jump to $G \cup \{xk\}$ with rate $w(\deg(x, G))$ where $x \in G$ and $k = \deg(x, G) + 1$. This means that each existing vertex $x \in \Upsilon(t)$ 'gives birth to a child' with rate $w(\deg(x, \Upsilon(t)))$ independently of the others.

Note that condition (M) implies

$$\sum_{k=0}^\infty \frac{1}{w(k)} = \infty$$

and hence it follows that the Markov chain $\Upsilon(t)$ is well defined for $t \in [0, \infty)$, it does not blow up in finite time.

We will use τ_x to denote the birth time of vertex x,

$$\tau_x := \inf\{t > 0 \ : \ x \in \Upsilon(t)\} \ .$$

Connection between the discrete and continuous time models

If we only look at our process at the stopping times when a new vertex is just added to the randomly growing tree:

$$T_n := \inf\{t : |\Upsilon(t)| = n + 1\}$$

then we get the discrete time model: $\Upsilon(T_n)$ has the same distribution as $\Upsilon^d(n)$, the discrete time model at time n.

3 Local properties

3.1 Questions, context

In the present Section we investigate the *local properties* of the random tree after a long time of its evolution. We ask questions about the neighborhood of the "typical" vertex (e.g. sampled uniformly randomly) of the random tree, after a long time.

Our main results are the following. We determine the asymptotic distribution of the degree sequence, which equivalently gives the limiting distribution of the degree of a (uniformly) randomly selected vertex. We also look deeper into the structure of the tree: we give the asymptotic distribution of the subtree under a randomly selected vertex. Moreover, we present the asymptotic distribution of the whole tree, seen from a randomly selected vertex. For a general approach for asymptotic distribution of random subtrees of random trees, see [2]. These results give greater insight to the limiting structure of the random tree.

The key of our method is to place the process into continuous time, as already introduced in Section 2.2. The greatest advantage of this setting is definitely that it reveals the connection between the original, discrete time random tree model and the extensively studied framework of general branching processes. Our main *local* results gain their proofs through this relation. As an earlier application of a similar idea, see the paper [38] of B. Pittel, in which the author establishes the connection with a Crump–Mode branching process, and proves his results about the height of the uniform and general ordered recursive tree, and also for a random m-ary search tree.

3.2 Results

From condition (M) it follows that the equation

$$\widehat{\rho}(\lambda) = 1 \tag{3}$$

has a unique root λ^*.

Now we are ready to state our first theorem.

Theorem 3.1. *Consider a weight function w satisfying condition* (M) *and let λ^* be defined as above. Consider a bounded function $\varphi : \mathcal{G} \to \mathbb{R}$. Then the following limit holds almost surely:*

$$\lim_{t\to\infty} \frac{1}{|\Upsilon(t)|} \sum_{x\in\Upsilon(t)} \varphi(\Upsilon(t)_{\downarrow x}) = \lambda^* \int_0^\infty e^{-\lambda^* t} \mathbf{E}\big(\varphi(\Upsilon(t))\big)\,\mathrm{d}t.$$

From Theorem 3.1 several statements follow, regarding the asymptotic behavior of our random tree *as seen from a randomly selected vertex* ζ, chosen uniformly from $\Upsilon(t)$. As typical examples we determine the asymptotic distribution of the number of children, respectively, that of the whole subtree under the randomly chosen vertex, its k^{th} ancestor, respectively. That is: the asymptotic distribution of $\deg(\zeta, \Upsilon(t)) \in \mathbb{N}$, $\Upsilon(t)_{\downarrow\zeta} \in \mathcal{G}$ and $(\Upsilon(t)_{\downarrow\zeta^{(k)}}, \zeta_{\downarrow k}) \in \mathcal{G}^{(k)}$.

In order to formulate these consequences of Theorem 3.1 we need to introduce some more notation.

We call a probability measure $\boldsymbol{\pi}$ on $\mathcal{G}$ *steady* if

$$\sum_{H\in\mathcal{G}} \boldsymbol{\pi}(H) \sum_{x\in H_{[1]}} \mathbb{1}\{H_{\downarrow x} = G\} = \boldsymbol{\pi}(G). \tag{4}$$

Also, let $G \in \mathcal{G}$ and one of its historical orderings $s = (s_0, s_1, \ldots, s_{|G|-1}) \in \mathcal{S}(G)$ be fixed. The historical sequence of total weights are defined as

$$W(G, s, i) := W(G(s, i))$$

for $0 \leq i \leq |G| - 1$ while the respective weights of the appearing vertices are defined as

$$w(G, s, i) := w\left(\deg\left((s_i)^1, G(s, i-1)\right)\right).$$

for $1 \leq i \leq |G|-1$. Since $\deg\left((s_i)^1, G(s, i-1)\right)$ is the degree of s_i's parent just before s_i appeared, $w(G, s, i)$ is the rate with which our random tree process jumps from $G(s, i-1)$ to $G(s, i)$.

Given the weight function $w : \mathbb{N} \to \mathbb{R}_+$ satisfying condition (M) and λ^* defined as before define

$$\begin{aligned}
\boldsymbol{p}_w(k) &:= \frac{\lambda^*}{\lambda^* + w(k)} \prod_{i=0}^{k-1} \frac{w(i)}{\lambda^* + w(i)}, \\
\boldsymbol{\pi}_w(G) &:= \sum_{s\in\mathcal{S}(G)} \frac{\lambda^*}{\lambda^* + W(G)} \prod_{i=0}^{|G|-2} \frac{w(G, s, i+1)}{\lambda^* + W(G, s, i)}.
\end{aligned}$$

Theorem 3.2. *Consider a weight function w which satisfies condition* (M) *and let λ^* be defined as before. Then the following limits hold almost surely:*

(a) For any fixed $k \in \mathbb{N}$

$$\lim_{t\to\infty} \frac{|\{x \in \Upsilon(t) : \deg(x, \Upsilon(t)) = k\}|}{|\Upsilon(t)|} = \boldsymbol{p}_w(k).$$

(b) For any fixed $G \in \mathcal{G}$

$$\lim_{t\to\infty} \frac{\left|\{x \in \Upsilon(t) : \Upsilon(t)_{\downarrow x} = G\}\right|}{|\Upsilon(t)|} = \boldsymbol{\pi}_w(G).$$

(c) For any fixed $(G,u) \in \mathcal{G}^{(k)}$

$$\lim_{t\to\infty} \frac{\left|\{x \in \Upsilon(t) : (\Upsilon(t)_{\downarrow x^{(k)}}, x_{\downarrow k}) = (G,u)\}\right|}{|\Upsilon(t)|} = \boldsymbol{\pi}_w(G).$$

Furthermore, the functions $\boldsymbol{p}_w, \boldsymbol{\pi}_w$ *are probability distributions on* $\mathbb{N}$ *and* $\mathcal{G}$*, respectively, and* $\boldsymbol{\pi}_w$ *is steady (i.e. identity* (4) *holds).*

Remark 3.1. *Parts (a), (b) and (c) of Theorem 3.2, in turn, give more and more information about the asymptotic shape of the randomly growing tree* $\Upsilon(t)$*, as seen from a random vertex* ζ *chosen with uniform distribution. Part (a) identifies the a.s. limit as* $t \to \infty$*, of the degree distribution of* ζ*. Part (b) identifies the a.s. limit as* $t \to \infty$*, of the distribution of the progeny of* ζ*. Finally, part (c) does the same for the distribution of the progeny of the* k^{th} *ancestor of the randomly selected vertex with the position of this vertex marked.*

Remark 3.2. *From part (c) it is easy to derive the asymptotic distribution of the progeny of the* k^{th} *ancestor of the randomly selected vertex (as a rooted ordered tree without any marked vertices):*

$$\lim_{t\to\infty} \frac{\left|\{x \in \Upsilon(t) : \Upsilon(t)_{\downarrow x^{(k)}} = G\}\right|}{|\Upsilon(t)|} = \boldsymbol{\pi}_w(G) \left|G_{[k]}\right|.$$

The limit is the size-biased version of $\boldsymbol{\pi}_w(G)$*, with the biasing done by the size of the* k^{th} *generation.*

Remark 3.3. *Since the distribution* $\boldsymbol{\pi}_w$ *is steady, part (c) identifies the asymptotic distribution of the* whole family tree *of the randomly selected vertex* ζ *(relatives of arbitrary degree included). Hence asymptotically, as* $t \to \infty$*, the tree* $\Upsilon(t)$ *viewed from a random vertex* ζ *will have the following structure (we omit the precise formulation):*
– there exists an infinite path of ancestors $\zeta^1, \zeta^2, \zeta^3, \ldots$ *'going back in time',*
– we have finite ordered random trees rooted at each vertex of this path,
– the tree rooted at ζ^k *with the position of* ζ *marked on it has distribution* $\boldsymbol{\pi}_w^{(k)}$ *on* $\mathcal{G}^{(k)}$ *where* $\boldsymbol{\pi}_w^{(k)}(G,u) = \boldsymbol{\pi}_w(G)$*.*

4 Global properties

4.1 Questions, context

It is natural to pose the following question. Let us fix a vertex, say the first vertex in the first generation (first child of the root). What is the "limiting success level" of this vertex, compared to the other vertices in the same generation? What we mean by this is the number of descendants of this vertex, after a long time of tree evolution, compared to the number of descendants of its brothers.

Another formulation of the same question is to fix a vertex, let the tree grow for a long time, then choose a vertex uniformly at random from the big tree, and ask the probability that this random vertex is descendant of the fixed vertex. Clearly, if we look at these limiting probabilities for let us say the first generation, we get a distribution, itself being random, that codes an important information of the evolution of the tree.

If one looks at the system of these limiting (as time evolution of the tree tends to infinity) random distributions on the different generations of the tree, it is tempting to ask something about the limiting measure of this system, when letting the generation level tend to infinity. We will define the above concepts properly, and will denote this overall limiting measure by μ.

Having a random measure in our hand, which describes a global property of the limiting infinite system, it is natural to ask about the *Hausdorff (and packing) dimension of this measure*. On the other hand, the dimension of the measure depends on a parameter of the underlying metric, which is arbitrary. To rule out this (trivial) dependence, it is usual to ask about the *entropy of the limiting measure*, which depends on the growth process only. This is the natural equivalent of the dimension from a dynamical point of view.

The key to our results is a Markov process appearing naturally in the construction of a μ-typical leaf of the tree. After some discussion of the tree structure, the Markov property will be easy to see. Some technical difficulties will arise from the non-compactness of the state space.

The model choice is special in the sense that we only allow a finite degree for each vertex, but it is general in the sense that after having fixed the maximum number of children K a vertex may have, the weight function w, which determines the rule of attachment, can be any positive-valued function on $\{0, 1, \ldots, K-1\}$.

4.2 Additional notation and the choice of w

We restrict our weight function to the following class of functions.

We fix a positive integer $K > 2$, and we require the weight function to be zero above K:

$$w(k) = 0, \quad k \geq K. \tag{5}$$

As is clear from the model definition, see Section 2.2, this restriction makes sure that any vertex can have at most K children. The vertices of the random tree are labeled with elements of

$$\mathcal{N} := \bigcup_{n=0}^{\infty} \mathbb{I}^n,$$

as in (1), but now with the choice of $\mathbb{I} = \{1, 2, \ldots, K\}$. Since we require (5), the weight function automatically fulfils (M).

4.3 Limiting objects

Recall from (3) that the equation

$$\widehat{\rho}(\lambda) = 1$$

has a unique root $\lambda^* > 0$, the Malthusian parameter. This λ^* gives the rate of exponential growth of the tree size almost surely. The normalized size of the tree converges almost surely to a random variable, which we denote by

$$\Theta := \lim_{t\to\infty} e^{-\lambda^* t} |\Upsilon(t)| \ .$$

For every $x \in \mathcal{N}$, we introduce the variables Θ_x, corresponding to the growth of the subtree under x analogously, recall the notation (2) for subtrees,

$$\Theta_x := \lim_{t\to\infty} e^{-\lambda^*(t-\tau_x)} \, |\Upsilon_{\downarrow x}(t)| \ .$$

Clearly, for every $x \in \mathcal{N}$, the random variables Θ_x are identically distributed. The basic relation between the different Θ_x variables in the tree is that for any $x \in \mathcal{N}$,

$$\Theta_x = \sum_{i=1}^{K} e^{-\lambda^*(\tau_{xi}-\tau_x)} \Theta_{xi} \ ,$$

which is straightforward from $|\Upsilon_{\downarrow x}(t)| = 1 + \sum_{i=1}^{K} |\Upsilon_{\downarrow xi}(t)|$.

Now let us ask the following question. Fix a vertex $x \in \mathcal{N}$, and at time t, draw a vertex ζ_t uniformly randomly from $\Upsilon(t)$. What is the probability that ζ_t is descendant of x, so $x \prec \zeta_t$? As shown in (6) below, this probability tends to an almost sure limit Δ_x as $t \to \infty$, which can be expressed using the τ and Θ random variables,

$$\Delta_x := \lim_{t\to\infty} \frac{|\Upsilon_{\downarrow x}(t)|}{|\Upsilon(t)|} = e^{-\lambda^*\tau_x} \lim_{t\to\infty} \frac{e^{-\lambda^*(t-\tau_x)}|\Upsilon_{\downarrow x}(t)|}{e^{-\lambda^* t}|\Upsilon(t)|} = \frac{e^{-\lambda^*\tau_x}\Theta_x}{\Theta_\emptyset} \ . \tag{6}$$

We can now, for any $n \in \mathbb{N}$, define a random measure μ_n on the finite set $\mathbb{I}^n = \{x : |x| = n\}$, the n^{th} generation of the full tree, by

$$\mu_n(\{x\}) := \Delta_x \ .$$

This is a probability measure almost surely, which follows from the facts $\Delta_\emptyset = 1$ and $\Delta_y = \sum_{i=1}^{K} \Delta_{yi}$.

Let H_n denote the entropy of μ_n, that is

$$H_n = -\sum_{|x|=n} \Delta_x \log \Delta_x \ .$$

A measure as the limiting object for the tree

Let $\partial\mathcal{N}$ denote the set of leaves of the complete tree: $\partial\mathcal{N} = \{1, 2, \ldots, K\}^\infty$. The concatenation xy makes sense for $x \in \mathcal{N}$ and $y \in \partial\mathcal{N}$, and then $xy \in \partial\mathcal{N}$. Also, for $x \in \mathcal{N}$ and $z \in \partial\mathcal{N}$, we write $x \prec z$ if $\exists y \in \partial\mathcal{N}$ such that $z = xy$. For $x \in \mathcal{N}$ we denote the set of leaves under x by $\partial\mathcal{N}(x) = \{z \in \partial\mathcal{N} \,:\, x \prec z\}$.

Let $x_{|l}$ denote the first l letters of the string x, or in different words, let it denote the ancestor of x on the l-th level of the tree, and let $\partial\mathcal{N}$ be equipped with the usual metric

$$d(x, y) = \Lambda^{\max\{n\in\mathbb{N}\,:\,x_{|n}=y_{|n}\}} \,, \tag{7}$$

where $0 < \Lambda < 1$ is an arbitrary constant. This constant is often chosen to be $1/e$, which makes certain formulae appear simpler. Yet we will not fix the value, so that our formulae express the dependence of the studied quantities on this arbitrary choice.

With the help of the μ_n random limiting measures, we define μ on the cylinder sets $\partial\mathcal{N}(x)$ of $\partial\mathcal{N}$ by

$$\mu(\partial\mathcal{N}(x)) := \mu_n(\{x\}) = \Delta_x \;, \text{ if } |x| = n \;,$$

and then we extend μ from $\{\partial\mathcal{N}(x) \,:\, x \in \mathcal{N}\}$ to the sigma-algebra generated (on $\partial\mathcal{N}$). Our results concern the properties of this extended random measure μ.

4.4 Results

Theorem 4.1. *The limiting entropy*

$$h := \lim_{n\to\infty} \frac{1}{n} H_n$$

exists and is constant with probability one.

Theorem 4.2. *The Hausdorff dimension* $\dim_H \mu$ *and the packing dimension* $\dim_P \mu$ *of the measure* μ *are constant and equal with probability one, and* h *and the dimensions satisfy the relation*

$$\dim_H \mu = \dim_P \mu = \frac{h}{-\log\Lambda},$$

where Λ *is from (7). Moreover, the local dimension of* μ *equals* $\dim_H \mu = \dim_P \mu$ *at* μ*-almost every point.*

Theorem 4.3. *Furthermore, an explicit formula for* h *is given:*

$$h = \mathbf{E}\left(\sum_{i=1}^{K} \lambda^* \tau_i e^{-\lambda^* \tau_i}\right).$$

This can be computed given the weight function w.

References

[1] David Aldous. *Probability Approximations via the Poisson Clumping Heuristic.* Springer, 1989.

[2] David Aldous. Asymptotic fringe distributions for general families of random trees. *Ann. Appl. Probab.*, 1(2):228–266, 1991.

[3] Albert-László Barabási and Réka Albert. Emergence of scaling in random networks. *Science*, 286(5439):509–512, 1999.

[4] Itai Benjamini and Oded Schramm. Recurrence of distributional limits of finite planar graphs. *Electronic J. Probab.*, 6(paper 23):1–13, 2001.

[5] Julien Berestycki. Multifractal spectra of fragmentation processes. *Journal of Statistical Physics*, 113(3):411–430, 2003.

[6] Jean Bertoin. *Random fragmentation and coagulation processes.* Cambridge Univ Pr, 2006.

[7] J. D. Biggins. Martingale convergence in the branching random walk. *Journal of Applied Probability*, 14(1):25–37, 1977.

[8] Béla Bollobás, Oliver Riordan, Joel Spencer, and Gábor Tusnády. The degree sequence of a scale-free random graph process. *Random Structures Algorithms*, 18(3):279–290, 2001.

[9] Béla Bollobás and Oliver M. Riordan. Mathematical results on scale-free random graphs. In *Handbook of graphs and networks*, pages 1–34. Wiley-VCH, Weinheim, 2003.

[10] Fan Chung, Shirin Handjani, and Doug Jungreis. Generalizations of Polya's urn problem. *Ann. Comb.*, 7(2):141–153, 2003.

[11] Steffen Dereich and Peter Mörters. Random networks with sublinear preferential attachment: degree evolutions. *Electron. J. Probab.*, 14(43):1222–1267, 2009.

[12] Rui Dong, Christina Goldschmidt, and James B. Martin. Coagulation-fragmentation duality, Poisson-Dirichlet distributions and random recursive trees. *Ann. Appl. Probab.*, 16(4):1733–1750, 2006.

[13] Joseph Leo Doob. *Stochastic Processes.* Wiley, 1953.

[14] Michael Drmota. *Random trees.* SpringerWienNewYork, Vienna, 2009. An interplay between combinatorics and probability.

[15] T. Duquesne. Packing and Hausdorff measures of stable trees. *Lévy Matters I*, pages 93–136, 2010.

[16] Thomas Duquesne and Jean-Franois Le Gall. Probabilistic and fractal aspects of Lévy trees. *Probability Theory and Related Fields*, 131:553–603, 2005.

[17] Richard Durrett. *Random Graph Dynamics*. Cambridge University Press, Cambridge, 2007. Cambridge Series in Statistical and Probabilistic Mathematics.

[18] Kenneth Falconer. *Techniques in Fractal Geometry*. Wiley, 1997.

[19] B. Haas and G. Miermont. The genealogy of self-similar fragmentations with negative index as a continuum random tree. *Electronic Journal of Probability*, 9(paper 4):57, 2004.

[20] B. Haas and G. Miermont. Scaling limits of Markov branching trees, with applications to Galton-Watson and random unordered trees. 2010. Arxiv preprint, `http://arxiv.org/abs/1003.3632`.

[21] B. Haas, G. Miermont, J. Pitman, and M. Winkel. Continuum tree asymptotics of discrete fragmentations and applications to phylogenetic models. *The Annals of Probability*, 36(5):1790–1837, 2008.

[22] Peter Jagers. *Branching processes with biological applications*. Wiley-Interscience [John Wiley & Sons], London, 1975. Wiley Series in Probability and Mathematical Statistics —Applied Probability and Statistics.

[23] Peter Jagers and Olle Nerman. The growth and composition of branching populations. *Adv. in Appl. Probab.*, 16(2):221–259, 1984.

[24] H. Kesten and B. P. Stigum. A limit theorem for multidimensional Galton-Watson processes. *Ann. Math. Statist.*, 37:1211–1223, 1966.

[25] P. L. Krapivsky and S. Redner. Organization of growing random networks. *Phys. Rev. E*, 63(6):066123, May 2001.

[26] P. L. Krapivsky, S. Redner, and F. Leyvraz. Connectivity of growing random networks. *Phys. Rev. Lett.*, 85(21):4629–4632, Nov 2000.

[27] Norbert Kusolitsch. Why the theorem of Scheffé should be rather called a theorem of Riesz. *Period. Math. Hungar.*, 61(1-2):225–229, 2010.

[28] Jean-François Le Gall. Processus de branchement, arbres et superprocessus. In *Development of mathematics 1950–2000*, pages 763–793. Birkhäuser, Basel, 2000.

[29] László Lovász. *Very large graphs*. 2008. Arxiv preprint, `http://arxiv.org/abs/0902.0132`.

[30] R. Lyons. A simple path to Biggins' martingale convergence for branching random walk. In K.B. Athreya and P. Jagers, editors, *Classical and modern branching processes*, The IMA volumes in mathematics and its applications. Springer, 1997.

[31] R. Lyons, R. Pemantle, and Y. Peres. Conceptual proofs of l log l criteria for mean behavior of branching processes. *The Annals of Probability*, 23(3):1125–1138, 1995.

[32] T. F. Móri. On random trees. *Studia Sci. Math. Hungar.*, 39(1-2):143–155, 2002.

[33] T. F. Móri. The maximum degree of the Barabási–Albert random tree. *Comb. Probab. Computing*, 14:339–348, 2005.

[34] T. F. Móri. A surprising property of the Barabási-Albert random tree. *Studia Sci. Math. Hungar.*, 43:265–273, 2006.

[35] Olle Nerman. On the convergence of supercritical general (C-M-J) branching processes. *Z. Wahrsch. Verw. Gebiete*, 57(3):365–395, 1981.

[36] Roberto Oliveira and Joel Spencer. Connectivity transitions in networks with superlinear preferential attachment. *Internet Math.*, 2(2):121–163, 2005.

[37] Peter Olofsson. The $x \log x$ condition for general branching processes. *J. Appl. Probab.*, 35(3):537–544, 1998.

[38] Boris Pittel. Note on the heights of random recursive trees and random m-ary search trees. *Random Struct. Alg.*, 5(2):337–347, 1994.

[39] Anna Rudas. Random tree growth with general weight function. 2004. Arxiv preprint, `http://arxiv.org/abs/math/0410532`.

[40] Anna Rudas and Bálint Tóth. Random tree growth with branching processes - a survey. In B Bollobás, R Kozma, and D Miklós, editors, *Handbook of Large-Scale Random Networks*, volume 18 of *Bolyai Society Mathematical Studies*, chapter 4. Springer, 2009.

[41] Anna Rudas, Bálint Tóth, and Benedek Valkó. Random trees and general branching processes. *Random Struct. Algorithms*, 31(2):186–202, 2007.

[42] Anna Rudas and Imre Péter Tóth. Entropy and Hausdorff dimension in random growing trees. *Stochastics and Dynamics*, Accepted for publication: 2011. 12. 15. DOI Number: 10.1142/S0219493712500104.

[43] Henry Scheffé. A useful convergence theorem for probability distributions. *Ann. Math. Statist.*, 18(3):434–438, 1947.

[44] Robert T. Smythe and Hosam M. Mahmoud. A survey of recursive trees. *Teor. Ĭmovīr. Mat. Stat.*, (51):1–29, 1994.

[45] G. Udny Yule. A mathematical theory of evolution, based on the conclusions of Dr. J. C. Willis, F.R.S. *Royal Society of London Philosophical Transactions Series B*, 213:21–87, 1925.

[46] Remco van der Hofstad. *Random Graphs and Complex Networks.* 2011, `http://www.win.tue.nl/~rhofstad/NotesRGCN2011.pdf`. in preparation.