Scaling limits of large random trees

Lecture notes of the Simposio de Probabilidad y Procesos Estocásticos

16 – 20 novembre 2015, Mérida, Yucatán

Bénédicte Haas,[*] Université Paris 13

Abstract

The goal of these lectures is to survey some of the recent progress on the description of large-scale structure of random trees. We will use the framework of Markov branching sequences of trees and develop several applications, to combinatorial trees, Galton-Watson trees, some dynamical models of randomly growing trees, cut trees, etc.

This is a rough draft – to be completed – all comments are welcome!

Contents

---

[*]E-mail: haas@math.univ-paris13.fr

1   Introduction

The main goal of these lectures is to survey some of the recent progress on the description of large-scale structure of random trees. Describing the structure of large (random) trees, and more generally large graphs, is an important goal of modern probabilities and combinatorics. The question we will typically be interested in is the following. For $(T_n, n \geq 1)$ a sequence of random trees, where for each n T_n is a tree of "size" n (the size of a tree may be its number of vertices, or its number of leaves; other variants are possible): does there exist a deterministic sequence $(a_n, n \geq 1)$ and a "continuous" random tree $\mathcal{T}$ such that

$$\frac{T_n}{a_n} \underset{n \to \infty}{\longrightarrow} \mathcal{T} \ ?$$

To make sense of this question, we will view T_n as a metric space by "replacing" its edges with segments of length 1, and then use the notion of Gromov-Hausdorff distance to compare compact metric spaces. When such a convergence holds, the continuous limit highlights some properties of the discrete objects that approximate it, and vice-versa. Such a result should be compared with the convergence of rescaled sums of i.i.d. random variables towards the normal distribution and its functional analog, the convergence of random walks towards the Brownian motion.

As a first example, consider T_n a tree picked uniformly at random in the set of trees with n vertices, labelled by $\{1, \ldots, n\}$. The tree T_n has to be understood as a typical representative of this set of trees. In this case the answer to the previous question dates back to a series of works by Aldous in the beginning of the 90's [3, 4, 5]: Aldous showed that

$$\frac{T_n}{2\sqrt{n}} \underset{n \to \infty}{\overset{(d)}{\longrightarrow}} \mathcal{T}_{\mathrm{Br}} \tag{1}$$

where the limiting tree is called the Brownian Continuum Random Tree. This result has various interesting consequences, e.g. its gives the asymptotics of the diameter, the height (if we consider rooted versions of the trees) and several other statistics related to the tree T_n. Consequently it also gives the asymptotic proportion of trees with n labelled vertices that have a diameter larger than $x\sqrt{n}$, etc. Some of these questions on statistics of uniform trees were already treated in previous works, the strength of Aldous' result is that is describes the asymptotic of the *whole* tree T_n. Aldous actually established the result (1) in a much broader context, that of *conditioned Galton–Watson trees* with finite variance. In this situation, T_n is the genealogical tree of a Galton-Watson process (with a given, fixed offspring distribution with finite variance)

conditioned on having a total number of vertices equal to $n, n \geq 1$. It turns out that the above sequence of uniform labelled trees can be seen as a sequence of conditioned Galton-Watson trees (when the offspring distribution is a Poisson distribution) and more generally that several sequences of *combinatorial* trees reduce to conditioned Galton-Watson trees. In the early 2000s, Duquesne [19] extended Aldous's result to conditioned Galton-Watson trees with an offspring distribution in the domain of attraction of a stable law. In most of these cases the scaling sequences (a_n) are much smaller $a_n \ll \sqrt{n}$, and a family of continuous trees arises in the limit: the so-called family of *stable Lévy trees*. These results on conditioned Galton-Watson trees are now very well–known and well–established and have a lot of applications in the study of large random graphs (e.g. Erdös-Rényi random graph, large random maps, etc.).

In these lectures, we will focus on a more recent point of view, that of sequences of random trees that satisfy a certain *Markov branching property*, which appears naturally in a large set of models and includes conditioned Galton–Watson trees. This Markov branching property is a sort of discrete fragmentation property which roughly says that the subtrees above some given height are independent with a law that depends only on their total size. Under appropriate assumptions, we will see that these sequences of trees, suitably rescaled, converge to continuous trees that belong to the class of so-called *self-similar fragmentation trees*. This class of trees is related to the self-similar fragmentation processes introduced and studied by Bertoin in the 2000s [8], which are models used to describe the evolution of objects that randomly split as times passes.

In Section 2 below, we will start with a series of definitions related to discrete trees and then discuss several examples of sequences of trees $(T_n, n \geq 1)$. We will also introduce there the notion of Markov branching sequences of trees. We will then present in Section 3 the results of Aldous [4] and Duquesne [19] on conditioned Galton-Watson trees, after having introduced the notions of real trees and Gromov-Hausdorff topology, in order to settle the topological framework in which we will work. Section 4 is the core of these lectures. We present there the main results on the scaling limits of Markov branching trees, and give the main ideas of the proofs. Various applications will then be discussed in Section 5: to conditioned Galton-Watson trees with arbitrary degree constraints, to Pólya trees, to several examples of dynamical models of tree growth and to the cut-trees of Cayley trees.

The main results on Markov branching trees presented here were developed in the paper [30], which has its roots in the earlier paper [31]. Several applications have been developed in these papers, and in more recent works. One objective of these notes is to survey and regroup all these results.

All the trees we will encounter here will have a power growth. We should emphasize that there is however a large set of random trees that naturally arise in applications and that do not have such a behavior. In particular, many models of trees arising in the analysis of algorithms have a logarihmic growth.

2 Discrete trees, examples and motivations

2.1 Discrete trees

A **discrete tree** (or graph-theoretic tree) is a finite or countable graph (V, E) that is connected and without cycle. Here V denotes the set of vertices of the graph and E its set of edges.

Note that two vertices are then connected by exactly one path and that $\#V = \#E + 1$ when the tree is finite.

A discrete tree can be seen as a metric space, when endowed with the **graph distance** d_{gr}: given two vertices $u, v \in T$, $d_{\mathrm{gr}}(u, v)$ is defined as the number of edges of the path from u to v.

A **rooted tree** (T, ρ) is a tree T with a vertex ρ that has been distinguished. The vertex ρ is then called the **root** of T. This gives a genealogical structure to the tree. The root corresponds to the generation 0, its neighbors can be interpreted as its children and form the generation 1, the children of its children form the generation 2, etc. We will usually call the **height** of a vertex its generation, and denote it by $\mathsf{ht}(v)$ (the height of a vertex is therefore its distance to the root). The height of the tree is then

$$\mathsf{ht}(T) = \sup_{v \in T} \mathsf{ht}(v).$$

and its diameter

$$\mathsf{diam}(T) = \sup_{u, v \in T} d_{\mathrm{gr}}(u, v).$$

In these lectures, we will mainly work with rooted trees.

The **degree of a vertex** $v \in T$ is the number of connected components obtained when removing v (in other words, it is the number of neighbors of v). A vertex $v \neq \rho$ and with degree 1 is called a **leaf**. A (strict) binary tree is a rooted tree where all vertices but the leaves have two children, i.e. the root has degree 2 and all other non-leaf vertices have degree 3.

In the context of rooted trees, one may need to order the children of the root, and then, recursively, the children of each vertex in the tree. This gives an **ordered** (or planar) tree. Formally, we generally see such a tree as a subset of the infinite Ulam-Harris tree

$$\mathcal{U} = \bigcup_{n=0}^{\infty} \mathbb{N}^n$$

where $\mathbb{N} := \{1, 2 \ldots\}$ and $\mathbb{N}^0 = \{\varnothing\}$. The element $\varnothing$ is the root of the Ulam-Harris tree, and all other $u = u_1 u_2 \ldots u_n \in \mathcal{U} \backslash \{\varnothing\}$ is connected to the root via the unique shortest path

$$\varnothing \to u_1 \to u_1 u_2 \to \ldots \to u_1 \ldots u_n.$$

The height (or generation) of such a sequence u is therefore its length, n. We then say that $T \subset \mathcal{U}$ is a (finite or infinite) rooted *ordered* tree if:

- $\varnothing \in T$
- if $u = u_1 \dots u_n \in T \backslash \{\varnothing\}$, then $u = u_1 \dots u_{n-1} \in T$ (the parent of an individual in T that is not the root is also in T)
- if $u = u_1 \dots u_n \in T$, there exists an integer $c_u(T) \geq 0$ such that the element $u_1 \dots u_n j \in T$ if and only if $1 \leq j \leq c_u(T)$.

The number $c_u(T)$ represents the number of children of u in T.

We will also sometimes consider (unordered) **labelled** trees. In these cases, the vertices are labelled, typically by $\{1, \dots, n\}$ if there are n vertices. Whereas in an unlabelled tree, the vertices are indistinguishable. Partial labelling is also possible, e.g. by labelling only the leaves of the tree.

In the following we will always specified when a tree is labelled and/or ordered. When not specified, it is implicitly unlabelled, unordered (and rooted).

2.2 First examples

We now present some classical models of random trees. Our goal will be to describe their asymptotic behavior when the size of the tree grows, as discussed in the Introduction.

Combinatorial trees. Let $\mathbb{T}_n$ be a set of trees of n vertices, all sharing some structural properties. E.g. $\mathbb{T}_n$ may be the set of all rooted trees with n vertices, or the set of all ordered rooted trees with n vertices, or the set of all binary trees with n vertices, etc. Pick a tree uniformly at random in this set and denote it by T_n. This tree is viewed as a "typical element" of $\mathbb{T}_n$ and its global behavior will represent some of the features shared by most of the trees. For example, if the probability that the height of T_n is larger than $n^{\frac{1}{2}+\varepsilon}$ tends to 0 as $n \to \infty$, this means that the proportion of trees in the set that have a height larger than $n^{\frac{1}{2}+\varepsilon}$ is asymptotically negligeable. We will more specifically be interested in the following cases:

- T_n is a uniform rooted tree with n vertices (i.e. a uniform Pólya tree)
- T_n is a uniform rooted ordered tree with n vertices
- T_n is a uniform tree with n labelled vertices (i.e. a uniform Cayley tree)
- T_n is a uniform rooted ordered binary tree with n vertices
- T_n is a uniform rooted binary tree with n vertices, etc.

Similarly one may consider trees picked uniformly amongst sets of trees with a given structure and n *leaves*, or more general degree constraints.

Galton–Watson trees. A Galton-Watson tree is a tree describing the genealogical structure of a Galton-Watson process. Let η be a probability on $\mathbb{Z}_+ - \eta$ is called the *offspring distribution*

– and let $m := \sum_{i \geq 1} i\eta(i) \in [0, \infty]$ denote its mean. We will always assume that $\eta(1) < 1$ in order to avoid the trivial case where each individual has a unique child. A η-Galton–Watson tree τ is usually seen as an ordered rooted tree and defined as follows (recall the Ulam-Harris notation $\mathcal{U}$):

- $c_\varnothing(\tau)$ is distributed according to η

- conditionally on $c_\varnothing(\tau) = p$, the p ordered subtrees $\tau_i = \{u \in \mathcal{U} : iu \in \tau\}$ descending from $i = 1, \ldots, p$ are independent and distributed as τ.

From this construction, one sees that the distribution of τ is given by:

$$\mathbb{P}(\tau = \mathsf{t}) = \prod_{v \in \mathsf{t}} \eta_{c_v(\mathsf{t})} \tag{2}$$

for all rooted ordered tree t. This definition of Galton–Watson trees as ordered trees is the simplest, avoiding any symmetry problems. However in the following we will mainly see these trees as metric spaces, which means that we can "forget the order".

Clearly, if we call Z_k the number of individuals at height k, then $(Z_k, k \geq 1)$ is a Galton–Watson process starting from $Z_0 = 1$. It is well-known that the extinction time of this process,

$$\inf\{k \geq 0 : Z_k = 0\}$$

if finite with probability 1 when $m \leq 1$ and with a probability $\in [0, 1)$ when $m > 1$. The offspring distribution η and the tree τ are said to be *subcritical* when $m < 1$, *critical* when $m = 1$ and *supercritical* when $m > 1$. From now on, we assume that

$$m = 1$$

and for integers n s.t. $\mathbb{P}(\#\tau = n) > 0$, we let T_n^η denote a *non-ordered* version of the Galton–Watson tree τ conditioned to have n vertices. Sometimes, we will need to keep the order and we will let $T_n^{\eta,\mathrm{ord}}$ denote this ordered conditioned version.

It turns out that conditioned Galton–Watson trees are closely related to combinatorial trees. Indeed, one can check with (2) that:

- if $\eta \sim \mathrm{Geo}(1/2)$, $T_n^{\eta,\mathrm{ord}}$ is uniform amongst the set of rooted ordered trees with n vertices

- if $\eta \sim \mathrm{Poisson}(1)$, T_n^η is uniform amongst the set of rooted trees with n labelled vertices

- if $\eta \sim \frac{1}{2}(\delta_0 + \delta_2)$, $T_n^{\eta,\mathrm{ord}}$ is uniform amongst the set of rooted ordered binary trees with n vertices.

We refer to Aldous [4] for details and additional examples.

Hence, studying the large scaled structure of Galton-Watson trees will also lead to results in the context of combinatorial trees. As mentioned in the Introduction, the description of large conditioned Galton-Watson trees is now very well known. It has been initiated by Aldous [3, 4, 5]

and then expanded by Duquesne [19], relying on previous works by Le Gall and Le Jan [38] and by Duquesne and Le Gall [21]. This will be reviewed in Section 3. However, there are also some natural sequences of combinatorial trees that *cannot* be reinterpreted as Galton-Watson trees, starting with the example of the uniform rooted tree with n vertices. Studying the asymptotics of this tree remained open for a while, because of the absence of symmetry properties.

In another direction, one may also wonder what happens when considering versions of Galton-Watson trees conditioned to have n leaves, instead of n vertices, or more generally a given number of vertices in a given subset $A \subset \mathbb{Z}_+$.

Dynamical models of tree growth. We now consider several sequences of finite rooted trees $(T_n, n \geq 1)$ that are built recursively by adding at each step new edges on the pre-existing tree. We start with a well-known algorithm that Rémy [47] introduced to generate uniform binary trees with n leaves.

Rémy's algorithm. The sequence $(T_n(\mathrm{R}), n \geq 1)$ is constructed recursively as follows:

- Step 1: $T_1(\mathrm{R})$ is the tree with one edge and two vertices: one root, one leaf
- Step n: given $T_{n-1}(\mathrm{R})$, choose uniformly at random one of its edges and graft on "its middle" one new edge-leaf, that is split the selected edge into two so as to obtain two edges separated by a new vertex, and then add a new edge-leaf to the new vertex. This gives $T_n(\mathrm{R})$.

It turns out (see e.g. [40]) that the tree $T_n(\mathrm{R})$, to which has been subtracted the edge between the root and the first branch point, is distributed as a binary critical Galton-Watson tree conditioned to have $2n - 1$ vertices, or equivalently n leaves (after forgetting the order in GW-tree). As so, we will deduce its asymptotic behavior from that of Galton-Watson trees. However this model can be extended in several directions, most of which are not related to Galton-Watson trees. We detail three of them.

Ford's α-model [27]. Let $\alpha \in [0, 1]$. We construct a sequence $(T_n(\alpha), n \geq 1)$ by modifying Rémy's algorithm as follows:

- Step 1: $T_1(\alpha)$ is the tree with one edge and two vertices: one root, one leaf
- Step n: given $T_{n-1}(\alpha)$, give a weight $1 - \alpha$ to each edge connected to a leaf, and α to all other edges (the internal edges). The total weight is $n - \alpha$. Now choose an edge at random with a probability proportional to its weight and graft on "its middle" one new edge-leaf. This gives $T_n(\alpha)$.

Note that when $\alpha = 1/2$ the weights are identical on all edges and we recover Rémy's algorithm. When $\alpha = 0$, the new edge is always graft uniformly on an edge-leaf, hence $T_n(0)$ is a *Yule tree* with n leaves. When $\alpha = 1$, we obtain a deterministic tree called the *comb tree*. This α-family of sequences of trees was introduced by Ford [27] in order to interpolate between the Yule, the uniform and the comb models. His goal was to propose new models for phylogenetic trees.

k-ary growing trees [34]. Here is another extension of Rémy's algorithm, where now several edges are added at each step. Consider an integer $k \geq 2$. The sequence $(T_n(k), n \geq 1)$ is constructed recursively as follows:

- Step 1: $T_1(k)$ is the tree with one edge and two vertices: one root, one leaf
- Step n: given $T_{n-1}(k)$, choose uniformly at random one of its edges and graft on "its middle" $k-1$ new edges–leaf. This gives $T_n(k)$.

When $k = 2$, we recover Rémy's algorithm. For larger k, there is no connection with Galton–Watson trees.

Marginals of stable trees – Marchal's algorithm [41]. We can also attribute weights to the vertices. Fix a parameter $\beta \in (1, 2]$ and construct the sequence $(T_n(\beta), n \geq 1)$ as follows:

- Step 1: $T_1(\beta)$ is the tree with one edge and two vertices: one root, one leaf
- Step n: given $T_{n-1}(\beta)$, attribute the weight
  - $\beta - 1$ on each edge
  - $d - 1 - \beta$ on each vertex of degree $d \geq 3$.

  The total weight is $n\beta - 1$. Then select at random an edge or vertex with a probability proportional to its weight and graft on it a new edge-leaf. This gives $T_n(\beta)$.

Note that when $\beta = 2$, vertices of degree 3 are never selected (their weight is 0). So the trees $T_n(\beta), n \geq 1$ are all binary, and we recover Rémy's algorithm.

Of course, several other extensions may be considered. In [15], Chen, Ford and Winkel consider models generalizing both Ford's and Marchal's algorithms, where the weights are attributed to vertices and edges, with possibly different weights on the edges. However, models with totally arbitrary weights are difficult to study. All the sequences mentioned above share a nice property, the Markov branching property, that will studied in the next section. We also mention that in a work in preparation [33], models where finite trees (instead of a group of $k-1$ edges) are grafted at each step on a uniformly selected edge are studied.

Remark. In these dynamical models of tree growth, we build *on a same probability space* the sequence of trees, contrary to the examples of Galton-Watson trees or combinatorial trees that give sequences of *distributions* of trees. In this situation, one may expect to have more than a convergence in distribution for the rescaled sequence of trees. We will see that it is indeed the case.

2.3 Markov branching trees

Markov branching trees were introduced by Aldous [6] as a class of random binary trees for phylogenetic models and later extended to non-binary cases in Broutin & al. [14], and Haas

& al. [31]. It turns out that many natural models of sequence of trees satisfy the **Markov branching property** (**MB– property** for short), starting with the example of conditioned Galton–Watson trees and most of the examples of the previous section.

Consider

$$(T_n, n \geq 1)$$

a sequence of trees where T_n is a rooted (unordered, unlabelled) tree with n *leaves*. The MB–property is a property of the sequence of *distributions* $(\mathcal{L}(T_n), n \geq 1)$. Informally, the MB–property says that for each n, *given that*

"the root of T_n splits in p subtrees with respectively $\lambda_1 \geq \ldots \geq \lambda_p$ leaves",

T_n is distributed as the tree obtained by gluing on a common root p independent trees with respective distributions those of $T_{\lambda_1}, \ldots, T_{\lambda_p}$. Such a sequence of trees is therefore characterized by a sequence of probabilities $(q_n, n \geq 1)$, where for each n, q_n is a probability on the set of partitions of the integer n: if $n \geq 2$,

$$\mathcal{P}_n := \left\{ \lambda = (\lambda_1, \ldots, \lambda_p), \lambda_i \in \mathbb{N}, \lambda_1 \geq \ldots \geq \lambda_p \geq 1 : \sum_{i=1}^{p} \lambda_i = n \right\}.$$

For a partition $\lambda \in \mathcal{P}_n$, we denote by $p(\lambda)$ its length, i.e. the number of terms in the sequence λ. If $n = 1$, $\mathcal{P}_1 = \{(1), \emptyset\}$ by convention (we need to have a cemetery point). The probability q_n determines how the n leaves of T_n are distributed into the subtrees above the root of T_n. We call such a probability a *splitting distribution*. The reason why we add an extra partition $\emptyset$ in $\mathcal{P}_1$ is that there are two types of trees with one leaf: (a) trees with more than two vertices, in which cases the root has one subtree, with one leaf (this corresponds to the partition (1)) and (b) the trees with is reduced to a unique vertex, with is both the leaf and the root, in which case the root has 0 subtree (this corresponds to the partition $\emptyset$).

In order that effective splittings occur, we will always assume that

$$q_n((n)) < 1, \quad \forall n \geq 1.$$

We now turn to a precise definition of the MB–property. In that aim, we need first to define a notion of *gluing* of trees. Consider $t_1, \ldots, t_p$, p discrete rooted (unordered) trees. Informally, we want to glue them on a same common root in order to form a tree $\langle t_1, \ldots, t_p \rangle$ whose root splits into the p subtrees $t_1, \ldots, t_p$. Formally, this can be done as follows. We first consider ordered versions of the trees $t_1^{\mathrm{ord}}, \ldots, t_p^{\mathrm{ord}}$ seen as subsets of the Ulam-Harris tree $\mathcal{U}$ and then define a new ordered tree by

$$\langle t_1^{\mathrm{ord}}, \ldots, t_p^{\mathrm{ord}} \rangle := \{\varnothing\} \cup_{i=1}^{p} i t_i^{\mathrm{ord}}.$$

The tree $\langle t_1, \ldots, t_p \rangle$ is then defined as the unordered version of $\langle t_1^{\mathrm{ord}}, \ldots, t_p^{\mathrm{ord}} \rangle$.

Definition 2.1. *Consider* $(q_n, n \geq 1)$ *a sequence of probabilities, with* q_n *a probability on* $\mathcal{P}_n$ *such that* $q_n((n)) < 1 \ \forall n \geq 1$. *We construct recursively a sequence of distributions of trees* $(\mathcal{L}_n^{\mathbf{q}}, n \geq 1)$ *as follows:*
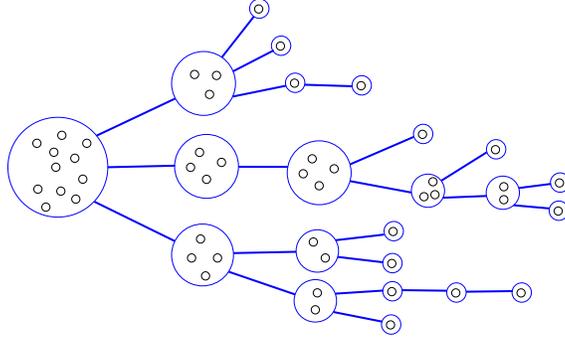
9

Figure 1: A sample tree T_{11}. The first splitting arises with probability $q_{11}(4, 4, 3)$.

(i) $\mathcal{L}_1^{\mathbf{q}}$ is the distribution of a line–tree with $G+1$ vertices and G edges where G is a geometric distribution:

$$\mathbb{P}(G = k) = p_1(\emptyset)(1 - p_1(\emptyset))^k.$$

(ii) For $n \geq 2$, $\mathcal{L}_n^{\mathbf{q}}$ is the distribution of

$$\langle T_1, \ldots, T_{p(\Lambda)} \rangle$$

where $\Lambda \sim p_n$ and given Λ, $T_1, \ldots, T_{p(\Lambda)}$ are independent trees with respective distributions $\mathcal{L}_1^{\mathbf{q}}, \ldots, \mathcal{L}_{p(\Lambda)}^{\mathbf{q}}$.

A sequence $(T_n, n \geq 1)$ of random rooted trees such that $T_n \sim \mathcal{L}_n^{\mathbf{q}}$ for each $n \in \mathbb{N}$ is called a MB–sequence of trees *indexed by the leaves*, with splitting distributions $(q_n, n \geq 1)$.

This construction may be re-interpreted as follows: We start from a collection of n *indistinguishable* balls, and with probability $q_n(\lambda_1, \ldots, \lambda_p)$, split the collection into p sub-collections with $\lambda_1, \ldots, \lambda_p$ balls. Note that there is a chance $q_n((n)) < 1$ that the collection remains unchanged during this step of the procedure. Then, re-iterate the splitting operation independently for each sub-collection using this time the probability distributions $q_{\lambda_1}, \ldots, q_{\lambda_p}$. If a sub-collection consists of a single ball, it can remain single with probability $q_1((1))$ or get wiped out with probability $q_1(\emptyset)$. We continue the procedure until all the balls are wiped out. The tree T_n is then the genealogical tree associated with this process: it is rooted at the initial collection of n balls and its n leaves correspond to the n isolated balls just before they are wiped out, See Figure 1 for an illustration.

Similarly, we will consider a MB–property for sequences of (distributions) of trees indexed by their number of vertices. Let here

$$(T_n, n \geq 1)$$

be a sequence of trees where T_n is a rooted (unordered) tree with n *vertices*. Here we consider a sequence of probabilities $(p_n, n \geq 1)$ with p_n a probability on $\mathcal{P}_n$ with no restriction but

$$p_1((1)) = 1.$$

10

Mimicking the previous balls in urns construction, we start from a collection of n indistinguishable balls, we first remove a ball, split the $n-1$ remaining balls in sub-collections with $\lambda_1, \ldots, \lambda_p$ balls with probability $p_{n-1}((\lambda_1, \ldots, \lambda_p))$, and iterate independently on sub-collections until no ball remains. Formally, this gives:

Definition 2.2. *Consider $(p_n, n \geq 1)$ a sequence of probabilities, with p_n a probability on $\mathcal{P}_n$ $\forall n \geq 1$ and $p_1(1) = 1$. We construct recursively a sequence of distributions of trees $(\mathcal{L}_n^{\mathbf{P}}, n \geq 1)$ as follows:*

(i) *$\mathcal{L}_1^{\mathbf{P}}$ is the deterministic distribution of the tree reduced to one vertex.*

(ii) *For $n \geq 2$, $\mathcal{L}_n^{\mathbf{P}}$ is the distribution of*

$$\langle T_1, \ldots, T_{p(\Lambda)} \rangle$$

where $\Lambda \sim p_{n-1}$ and given Λ, $T_1, \ldots, T_{p(\Lambda)}$ are independent trees with respective distributions $\mathcal{L}_1^{\mathbf{P}}, \ldots, \mathcal{L}_{p(\Lambda)}^{\mathbf{P}}$.

A sequence $(T_n, n \geq 1)$ of random rooted trees such that $T_n \sim \mathcal{L}_n^{\mathbf{P}}$ for each $n \in \mathbb{N}$ is called a MB–sequence of trees indexed by the vertices, with splitting distributions $(p_n, n \geq 1)$.

More generally, the MB-property can be extended to sequences of trees $(T_n, n \geq 1)$ with arbitrary degree restriction, i.e. such that for all n, T_n has n vertices in A, where A is a given subset of $\mathbb{Z}$. We will not develop this here and refer the interested reader to [48] for details.

Some examples.

1. **A deterministic example.** Consider the splitting distributions on $\mathcal{P}_n$

$$q_n(\lceil n/2 \rceil, \lfloor n/2 \rfloor) = 1, \quad n \geq 1$$

and $(T_n, n \geq 1)$ the corresponding MB-sequence indexed by leaves. Then T_n is a deterministic discrete binary tree, whose root splits in two subtrees with both $n/2$ leaves when n is even, and respectively $(n+1)/2, (n-1)/2$ leaves when n is odd. Clearly, when $n = 2^k$, the height of T_n is exactly k, and more generally for large n, $\mathrm{ht}(T_n) \sim \ln(n)/\ln(2)$.

2. **A basic example.** Let q_n be the probability on $\mathcal{P}_n$ defined by

$$q_n((n)) = 1 - \frac{1}{n^\gamma} \quad \text{and} \quad q_n(\lceil n/2 \rceil, \lfloor n/2 \rfloor) = \frac{1}{n^\gamma} \quad \text{for some } \gamma > 0,$$

and let $(T_n, n \geq 1)$ be a MB-sequence indexed by leaves with splitting distributions (q_n). Then T_n is a discrete tree with vertices with degrees $\in \{1, 2, 3\}$ where the distance between the root and the first branching point (i.e. the first vertex of degree 3) is a Geometric distribution on $\mathbb{Z}_+$ with success parameter $n^{-\gamma}$. The two subtrees above this branching point are then independent subtrees, independent of the Geometric r.v. just mentioned, and whose respective distances between the root and first branching point are Geometric

11

distributions with respectively $(\lceil n/2 \rceil)^{-\gamma}$ and $(\lfloor n/2 \rfloor)^{-\gamma}$ parameters. Noticing the weak convergence

$$\frac{\mathrm{Geo}(n^{-\gamma})}{n^\gamma} \xrightarrow[n\to\infty]{} \mathrm{Exp}(1)$$

one may expect that $n^{-\gamma}T_n$ has a limit in distribution. We will later see that it is indeed the case.

3. **Galton–Watson trees**. Let T_n^η be a Galton–Watson tree with offspring distribution η, conditioned on having n leaves, for integers n for which this is possible. The branching property is then preserved by conditioning and the sequence $(T_n^\eta, n : \mathbb{P}(\#_{\mathrm{leaves}}T^\eta) > 0)$ is Markov branching, with splitting distributions

$$q_n^{\mathrm{GW}-\eta}(\lambda) = \eta(p) \times \frac{p!}{\prod_{i=1}^{p} m_i(\lambda)!} \times \frac{\prod_{i=1}^{p} \mathbb{P}(\#_{\mathrm{leaves}}T^\eta = \lambda_i)}{\mathbb{P}(\#_{\mathrm{leaves}}T^\eta = n)}$$

for all $\lambda \in \mathcal{P}_n$, where $\#_{\mathrm{leaves}}T^\eta$ is the number of leaves of the unconditioned GW tree T^η, and $m_i(\lambda) = \#\{1 \le j \le p : \lambda_j = i\}$.

Similarly, if T_n^η now denotes a Galton–Watson tree with offspring distribution η, conditioned on having n *vertices*, the sequence $(T_n^\eta, \mathbb{P}(\#_{\mathrm{vertices}}T^\eta) > 0)$ is Markov branching, with splitting distributions

$$p_{n-1}^{\mathrm{GW}-\eta}(\lambda) = \eta(p) \times \frac{p!}{\prod_{i=1}^{p} m_i(\lambda)!} \times \frac{\prod_{i=1}^{p} \mathbb{P}(\#_{\mathrm{vertices}}T^\eta = \lambda_i)}{\mathbb{P}(\#_{\mathrm{vertices}}T^\eta = n)}$$

for all $\lambda \in \mathcal{P}_{n-1}$, where $\#_{\mathrm{vertices}}T^\eta$ is the number of leaves of the unconditioned GW tree T^η.

4. **Dynamical models of tree growth.** Rémy's, Ford's, Marchal's and the k-ary algorithms all lead to MB-sequences of trees indexed by leaves. Roughly, this can be seen by induction on n. By construction, the distribution of the leaves in the subtrees above the root is closely connected to urns models. We have the following expressions for the splitting distributions:

Ford's model: for $k \ge \frac{n}{2}$,

$$q_n^{\mathrm{Ford}-\alpha}(k, n-k) = \left(1 + \mathbb{1}_{k\neq \frac{n}{2}}\right) \frac{\Gamma(k-\alpha)\Gamma(n-k-\alpha)}{\Gamma(n-\alpha)\Gamma(1-\alpha)} \left(\frac{\alpha}{2}\binom{n}{k} + (1-2\alpha)\binom{n-2}{k-1}\right),$$

See [27] for a proof of the MB-property and calculation details. In particular, taking $\alpha = 1/2$ we see that

$$q_n^{\mathrm{Rémy}}(k, n-k) = \frac{\alpha}{2}\left(1 + \mathbb{1}_{k\neq \frac{n}{2}}\right) \frac{\Gamma(k-\alpha)\Gamma(n-k-\alpha)}{\Gamma(n-\alpha)\Gamma(1-\alpha)}\binom{n}{k}, \quad k \ge \frac{n}{2}.$$

k-ary model: for $\lambda = (\lambda_1, \ldots, \lambda_p) \in \mathcal{P}_n$

$$q_n^k(\lambda) = \sum_{\mathbf{n}=(n_1,\ldots,n_p)\in\mathbb{N}^p : \mathbf{n}^\downarrow = \lambda} \bar{q}_n(\mathbf{n})$$

12

where $\mathbf{n}^{\downarrow}$ is the decreasing rearrangement of the elements of $\mathbf{n}$ and

$$\bar{q}_n(\mathbf{n}) = \frac{1}{k(\Gamma(\frac{1}{k}))^{k-1}} \left(\prod_{i=1}^{k} \frac{\Gamma(\frac{1}{k} + n_i)}{n_i!} \right) \frac{n!}{\Gamma(\frac{1}{k} + n + 1)} \left(\sum_{j=1}^{n_1+1} \frac{n_1!}{(n_1 - j + 1)!} \frac{(n - j + 1)!}{n!} \right).$$

See [34, Proposition 3.3] for a proof of the MB-property and calculation details.

Marchal's model: for $\lambda = (\lambda_1, \ldots, \lambda_p) \in \mathcal{P}_n$,

$$q_n^{\text{Marchal}-\beta}(\lambda) = \frac{n!}{\lambda_1! \ldots \lambda_p! m_1(\lambda)! \ldots m_n(\lambda)!} \frac{\beta^{2-p} \Gamma(2 - \beta^{-1}) \Gamma(p - \beta)}{\Gamma(n - \beta^{-1}) \Gamma(2 - \beta)} \prod_{i=1}^{p} \frac{\Gamma(\lambda_j - \beta^{-1})}{\Gamma(1 - \beta^{-1})}$$

where $m_i(\lambda) = \#\{1 \le j \le p : \lambda_j = i\}$. See [21, Theorem 3.2.1] and [42, Lemma 5] for a proof of the MB-property (here the trees are GW-trees conditioned to have a fixed number of leaves) and calculation details.

5. **Cut-tree of a uniform Cayley tree**. Consider C_n a uniform Cayley tree of size n (i.e. a uniform rooted labelled tree with n vertices). It has the following recursive property (see Pitman [46, Theorem 5]): removing an edge uniformly at random in C_n gives two trees, which given their numbers of vertices, $k, n - k$ say, are independent uniform Cayley trees of respective sizes $k, n - k$. Bertoin [9] used this fact to study the _cut-tree_ T_n of C_n. The tree T_n is the genealogical tree of the following deletion procedure: remove in C_n one edge uniformly at random, then remove another edge in the remaining set of $n - 2$ edges uniformly at random and so on until all edges have been removed. We refer to [9] for a precise construction of T_n. Let us just mention here that T_n is a rooted binary tree with n leaves, and that Pitman's recursive property implies that $(T_n, n \ge 1)$ is MB. The corresponding splitting probabilities are:

$$q_n^{\text{Cut}-\text{tree}}(k, n - k) = \frac{(n - k)^{n-k-1}}{(n - k)!} \frac{k^{k-1}}{k!} \frac{(n - 2)!}{n^{n-3}}, \quad k > n/2,$$

the calculations are detailed in [9, 45].

Remark. The first example is a simple example of models where _macroscopic branchings are frequent_, unlike the second example where _macroscopic branchings are rare_ (they occur with probability $n^{-\gamma} \to 0$). Although its not completely obvious yet, all other examples above have rare macroscopic branchings (in a sense that will be specified later) and this is typically the context in which we will study the scaling limits of MB-trees. Typically the tree T_n will then grow in a power of n. When macroscopic branchings are frequent, there is no scaling limit in general for the Gromov-Hausdorff topology, a topology introduced in the next section. However it is known that the height of the tree T_n is then often of order $c \ln(n)$. This has been studied in [14].

3 The Galton–Watson example and topological framework

We start with an informal version of the prototype result of Aldous on the description of the asymptotics of conditioned Galton-Watson trees. Let η be a critical offspring distribution with *finite variance* $\sigma^2 \in (0, \infty)$, and let T_n^η denote a Galton-Watson tree with offspring distribution η, conditioned to have n vertices (in the following it is implicit that we only consider integers n such that this conditioning is possible). Then Aldous [5] showed that

$$\frac{\sigma}{2} \times \frac{T_n^\eta}{n^{1/2}} \xrightarrow[n \to \infty]{(\mathrm{d})} \mathcal{T}_{\mathrm{Br}} \tag{3}$$

where the continuous tree $\mathcal{T}_{\mathrm{Br}}$ arising in the limit is called the Continuum Random Tree (CRT for short), or sometimes, more precisely the Brownian CRT, or Brownian tree. Note that the limit only depends on η via its variance σ.

This result by Aldous was a breakthrough in the study of large random trees, since it was the first to describe the behavior of the tree as a whole. We will discuss this in more details in Section 3.2. Let us first formalize the topological framework in order to make sense of this convergence.

3.1 Real Trees and the Gromov–Hausdorff topology

Since the pioneering works of Evans, Pitman and Winter [25] in 2003 and Duquesne and Le Gall [22] in 2005, the theory of *real trees* (or $\mathbb{R}$-trees) has been intensively used in probability. These trees are metric spaces having a "tree property" (roughly, this means that for each pair of points x, y in the metric space, there is a unique path going from x to y – see below for a precise definition). This point of view allows behavior such as infinite total length of the tree, vertices with infinite degree, and density of the sets of leaves in the tree.

In these lectures, all the real trees we will deal with are compact metric spaces. For this reason, we restrict ourselves to the theory of compact real trees. We now briefly recall background on real trees and the Gromov-Hausdorff and Gromov-Hausdorff-Prokhorov distances, and refer to [26, 36] for an overview on this topic.

Real trees. A *real tree* is a metric space $(\mathcal{T}, d)$ such that, for any points x and y in $\mathcal{T}$,

- there is an isometry $\varphi_{x,y} : [0, d(x, y)] \to \mathcal{T}$ such that $\varphi_{x,y}(0) = x$ and $\varphi_{x,y}(d(x, y)) = y$
- for every continuous, injective function $c : [0, 1] \to \mathcal{T}$ with $c(0) = x$, $c(1) = y$, one has $c([0, 1]) = \varphi_{x,y}([0, d(x, y)])$.

Note that a discrete tree may be seen as a real tree by "replacing" its edges by line segments. Unless specified, it will be implicit in all these notes that these line segments are all of length 1.

We denote by $[[x, y]]$ the line segment $\varphi_{x,y}([0, d(x, y)])$ between x and y, and also write $]]x, y]]$ or $[[x, y[[$ when we want to exclude x or y. Our trees will always be *rooted* at a point $\rho \in \mathcal{T}$. The height of a point $x \in \mathcal{T}$ is defined by

$$\mathsf{ht}(x) = d(x, \rho)$$

and the height of the tree itself is the supremum of the heights of its points. The *degree* of x is the number of connected components of $\mathcal{T} \backslash \{x\}$. We call *leaves* of $\mathcal{T}$ all the points which have degree 1, excluding the root. A k-ary tree is a tree whose points have degrees in $\{1, 2, k + 1\}$ (with at least one point of degree $k + 1$). Given two points x and y, we define $x \wedge y$ as the unique point of $\mathcal{T}$ such that $[[\rho, x]] \cap [[\rho, y]] = [[\rho, x \wedge y]]$. It is called the *branch point* of x and y if its degree is larger or equal to 3. For $a > 0$, we define the rescaled tree $a\mathcal{T}$ as $(\mathcal{T}, ad)$ (the metric d thus being implicit and dropped from the notation).

As mentioned above, we will only consider compact real trees in this work. We now want to measure how close two such metric spaces are. We start by recalling the definition of Hausdorff distance between compact subsets of a metric space.

Hausdorff distance. If A and B are two nonempty compact subsets of a metric space (E, d), the Hausdorff distance between A and B is defined by

$$d_{E,\mathrm{H}}(A, B) = \inf \{\varepsilon > 0 \,;\, A \subset B^{\varepsilon} \text{ and } B \subset A^{\varepsilon}\},$$

where A^{ε} and B^{ε} are the closed ε-enlargements of A and B, i.e. $A^{\varepsilon} = \{x \in E : d(x, A) \leq \varepsilon\}$ and similarly for B^{ε}.

The Gromov-Hausdorff convergence generalizes this and allows us to talk about convergence of compact $\mathbb{R}$-trees.

Gromov–Hausdorff distance. Given two compact rooted trees $(\mathcal{T}, d, \rho)$ and $(\mathcal{T}', d', \rho')$, let

$$d_{\mathrm{GH}}(\mathcal{T}, \mathcal{T}') = \inf \left[\max(d_{\mathcal{Z},\mathrm{H}}(\phi(\mathcal{T}), \phi'(\mathcal{T}')), d_{\mathcal{Z}}(\phi(\rho), \phi'(\rho'))) \right],$$

where the infimum is taken over all pairs of isometric embeddings ϕ and ϕ' of $\mathcal{T}$ and $\mathcal{T}'$ in the same metric space $(\mathcal{Z}, d_{\mathcal{Z}})$, for all choices of metric spaces $(\mathcal{Z}, d_{\mathcal{Z}})$.

We will also be concerned with *measured* trees, that is $\mathbb{R}$-trees equipped with a probability measure on their Borel sigma-field. To this effect, recall first the definition of the Prokhorov distance between two probability measures μ and μ' on a metric space (E, d):

$$d_{E,\mathrm{P}}(\mu, \mu') = \inf \{\varepsilon > 0 \,;\, \forall A \in \mathcal{B}(E), \mu(A) \leq \mu'(A^{\varepsilon}) + \varepsilon \text{ and } \mu'(A) \leq \mu(A^{\varepsilon}) + \varepsilon\}.$$

This distance metrizes the weak convergence on the set of probability measures on (E, d).

Gromov–Hausdorff–Prohorov distance. Now, given two measured compact rooted trees $(\mathcal{T}, d, \rho, \mu)$ and $(\mathcal{T}', d', \rho', \mu')$, we let

$$d_{\mathrm{GHP}}(\mathcal{T}, \mathcal{T}') = \inf \left[\max(d_{\mathcal{Z},\mathrm{H}}(\phi(\mathcal{T}), \phi'(\mathcal{T}')), d_{\mathcal{Z}}(\phi(\rho), \phi'(\rho')), d_{\mathcal{Z},\mathrm{P}}(\phi_* \mu, \phi'_* \mu')) \right],$$

where the infimum is taken on the same space as before and $\phi_*\mu$, $\phi'_*\mu'$ are the push-forwards of μ, μ' by ϕ, ϕ'.

As shown in [25] and [2], the space of compact rooted $\mathbb{R}$-trees (respectively compact measured rooted $\mathbb{R}$-trees), taken up to root-preserving isomorphisms (resp. root-preserving and measure-preserving) and equipped with the GH (resp. GHP) metric is Polish. We will always identify two (measured) rooted $\mathbb{R}$-trees when their are isometric and still use the notation $(\mathcal{T}, d)$ (or $\mathcal{T}$ when the metric is clear) to design their isometry class.

3.2   Scaling limits of conditioned Galton–Watson trees

Wa can now turn to rigorous statements on the asymptotic behavior of conditioned Galton–Watson trees. We reformulate the above result (3) by Aldous in the finite variance case and then present a result by Duquesne [19] when the offspring distribution η is heavy tailed, in the domain of attraction of a stable distribution.

Let T_n^η be a η-GW tree conditoned to have n vertices, and μ_n^η be the uniform probability on its vertices. The following convergences hold for the Gromov–Hausdorff–Prohorov topology.

Theorem 3.1.   (i) (Aldous [5]) *Assume that η has a finite variance σ^2. Then, there exists a random compact real tree, called the Brownian tree and denoted $\mathcal{T}_{\mathrm{Br}}$, endowed with a probability measure μ_{Br} supported by its set of leaves, such that*

$$\left(\frac{\sigma T_n^\eta}{2n^{1/2}}, \mu_n^\eta \right) \xrightarrow[n\to\infty]{\mathrm{(d)}} \left(\mathcal{T}_{\mathrm{Br}}, \mu_{\mathrm{Br}} \right).$$

(ii) (Duquesne [19]) *If $\eta_k \sim Ck^{-1-\alpha}$ as $k \to \infty$ for $\alpha \in (1,2)$, then there exists a random compact real tree $\mathcal{T}_\alpha$ called the* stable Lévy tree with index α, *endowed with a probability measure μ_α supported by its set of leaves, such that*

$$\left(\frac{T_n^\eta}{n^{1-1/\alpha}}, \mu_n^\eta \right) \xrightarrow{\mathrm{(d)}} \left(\left(\frac{\alpha-1}{C\Gamma(2-\alpha)} \right)^{1/\alpha} \alpha^{1/\alpha-1} \cdot \mathcal{T}_\alpha, \mu_\alpha \right).$$

The result of Duquesne actually extends to cases where the offspring distribution η is in the domain of attraction of a stable distribution with index $\alpha \in (1,2]$. See [19].

The family of stable Lévy trees $(\mathcal{T}_\alpha, \alpha \in (1,2]$] – by convention $\mathcal{T}_2$ is the brownian tree $\mathcal{T}_{\mathrm{Br}}$ – was introduced by Duquesne and Le Gall [21, 22], building on earlier work of Le Gall and Le Jan [38]. These trees are intimately related to continuous state branching processes, fragmentation and coalescence processes, and appear as scaling limits of various models of trees and graphs. In the last few years, the geometric and fractal aspects of stable trees have been studied in great detail: Hausdorff and packing dimensions and measures [22, 23, 20, 28]; spectral dimension [16]; spinal decompositions and invariance under uniform re-rooting [32, 24]; fragmentation into

subtrees [42, 43]; and embeddings of stable trees into each other [17]. The stable trees are also related to Beta-coalescents [1, 7]; intervene in the description of other scaling limits of random maps [37, 44, 39]; and have dual graphs, called the stable looptrees [18], which also appear as scaling limits of natural combinatorial models.

Applications to combinatorial trees: one can then use the connections between some families of combinatorial trees and Galton–Watson trees mentioned in Section 2.1 to obtain that

(1) If T_n is uniform amongst the set of rooted ordered trees with n vertices,

$$n^{-1/2}T_n \xrightarrow{\text{(d)}} \mathcal{T}_{B_r}.$$

(2) If T_n is uniform amongst the set of rooted trees with n labelled vertices,

$$n^{-1/2}T_n \xrightarrow{\text{(d)}} 2\mathcal{T}_{B_r}.$$

(3) If T_n is uniform amongst the set of rooted trees with n labelled vertices,

$$n^{-1/2}T_n \xrightarrow{\text{(d)}} 2\mathcal{T}_{B_r}.$$

This global perspective provides the behavior of several statistics of the trees (maximal height, height of a typical vertex, diameter, etc.) that first interested combinatorists.

We will not present the proofs of Aldous [5] and Duquesne [19] of these results, but will rather focus on the fact that they may be recovered by using the Markov branching property. This is the goal of the next two sections, where we will present in a general setting some results on the scaling limits for MB-sequences of trees. The main idea of the proofs of Aldous [5] and Duquesne [19] is rather based on the study of the so-called *contour function* of the trees. We refer to their papers, as well as Le Gall's survey for details. See also Duquesne and Le Gall [21] and Kortchemski [35] for further related results.

In another direction, let us mention that in some particular cases, it is possible to construct the sequence on conditioned Galton-Watson trees on a same probability space, and to improve the convergence in distribution into an almost sure convergence. This will be discussed in Section 5. In that section, we will also present some results on sequences of Galton-Watson trees conditioned by their number of leaves or more general arbitrary degree restrictions.

4   Scaling limits of Markov-Branching trees

Our goal is to set up a criterion on the splitting probabilities (q_n) of a MB-sequence such that this sequence, suitably normalized, converges to a continuous limit. We follow here the

presentation of the work [30]. We also refer to [31] were similar results where proved under stronger assumptions.

The splitting probability q_n corresponds to a "discrete" fragmentation of the integer n into smaller integers, $\forall n \geq 1$. To set up the desired criterion, we first need to introduce a continuous counterpart for these partitions of integers, namely

$$\mathcal{S}^{\downarrow} = \left\{ \mathbf{s} = (s_1, s_2, \ldots) : s_1 \geq s_2 \geq \ldots \geq 0 \text{ and } \sum_{i \geq 1} s_i = 1 \right\}$$

which is endowed with the distance $d_{\mathcal{S}^{\downarrow}}(\mathbf{s}, \mathbf{s}') = \sup_{i \geq 1} |s_i - s_i'|$. Our main hypothesis on (q_n) then reads:

Hypothesis (H): $\exists \, \gamma > 0$ and ν a non-trivial $\sigma-$finite measure on $\mathcal{S}^{\downarrow}$ satisfying $\int_{\mathcal{S}^{\downarrow}} (1 - s_1)\nu(\mathrm{d}\mathbf{s}) < \infty$ and $\nu(1, 0, \ldots) = 0$, such that

$$n^{\gamma} \sum_{\lambda \in \mathcal{P}_n} q_n(\lambda) \left(1 - \frac{\lambda_1}{n}\right) f\left(\frac{\lambda_1}{n}, \ldots, \frac{\lambda_p}{n}, 0, \ldots\right) \xrightarrow[n \to \infty]{} \int_{\mathcal{S}^{\downarrow}} (1 - s_1) f(\mathbf{s}) \nu(\mathrm{d}\mathbf{s}).$$

for all continuous $f : \mathcal{S}^{\downarrow} \to \mathbb{R}$.

We will see that most of the examples of splitting probabilities introduced in Section 2.3 satisfies this hypothesis. As a first, easy, example, we consider the following case: $q_n((n)) = 1 - cn^{-\alpha}$ and $q_n(\lceil n/2 \rceil, \lfloor n/2 \rfloor) = cn^{-\alpha}$, $\alpha > 0$. Then, clearly, (H) is satisfied with

$$\gamma = \alpha \quad \text{and} \quad \nu(\mathrm{d}\mathbf{s}) = c\delta_{\left(\frac{1}{2}, \frac{1}{2}, 0, \ldots\right)}.$$

Informally, the interpretation of the hypothesis (H) is that macroscopic branchings are rare:

with proba.~ 1 :   with proba. $\sim \frac{\nu(\mathrm{d}\mathbf{s})}{n^{\gamma}}$:

This, of course, is a very rough traduction of (H), since the measure ν may be infinite. In such a case, to be a little more precise, the splitting events $n \mapsto n\mathbf{s}, \mathbf{s} \in \mathcal{S}^{\downarrow}$ with $s_1 < 1 - \varepsilon$ for some $\varepsilon \in (0, 1)$ occur asymptotically with a probability proportional to $n^{-\gamma} \mathbb{1}_{\{s_1 < 1 - \varepsilon\}} \nu(\mathrm{d}\mathbf{s})$.

The main result of this section is the following.

Theorem 4.1. *Let* $(T_n, n \geq 1)$ *be a* MB-*sequence indexed by the leaves and assume that its splitting probabilities satisfy* (H). *Then there exists a compact, measured real tree* $(\mathcal{T}_{\gamma, \nu}, \mu_{\gamma, \nu})$ *such that*

$$\left(\frac{T_n}{n^{\gamma}}, \mu_n\right) \xrightarrow[\mathrm{GHP}]{\mathrm{(d)}} (\mathcal{T}_{\gamma, \nu}, \mu_{\gamma, \nu}),$$

where μ_n *is the uniform probability on the leaves of* T_n.

18

The goal of this section is to detail the main steps of the proof of this result and to discuss some properties of the limiting tree, which belongs to the so-called family of *self-similar fragmentation trees*. In that aim we will first see how grows the height of a leaf chosen uniformly at random in T_n (Section 4.1 and Section 4.2). Then we will review some results on self-similar fragmentation trees (Section 4.3). And last we will show, by induction, that we can use the one-dimensional behavior (behavior of the height of a random leaf), to get the k-dimensional convergence (behavior of the subtree spanned by k leaves chosen independently), and conclude with a tightness criterion (Section 4.4).

There is a similar result for MB-sequences indexed by the vertices.

Theorem 4.2. *Let $(T_n, n \geq 1)$ be a* MB-*sequence indexed by the vertices and assume that its splitting probabilities satisfy* (H) *for some* $0 < \gamma < 1$. *Then there exists a compact, measured real tree* $(\mathcal{T}_{\gamma,\nu}, \mu_{\gamma,\nu})$ *such that*

$$\left(\frac{T_n}{n^{\gamma}}, \mu_n \right) \xrightarrow[\text{GHP}]{\text{(d)}} (\mathcal{T}_{\gamma,\nu}, \mu_{\gamma,\nu}),$$

where μ_n *is the uniform probability on the vertices of* T_n.

Theorem 4.2 is actually a direct corollary of Theorem 4.1, for the following reason. Consider a MB-sequence indexed by the vertices with splitting probabilities (p_n) and for all n, branch on each internal vertex of the tree T_n an edge-leaf. This gives a tree $\overline{T}_n$ with n vertices. It is then obvious that $(\overline{T}_n, n \geq 1)$ is MB-sequence indexed by the *leaves*, with splitting probabilities (p_n) defined by

$$q_n(\lambda_1, \ldots, \lambda_p, 1) = p_{n-1}(\lambda_1, \ldots, \lambda_p), \quad \text{for all } (\lambda_1, \ldots, \lambda_p) \in \mathcal{P}_{n-1}$$

(and $q_n(\lambda) = 0$ for all other $\lambda \in \mathcal{P}_n$). It is also clear that (q_n) satisfies (H) with parameters (γ, ν), $0 < \gamma < 1$, if and only if (p_n) does. Hence Theorem 4.1 implies Theorem 4.2.

Remark. These two statements are also valid when replacing in (H) and in the theorems the power sequence n^{γ} by any regularly varying sequence with index $\gamma > 0$. We recall that a sequence (a_n) varies regularly with index $\gamma > 0$ if for all $c > 0$,

$$\frac{a_{\lfloor cn \rfloor}}{a_n} \xrightarrow[n \to \infty]{} c^{\gamma}.$$

We refer to [13] for backgrounds on that topic. For simplicity, in the following we will only works with power sequences, but the reader should have in mind that everything holds similarly for regularly varying sequences.

4.1 A Markov chain in the MB–sequence of trees

Consider $(T_n, n \geq 1)$ a MB–sequence of trees indexed by the leaves and given T_n, mark one of its n leaves uniformly at random. We denote by $\star_n$ this marked leaf, and then by $\star_n(k)$ its
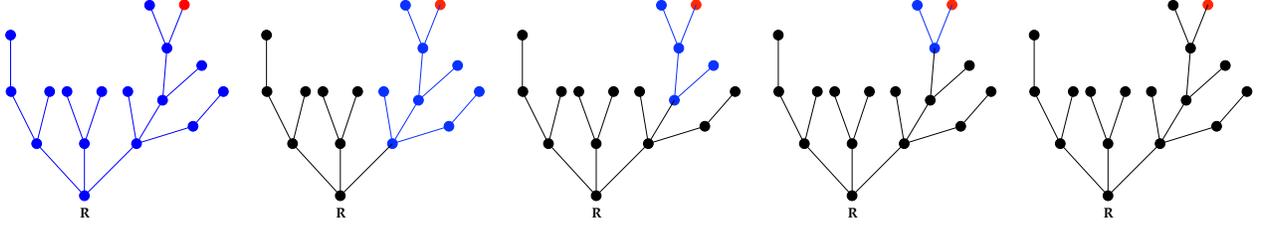
Figure 2: A Markov chain in the Markov branching tree: here, $n = 9$ and $X_9(0) = 9$, $X_9(1) = 5$, $X_9(2) = 3$, $X_9(3) = 2$, $X_9(4) = 1$ and $X_9(i) = 0, \forall i \geq 5$.

ancestor at generation k, $0 \leq k \leq n$ (so that $\star_n(0)$ is the root of T_n and $\star_n(\mathrm{ht}(\star_n)) = \star_n$). We then let $T_n^\star(k)$ denote the subtree composed by the descendants of $\star_n(k)$ in T_n, formally,

$$T_n^\star(k) := \{v \in T_n : \star_n(k) \in [[\rho, v]]\}, \quad k \leq \mathrm{ht}(\star)$$

and $T_n^\star(k) := \emptyset$ if $k > \mathrm{ht}(\star)$. We then set

$$X_n(k) := \# \{\text{leaves of } T_n^\star(k)\}, \quad \forall k \in \mathbb{Z}_+ \tag{4}$$

with the convention that $X_n(k) = 0$ for $k > \mathrm{ht}(\star)$.

Proposition 4.3. $(X_n(k), k \geq 0)$ *is a $\mathbb{Z}_+$–valued non–increasing Markov chain starting from $X_n(0) = n$, with transition probabilities*

$$p(i, j) = \sum_{\lambda \in \mathcal{P}_i} p_i(\lambda) m_j(\lambda) \frac{j}{i} \quad \text{for all } 0 \leq j \leq i, \text{ with } i > 1 \tag{5}$$

and $p_{1,0} = p_1(\emptyset) = 1 - p_{1,1}$.

Proof. The Markov property is a direct consequence of the Markov branching property. Given $X_n(1) = i_1, \ldots, X_n(k-1) = i_{k-1}$, the tree $T_n^\star(k-1)$ is distributed as $T_{i_{k-1}}$ if $i_{k-1} \geq 1$ and is the emptyset otherwise. In particular, when $i_{k-1} = 0$, the conditional distribution of $X_n(k)$ is the Dirac mass at 0. When $i_{k-1} \geq 1$, we use that, by definition, $\star_n$ is in $T_n^\star(k-1)$, hence, by still conditioning on the same event, we have that $\star_n$ is uniformly distributed amongst the i_{k-1} leaves of $T_{i_{k-1}}$. Otherwise said, given $X_n(1) = i_1, \ldots, X_n(k-1) = i_{k-1}$ with $i_{k-1} \geq 1$, $(T_n^\star(k-1), \star_n)$ is distributed as $(T_{i_{k-1}}, \star_{i_{k-1}})$ and consequently $X_n(k)$ is distributed as $X_{i_{k-1}}(1)$. Hence the Markov property of the chain $(X_n(k), k \geq 0)$. It remains to compute the transition probabilities:

$$p(n, k) = \mathbb{P}(X_n(1) = k) = \sum_{\lambda \in \mathcal{P}_n} q_n(\lambda) \mathbb{P}(X_n(1) = k | \Lambda_n = \lambda)$$

where Λ_n denotes the partition of n corresponding to the distribution of the leaves in the subtrees of T_n above the root. Since $\star_n$ is chosen uniformly amongst the set of leaves, we clearly have that

$$\mathbb{P}(X_n(1) = k | \Lambda_n = \lambda) = \frac{k}{n} \times \#\{i : \lambda_i = k\}.$$

$\square$

20

Hence studying the asymptotic behavior of the height of the marked leaf in the tree T_n reduced to studying the asymptotic behavior of the absorption time A_n of the Markov chain $(X_n(k), k \geq 0)$ at 0:

$$A_n := \inf\{k \geq 0 : X_n(k) = 0\}$$

(to be precise, this absorption time is equal to the height of the marked leaf $+1$). The study of the scaling limits of $((X_n(k), k \geq 1), A_n)$ as $n \to \infty$ is the goal of the next section. Before undertaking this task, let us notice that the hypothesis (H) on the splitting probabilities $(q_n, n \geq 1)$ of $(T_n, n \geq 1)$, together with (5) implies the following fact on the transition probabilities $(p(n,k), k \leq n)$:

$$n^\gamma \sum_{k=0}^n p(n,k) \left(1 - \frac{k}{n}\right) g\left(\frac{k}{n}\right) \longrightarrow \int_{[0,1]} g(x)\mu(\mathrm{d}x) \tag{6}$$

for all continuous functions $g : [0,1] \to \mathbb{R}$, where the measure μ in the limit is a finite, non–zero measure on $[0,1]$ defined by

$$\int_{[0,1]} g(x)\mu(\mathrm{d}x) = \int_{\mathcal{S}^\downarrow} \sum_{i \geq 1} s_i(1 - s_i)g(s_i)\nu(\mathrm{d}\mathbf{s}). \tag{7}$$

To see this, apply (H) to the continuous function defined by

$$f(\mathbf{s}) = \frac{\sum_{i \geq 2} s_i(1 - s_i)g(s_i)}{1 - s_1} \quad \text{for } \mathbf{s} \neq (1, 0, \ldots)$$

and $f(1, 0, \ldots) = g(1) + g(0)$.

4.2   Scaling limits of non–increasing Markov chain

According to the previous section, studying the height of *typical leaf* of T_n (i.e. a leaf marked uniformly at random) amounts to studying the absorption time at 0 of a $\mathbb{Z}_+$-valued non-increasing Markov chain. In this section, we study in a general framework the scaling limit of $\mathbb{Z}_+$-valued non–increasing Markov chains, under appropriate assumptions on the transition probabilities. We will then see how this applied to the height of a typical leaf in a MB–sequence of trees.

From now on,

$$(X_n(k), k \geq 0)$$

is a non–increasing $\mathbb{Z}_+$-valued Markov chain starting from n ($X_n(0) = n$), with transition probabilities $(p(i,j), 0 \leq j \leq i)$ such that

Hypothesis (H$'$): $\exists\, \gamma > 0$ and μ a non–trivial finite measure on $[0,1]$ such that

$$n^\gamma \sum_{k=0}^n p(n,k) \left(1 - \frac{k}{n}\right) f\left(\frac{k}{n}\right) \longrightarrow \int_{[0,1]} f(x)\mu(\mathrm{d}x)$$

21

for all continuous test functions $f : [0,1] \to \mathbb{R}$.

This hypothesis means that starting from n, "macroscopic" (i.e. with size proportional to n) jumps are rare, in the sens that the relative mean of the first jump tends to 0 as follows:

$$\mathbb{E}\left[\frac{n - X_n(1)}{n}\right] \sim \frac{\mu([0,1])}{n^\gamma}$$

and for a.e. $0 < \varepsilon \leq 1$, the probability to do a jump larger than εn is of order $c_\varepsilon n^{-\gamma}$ where $c_\varepsilon = \int_{[0,1-\varepsilon]} (1-x)^{-1} \mu(\mathrm{d}x)$ (note that this may tend to ∞ when ε tends to 0).

In the following, $\mathbb{D}([0,\infty), [0,\infty))$ denotes the set of non–negative càdlàg processes, endowed with the Skorokhod topology. Moreover, we let

$$A_n := \inf\{k \geq 0 : X_n(i) = X_n(k), \quad \forall i \geq k\}$$

be the first time at which the chain enter an absorption state (note that $A_n < \infty$ a.s. since the chain is non-increasing and $\mathbb{Z}_+$-valued).

Theorem 4.4 ([29]). *Assume* (H$'$).

(i) *Then,*

$$\left(\frac{X_n(\lfloor n^\gamma t \rfloor)}{n}, t \geq 0\right) \xrightarrow[n\to\infty]{(d)} \left(\exp(-\xi_{\tau(t)}), t \geq 0\right),$$

where ξ is a subordinator, i.e. a non-decreasing Lévy process. Its distribution is characterized by its Laplace transform $\mathbb{E}[\exp(-\lambda \xi_t)] = \exp(-t\phi(\lambda))$, with

$$\phi(\lambda) = \mu(\{0\}) + \mu(\{1\})\lambda + \int_{(0,1)} (1 - x^\lambda)\frac{\mu(\mathrm{d}x)}{1-x}, \quad \lambda \geq 0.$$

The time–change ρ (acceleration of time) is defined by

$$\tau(t) = \inf\left\{u \geq 0 : \int_0^u \exp(-\gamma\xi_r)\mathrm{d}r \geq t\right\}, t \geq 0.$$

(ii) *Moreover, jointly with the above convergence,*

$$\frac{A_n}{n^\gamma} \xrightarrow[n\to\infty]{(d)} \int_0^\infty \exp(-\gamma\xi_r)\mathrm{d}r = \inf\left\{t \geq 0 : \exp(-\xi_{\tau(t)}) = 0\right\}.$$

Main steps of the proof. – to be detailed – Assume (H$'$).

(1) Let : $Y_n(t) := n^{-1}X_n(\lfloor n^\gamma t \rfloor)$, then $(Y_n, n \geq 1)$ is tight

   to see this use (H') and Aldous' tightness criterion.

22

(2) It remains to prove that every possible limits in law of the subsequences of (Y_n) are distributed as X_∞.

Let Y' be such a limit: $\exists$ a subsequence $(n_k, k \geq 1)$ s.t. $Y_{n_k} \xrightarrow{(d)} Y'$

Let $\tau_{Y_n}(t) := \inf\{u : \int_0^u Y_n^{-\gamma}(r)\mathrm{d}r > t\}$, $\tau_{Y'}(t) := \inf\{u : \int_0^u (Y'(r))^{-\gamma}\mathrm{d}r > t\}$

$$Z_n(t) := Y_n(\tau_{Y_n}(t)) \quad \text{and} \quad Z'(t) = Y'(\tau_{Y'}(t))$$

Fact: $Y'(t) = Z'(\tau_{Y'}^{-1}(t)) = Z'(\inf\{u : \int_0^u Z'^\gamma(r)\mathrm{d}r > t\})$.

(a) Observe the following (easy!) fact: if P is the transition function of a Markov chain M with countable state space $\subset \mathbb{R}$, then for any positive function f such that $f^{-1}(\{0\})$ is absorbing,

$$f(M(k)) \prod_{i=0}^{k-1} \frac{f(M(i))}{Pf(M(i))}, \quad k \geq 0$$

is a martingale.

As a consequence: for all $\lambda \geq 0$ and $n \geq 1$, let $G_n(\lambda) := \mathbb{E}\left[(X_n(1)/n)^\lambda\right]$, then,

$$M_n^{(\lambda)}(t) := Z_n^\lambda(t) \left(\prod_{i=0}^{\lfloor n^\gamma \tau_{Y_n}(t)\rfloor - 1} G_{X_n(i)}(\lambda) \right)^{-1}, \quad t \geq 0$$

is a martingale.

(b) Since we have assumed $Y'_{n_k} \to Y'$ and (H'),

$$M_{n_k}^{(\lambda)} \xrightarrow{(d)} (Z')^\lambda \exp(\phi(\lambda)\cdot),$$

and the limit is also a martingale.

(c) $\Rightarrow -\ln Z'$ is an increasing Lévy process, with Laplace exponent ϕ,

(easy to see with Laplace transforms).

Hence $Y \overset{(d)}{=} X_\infty$. $\qquad\square$

This in particular leads to the following expected corollary on the asymptotics of the height of a marked leaf in MB-sequence of trees.

Corollary 4.5. *Let $(T_n, n \geq 1)$ be a MB-sequence indexed by the leaves, with splitting probabilities satisfying* (H), *with parameters (γ, ν). For each n, let $\star_n$ be a leaf chosen uniformly amongst the n leaves of T_n. Then,*

$$\frac{\mathsf{ht}(\star_n)}{n^\gamma} \xrightarrow[n\to\infty]{(d)} \int_0^\infty \exp(-\gamma\xi_r)\mathrm{d}r$$

where ξ is a subordinator with Laplace transform $\phi(\lambda) = \int_{\mathcal{S}^\downarrow} \sum_{i\geq 1} \left(1 - s_i^\lambda\right) s_i \nu(\mathrm{d}\mathbf{s})$.

23

Proof. We have seen at the end of the previous section that under (H), the transition probabilities of the Markov chain (4) satisfy assumption (H′) with parameters γ and μ, with μ defined by (7). The conclusion follows with Theorem 4.4 (ii). $\qquad\square$

To learn more. Apart from applications to Markov branching trees, Theorem 4.4 can be used to describe the asymptotic behavior (see [29]) of random walks with a barrier or of the number of collisions in Λ-coalescent processes. Recently, Bertoin and Kortchemski [11] set up similar results to *non-monotone* Markov chains and develop several applications, to random walks conditioned to stay positive, to the number of particles in some coagulation-fragmentations processes, to random planar maps (see [10] for this last point). Also in [33] (in progress) similar convergences for typed Markov chains towards "Lamperti time changed" Markov additive processes are studied. This will have applications to dynamical models of tree growth in a broader context than the one presented here, and more generally to multi-types MB–trees.

4.3 Self-similar fragmentation trees.

Self-similar fragmentation trees are random compact, measured real trees that describe the genealogical structure of some self-similar fragmentation processes. This set of trees corresponds to the set of trees arising as scaling limits of MB–trees. We start by introducing the self-similar fragmentation processes, following Bertoin [8], and then turn to the description of their genealogical trees, which were first introduced in [28] and then in [49] in a broader context.

4.3.1 Self-similar fragmentation processes

4.3.2 Self-similar fragmentation trees

To be completed.

4.4 Scaling limits of MB trees.

We will now see how we can use the previous section to prove Theorem 4.1. We only give a hint of the proof of the convergence of rescaled trees and refer to [30, Section 4.4] to see how to incorporate the measures. The proof of the convergence of rescaled trees proceeds in two main steps:

Convergence of finite-dimensional marginals. For all integers $k \geq 2$, let $T_n(k)$ be the subtree of T_n spanned by the root and k (different) leaves picked independently, uniformly at random. Similarly, let $\mathcal{T}_{\gamma,\nu}(k)$ be the subtree of $\mathcal{T}_{\gamma,\nu}$ spanned by the root and k leaves picked

independently at random according to the measure $\mu_{\gamma,\nu}$. Then (under (H))

$$\frac{T_n(k)}{n^\gamma} \xrightarrow[n\to\infty]{(d)} \mathcal{T}_{\gamma,\nu}(k). \tag{8}$$

This is what we call *the finite-dimensional marginals convergence*. The proof holds by induction on k. For $k = 1$, this is Corollary 4.5. For $k \geq 2$, the proof relies on the induction hypothesis and on the MB-property. Here is the main idea. Consider the decomposition of T_n into subtrees above its first branch point in $T_n(k)$ and take only into account the subtrees having marked leaves. We obtain $m \geq 2$ subtrees with, say, $n_1, \ldots, n_m$ leaves respectively ($\sum_{i=1}^m n_i \leq n$), and each of these trees there have $k_1 \geq 1, \ldots k_m \geq 1$ marked leaves ($\sum_{i=1}^m k_i = k$). Given m, the sizes $n_1, \ldots, n_m$, and the number of marked leaves $k_1 \geq 1, \ldots k_m \geq 1$, the MB-property ensures that the m subtrees are independent with respective distributions that of $T_{n_1}(k_1), \ldots, T_{n_m}(k_m)$. An application of the induction hypothesis to these subtrees leads to the expected result. We refer to [30, Section 4.2] for details.

A tightness criterion. To get the convergence for the GH-topology, the previous result must be completed with a tightness criterion. The idea is to use the following well-known result.

Theorem 4.6. *If $X_n, X, X_n(k), X(k)$ are r.v. in a metric space (E, d) such that $X_n(k) \xrightarrow[n\to\infty]{(d)} X(k), \forall k$ and $X(k) \xrightarrow[k\to\infty]{(d)} X$ and for all $\varepsilon > 0$,*

$$\lim_{k\to\infty} \limsup_{n\to\infty} \mathbb{P}\left(d(X_n, X_n(k)) > \varepsilon\right) = 0 \tag{9}$$

then $X_n \xrightarrow[n\to\infty]{(d)} X$.

See [12, Theorem 3.2]. In our context, the finite-dimensional convergence (8) has already been checked, and we know from Section 4.3 that $\mathcal{T}_{\gamma,\nu}(k) \to \mathcal{T}_{\gamma,\nu}$ almost surely as $k \to \infty$. It remains to to establish the tightness criterion (9) for $T_n, T_n(k)$. The main tool is the following bounds:

Proposition 4.7. *Under (H), for all $p > 0$, there exists a finite constant C_p such that*

$$\mathbb{P}\left(\frac{\mathsf{ht}(T_n)}{n^\gamma} \geq x\right) \leq \frac{C_p}{x^p}, \quad \forall x > 0, \forall n \geq 1.$$

The proof holds by induction on n, using (H) and the MB-property. We refer to [30, Section 4.3] for details and to see how, using again the MB-property, this helps to control the distance between T_n and $T_n(k)$, and to get the that for $\varepsilon > 0$:

$$\lim_{k\to\infty} \limsup_{n\to\infty} \mathbb{P}\left(d_{\mathrm{GH}}\left(\frac{T_n(k)}{n^\gamma}, \frac{T_n}{n^\gamma}\right) \geq \varepsilon\right) = 0$$

as required.

25

5 Applications

To be completed.

5.1 Galton–Watson trees

5.2 Pólya trees

5.3 Dynamical models of tree growth

5.4 Cut-tree of Cayley trees

References

[1] R. ABRAHAM AND J.-F. DELMAS, *β-coalescents and stable Galton-Watson trees*, ALEA Lat. Am. J. Probab. Math. Stat., 12 (2015), pp. 451–476.

[2] R. ABRAHAM, J.-F. DELMAS, AND P. HOSCHEIT, *A note on the Gromov-Hausdorff-Prokhorov distance between (locally) compact metric measure spaces*, Electron. J. Probab., 18(14) (2013), pp. 1–21.

[3] D. ALDOUS, *The continuum random tree. I*, Ann. Probab., 19 (1991), pp. 1–28.

[4] ——, *The continuum random tree. II. An overview*, in Stochastic analysis (Durham, 1990), vol. 167 of London Math. Soc. Lecture Note Ser., Cambridge Univ. Press, Cambridge, 1991, pp. 23–70.

[5] D. ALDOUS, *The continuum random tree III*, Ann. Probab., 21 (1993), pp. 248–289.

[6] D. ALDOUS, *Probability distributions on cladograms*, in Random discrete structures (Minneapolis, MN, 1993), vol. 76 of IMA Vol. Math. Appl., Springer, New York, 1996, pp. 1–18.

[7] J. BERESTYCKI, N. BERESTYCKI, AND J. SCHWEINSBERG, *Beta-coalescents and continuous stable random trees*, Ann. Probab., 35 (2007), pp. 1835–1887.

[8] J. BERTOIN, *Random fragmentation and coagulation processes*, vol. 102 of Cambridge Studies in Advanced Mathematics, Cambridge University Press, Cambridge, 2006.

[9] ——, *Fires on trees*, Ann. Inst. Henri Poincaré Probab. Stat., 48 (2012), pp. 909–921.

[10] J. BERTOIN, N. CURIEN, AND I. KORTCHEMSKI, *Random planar maps & growth-fragmentations*, (2015). preprint, arXiv:1507.02265.

[11] J. BERTOIN AND I. KORTCHEMSKI, *Self-similar scaling limits of markov chains on the positive integers*. Prépublication – arXiv:1412.1068.

[12] P. BILLINGSLEY, *Convergence of probability measures*, Wiley Series in Probability and Statistics: Probability and Statistics, John Wiley & Sons, Inc., New York, second ed., 1999. A Wiley-Interscience Publication.

[13] N. H. Bingham, C. M. Goldie, and J. L. Teugels, *Regular variation*, vol. 27 of Encyclopedia of Mathematics and its Applications, Cambridge University Press, Cambridge, 1989.

[14] N. Broutin, L. Devroye, E. McLeish, and M. de la Salle, *The height of increasing trees*, Random Structures Algorithms, 32 (2008), pp. 494–518.

[15] B. Chen, D. Ford, and M. Winkel, *A new family of Markov branching trees: the alpha-gamma model*, Electron. J. Probab., 14 (2009), pp. no. 15, 400–430.

[16] D. Croydon and B. Hambly, *Spectral asymptotics for stable trees*, Electron. J. Probab., 15 (2010), pp. 1772–1801, paper no. 57 (electronic).

[17] N. Curien and B. Haas, *The stable trees are nested*, Probab. Theory Related Fields, 157(1) (2013), pp. 847–883.

[18] N. Curien and I. Kortchemski, *Random stable looptrees*, Electron. J. Probab., 19 (2014), pp. 1–35, paper no. 108 (electronic).

[19] T. Duquesne, *A limit theorem for the contour process of conditioned Galton-Watson trees*, Ann. Probab., 31 (2003), pp. 996–1027.

[20] ——, *The exact packing measure of Lévy trees*, Stochastic Process. Appl., 122 (2012), pp. 968–1002.

[21] T. Duquesne and J.-F. Le Gall, *Random trees, Lévy processes and spatial branching processes*, Astérisque, (2002), pp. vi+147.

[22] ——, *Probabilistic and fractal aspects of Lévy trees*, Probab. Theory Related Fields, 131 (2005), pp. 553–603.

[23] ——, *The Hausdorff measure of stable trees*, ALEA Lat. Am. J. Probab. Math. Stat., 1 (2006), pp. 393–415.

[24] ——, *On the re-rooting invariance property of Lévy trees*, Electron. Commun. Probab., 14 (2009), pp. 317–326.

[25] S. Evans, J. Pitman, and A. Winter, *Rayleigh processes, real trees, and root growth with re-grafting.*, Probab. Theory Related Fields, 134(1) (2006), pp. 918–961.

[26] S. N. Evans, *Probability and real trees*, vol. 1920 of Lecture Notes in Mathematics, Springer, Berlin, 2008. Lectures from the 35th Summer School on Probability Theory held in Saint-Flour, July 6–23, 2005.

[27] D. Ford, *Probabilities on cladograms: introduction to the alpha model.* Prépublication – arXiv:math/0511246.

[28] B. Haas and G. Miermont, *The genealogy of self-similar fragmentations with negative index as a continuum random tree*, Electron. J. Probab., 9 (2004), pp. no. 4, 57–97 (electronic).

[29] B. Haas and G. Miermont, *Self-similar scaling limits of non-increasing Markov chains*, Bernoulli Journal, 17 (2011), pp. 1217–1247.

27

[30] ——, *Scaling limits of Markov branching trees with applications to Galton-Watson and random unordered trees*, Ann. Probab., 40 (2012), pp. 2589–2666.

[31] B. Haas, G. Miermont, J. Pitman, and M. Winkel, *Continuum tree asymptotics of discrete fragmentations and applications to phylogenetic models*, Ann. Probab., 36 (2008), pp. 1790–1837.

[32] B. Haas, J. Pitman, and M. Winkel, *Spinal partitions and invariance under re-rooting of continuum random trees*, Ann. Probab., 37 (2009), pp. 1381–1411.

[33] B. Haas and R. Stephenson, *Multitype Markov Branching trees*. En préparation.

[34] ——, *Scaling limits of k-ary growing trees*, Ann. IHP Proba. Stat., 51(4) (2015), pp. 1314–1341.

[35] I. Kortchemski, *Invariance principles for Galton-Watson trees conditioned on the number of leaves*, Stochastic Process. Appl., 122 (2012), pp. 3126–3172.

[36] J.-F. Le Gall, *Random real trees*, Ann. Fac. Sci. Toulouse Math. (6), 15 (2006), pp. 35–62.

[37] ——, *Uniqueness and universality of the Brownian map*, Ann. Probab., 41 (2013), pp. 2880–2960.

[38] J.-F. Le Gall and Y. Le Jan, *Branching processes in Lévy processes: the exploration process*, Ann. Probab., 26 (1998), pp. 213–252.

[39] J.-F. Le Gall and G. Miermont, *Scaling limits of random planar maps with large faces*, Ann. Probab., 39 (2011), pp. 1–69.

[40] P. Marchal, *Constructing a sequence of random walks strongly converging to Brownian motion*, in Discrete random walks (Paris, 2003), Discrete Math. Theor. Comput. Sci. Proc., AC, Assoc. Discrete Math. Theor. Comput. Sci., Nancy, 2003, pp. 181–190 (electronic).

[41] ——, *A note on the fragmentation of a stable tree*, in Fifth Colloquium on Mathematics and Computer Science, Discrete Math. Theor. Comput. Sci. Proc., AI, Assoc. Discrete Math. Theor. Comput. Sci., Nancy, 2008, pp. 489–499.

[42] G. Miermont, *Self-similar fragmentations derived from the stable tree. I. Splitting at heights*, Probab. Theory Related Fields, 127 (2003), pp. 423–454.

[43] ——, *Self-similar fragmentations derived from the stable tree. II. Splitting at nodes*, Probab. Theory Related Fields, 131 (2005), pp. 341–375.

[44] ——, *The Brownian map is the scaling limit of uniform random plane quadrangulations*, Acta Math., 210 (2013), pp. 319–401.

[45] J. L. Pavlov, *The asymptotic distribution of the maximum size of trees in a random forest*, Teor. Verojatnost. i Primenen., 22 (1977), pp. 523–533.

[46] J. Pitman, *Coalescent random forests*, J. Combin. Theory Ser. A, 85 (1999), pp. 165–193.

[47] J.-L. Rémy, *Un procédé itératif de dénombrement d'arbres binaires et son application à leur génération aléatoire*, RAIRO Inform. Théor., 19 (1985), pp. 179–195.

[48] D. RIZZOLO, *Scaling limits of Markov branching trees and Galton-Watson trees conditioned on the number of vertices with out-degree in a given set*, Ann. Inst. H. Poincaré Probab. Statist., 51 (2015), pp. 512–532.

[49] R. STEPHENSON, *General fragmentation trees*, Electron. J. Probab., 18(101) (2013), pp. 1–45.