Electronic Journal of Probability

Electron. J. Probab. **22** (2017), no. 95, 1–53.
ISSN: 1083-6489 DOI: 10.1214/17-EJP96

Local limits of Markov branching trees and their volume growth

Camille Pagnard*

Abstract

We are interested in the local limits of families of random trees that satisfy the Markov branching property, which is fulfilled by a wide range of models. Loosely, this property entails that given the sizes of the sub-trees above the root, these sub-trees are independent and their distributions only depend upon their respective sizes. The laws of the elements of a Markov branching family are characterised by a sequence of probability distributions on the sets of integer partitions which describes how the sizes of the sub-trees above the root are distributed.

We prove that under some natural assumption on this sequence of probabilities, when their sizes go to infinity, the trees converge in distribution to an infinite tree which also satisfies the Markov branching property. Furthermore, when this infinite tree has a single path from the root to infinity, we give conditions to ensure its convergence in distribution under appropriate rescaling of its distance and counting measure to a self-similar fragmentation tree with immigration. In particular, this allows us to determine how, in this infinite tree, the "volume" of the ball of radius R centred at the root asymptotically grows with R.

Our unified approach will allow us to develop various new applications, in particular to different models of growing trees and cut-trees, and to recover known results. An illustrative example lies in the study of Galton-Watson trees: the distribution of a critical Galton-Watson tree conditioned on its size converges to that of Kesten's tree when the size grows to infinity. If furthermore, the offspring distribution has finite variance, under adequate rescaling, Kesten's tree converges to Aldous' self-similar CRT and the total size of the R first generations asymptotically behaves like R^2.

Keywords: random trees; Markov branching trees; local limits; scaling limits; self-similar fragmentation trees; GHP topology; Galton-Watson trees.
AMS MSC 2010: 60J80; 60F17.
Submitted to EJP on August 26, 2016, final version accepted on August 19, 2017.
Supersedes arXiv:1608.06968v2.
Supersedes HAL:hal-01356126.

*Ceremade, UMR CNRS 7534 Université Paris-Dauphine, PSL Research university.
E-mail: pagnard@ceremade.dauphine.fr

1 Introduction

The focus of this work is to study the asymptotic behaviour of sequences of random trees which satisfy the Markov branching property first introduced by Aldous in [6, Section 4] and later extended for example in [17, 30, 31]. See Haas [28] for an overview of this general model and Lambert [40] for applications to models used in evolutionary biology. Our study will therefore encompass various models, like Galton-Watson trees conditioned on their total progeny or their number of leaves, certain models of cut-trees (see Bertoin [12, 13, 14]) or recursively built trees (see Rémy [45], Chen, Ford and Winkel [19], Haas and Stephenson [32]) as well as models of phylogenetic trees (Ford's α-model [24] and Aldous' β-splitting model [6]).

Informally, a sequence $(T_n)_n$ of random trees satisfies the Markov branching property if for all n, T_n has "size" n, and conditionally on the event "T_n has p sub-trees above its root with respective sizes $n_1 \geq \cdots \geq n_p$", these sub-trees are independent and for each $i = 1, \dots, p$, the i^{th} largest sub-tree is distributed like T_{n_i}. The sequence of distributions of $(T_n)_n$ is characterised by a family $q = (q_n)_n$ of probability distributions, referred to as "first-split distributions" (see next paragraph), where q_n is supported by the set of partitions of the integer n. We will detail two different constructions of Markov branching trees corresponding to a given sequence q for two different notions of size: the number of leaves or the number of vertices.

Let $(q_n)_n$ be a sequence of first-split distributions. A tree with n leaves with distribution in the associated Markov branching family is built with the following process. Consider a cluster of n identical particles and with probability $q_n(\lambda_1, \dots, \lambda_p)$, split it into p smaller clusters containing $\lambda_1, \dots, \lambda_p$ particles respectively. For each $i = 1, \dots, p$, independently of the other sub-clusters, split the i^{th} cluster according to q_{λ_i}. When a sub-cluster contains only 1 particle, with probability $q_1(1) < 1$, let it either give birth to a new sub-cluster which only contains 1 particle as well, or, with probability $1 - q_1(1)$, let the particle "die". Repeat this procedure until each of the particles are dead. The genealogy of these splits may be encoded as a tree with n leaves (which correspond to the death of each particle). We'll denote by $\mathrm{MB}_n^{\mathcal{L},q}$ the distribution of such a tree.

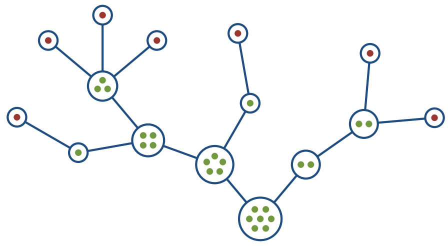

Figure 1: Example of a tree with 7 leaves (in red) and first-split equal to $(5, 2)$.

A Markov branching tree with a given number of vertices, say n, is built with a slightly different procedure and we will call MB_n^q its distribution. Section 2.2.1 will rigorously detail the constructions of both MB_n^q and $\mathrm{MB}_n^{\mathcal{L},q}$. Rizzolo [46] considered a more general notion of size and described the construction of corresponding Markov branching trees.

One way of looking at the behaviour of large trees is through the local limit topology. For a given tree $\mathtt{t}$ and $R \geq 0$, we denote by $\mathtt{t}|_R$ the subset of vertices of $\mathtt{t}$ at graph distance less than R from its root. We will say that a sequence $\mathtt{t}_n$ converges locally to a limit tree $\mathtt{t}_\infty$ if for any radius R, $\mathtt{t}_n|_R = \mathtt{t}_\infty|_R$ for sufficiently large n. There is considerable literature on the study of the local limits of certain classes of random trees or, more generally, of graphs. For instance, see Abraham and Delmas [1, 2], Stephenson [50], Stefánsson [47, 48] or a recent paper by Broutin and Mailler [18], as well as references therein, for studies related to our work.

Let us present in this Introduction the simplest, and most common, case in which Markov branching trees have local limits. Let $(T_n)_n$ be a sequence of Markov branching trees indexed by their size with corresponding family of first-split distributions $(q_n)_n$. Let p be a non-negative integer and $\lambda_1 \geq \cdots \geq \lambda_p > 0$ be a non increasing family of integers with sum L. For n large enough, consider $q_n(n-L, \lambda_1, \dots, \lambda_p)$, that is the probability that T_n gives birth to $p+1$ sub-trees among which the p smallest have respective sizes $\lambda_1, \dots, \lambda_p$. Assume that for any such p and λ, $q_n(n-L, \lambda_1, \dots, \lambda_p)$ converges to $q_*(\lambda_1, \dots, \lambda_p)$ for some probability measure q_* on the set of non-increasing finite sequences of positive integers. Under this natural assumption, we will prove in a rather straightforward way that T_n locally converges to some "infinite Markov branching tree" T_∞ with a single path from the root to infinity, called its *infinite spine*. The distribution of T_∞ is characterised by the family $(q_n)_n$ and the measure q_* which describes the distribution of the sizes of the finite sub-trees grafted on the spine of T_∞. See Theorem 2.9 for a more precise and general statement.

A drastically different approach to understand the behaviour of large random trees is that of scaling limits. Aldous was the first to study scaling limits of random trees as a whole, see [5], and notably introduced the celebrated Brownian tree as the limit of rescaled critical Galton-Watson trees conditioned on their size with any offspring law that has finite variance. See also Le Gall [41] for a survey on random "continuous" trees.

In this context, we will consider T_n as a metric space rescaled by some factor a_n, i.e. the edges of T_n will be viewed as real segments of length a_n, and denote by $a_n T_n$ this rescaled metric space. Scaling limits for Markov branching trees were studied in [30, 31] by Haas *et al.* Their main result (see Theorems 5 and 6 in [30]) is that under simple conditions on the sequence $(q_n)_n$ of first-split distributions, T_n converges in distribution, under appropriate rescaling, to a self-similar fragmentation tree. These objects were introduced by Haas and Miermont [29] and notably encompass Aldous' Brownian tree as well as Duquesne and Le Gall's stable trees [23].

Haas and Miermont's result from [30] in particular gives an asymptotic relation between the size and height of a finite Markov branching tree. When considering an infinite Markov branching tree T, we may wonder if a similar relation exists, namely how many vertices or leaves are typically found at height less than some large integer R. This seemingly simple question, the study of the integer sequence $(\#T|_R)_R$, leads us to consider the scaling limits of the weighted tree (T, μ_T), where μ_T is the counting measure on either the vertices of T or on its leaves.

In Theorem 4.2, we consider the case in which T is an infinite Markov branching tree with a unique infinite spine with distribution characterised by a family $(q_n)_n$ of first-split distributions and a probability measure q_* associated to the sizes of the finite sub-trees grafted on the spine. We prove that under the assumptions of Haas and Miermont's theorem on the family $(q_n)_n$ and an additional condition on the measure q_*, when R goes to infinity, the tree T/R endowed with the adequately rescaled measure μ_T converges in distribution to a self-similar fragmentation tree with immigration. These continuum random trees (CRTs) with infinite height were introduced by Haas [27]. They include Aldous' self-similar CRT [5] (which will appear as the limit in many of our applications) and Duquesne's Lévy trees with immigration [22].

As a result, under appropriate rescaling, the "volume" of the ball of radius R centred at the root of T converges in distribution to the measure of the ball with radius 1 centred at the root of a self-similar fragmentation tree with immigration. Proposition 4.3 actually gives the stronger convergence of the whole "volume growth" process.

The unified framework used here will yield multiple applications. As a first example, Theorem 2.9 will allow us to recover known results on the local limits of conditioned Galton-Watson trees towards Kesten's tree (see Abraham and Delmas [2] for instance) and Theorem 4.2 will give an alternative proof to Duquesne's results (see [22]) on the convergence of rescaled infinite critical Galton-Watson trees to Lévy trees with immigration. We will give similar results for some models of cut-trees, which encodes the genealogy of the random dismantling of trees, studied by Bertoin [12, 13, 14]. We will also study some models of sequentially growing trees described in [19, 32, 42, 45] and models of phylogenetic trees [6, 24].

This paper will be organised as follows. In Section 2, we will define finite and infinite Markov branching trees and give a natural criterion for their convergence under the local limit topology in Theorem 2.9. In Section 3 we will detail the background needed for our main result, Theorem 4.2, i.e. the study of the scaling limits of infinite Markov branching trees. Section 4 will focus on the proof of this result. Finally, Section 5 will give applications of our unified approach to various Markov branching models.

2 Markov branching trees and their local limits

2.1 Trees and partitions

2.1.1 Background on trees

Let $\mathcal{U} := \bigcup_{n\geq 0} \mathbb{N}^n$ be the set of finite words on $\mathbb{N}$ with the conventions $\mathbb{N} = \{1, 2, 3, \dots\}$ and $\mathbb{N}^0 = \{\varnothing\}$. We then call a *plane tree* or *ordered rooted tree* any non-empty subset $\mathtt{t} \subset \mathcal{U}$ such that:

- The empty word $\varnothing$ belongs to $\mathtt{t}$, it will be thought of as its "root",
- If $u = (u_1, \dots, u_n)$ is in $\mathtt{t}$, then its parent $\mathrm{pr}(u) := (u_1, \dots, u_{n-1})$ is also in $\mathtt{t}$,
- For all u in $\mathtt{t}$, there exists a finite integer $c_u(\mathtt{t}) \geq 0$ such that $u\,i := (u_1, \dots, u_n, i)$ is in $\mathtt{t}$ iff $1 \leq i \leq c_u(\mathtt{t})$. We will say that $c_u(\mathtt{t})$ is the number of children of u in $\mathtt{t}$.

Let $\mathtt{T}^{\mathrm{ord}}$ be the set of plane trees. Observe that if $\mathtt{t}$ is an infinite plane tree, this definition requires the number of children of each of its vertices to be finite.

Plane trees are endowed with a total order which is of limited interest to us. Because of this, we define an equivalence relation on $\mathtt{T}^{\mathrm{ord}}$ to allow us to consider as identical two trees which have the same "shape" but different vertex orderings.

Say that two plane trees $\mathtt{t}$ and $\mathtt{t}'$ are equivalent (written $\mathtt{t} \sim \mathtt{t}'$) iff there exists a bijection $\sigma : \mathtt{t} \to \mathtt{t}'$ such that $\sigma(\varnothing) = \varnothing$ and for all $u \in \mathtt{t} \setminus \{\varnothing\}$, $\mathrm{pr}[\sigma(u)] = \sigma[\mathrm{pr}(u)]$. Finally, set $\mathtt{T} := \mathtt{T}^{\mathrm{ord}}/\sim$. From now on, unless otherwise stated, we will only consider unordered trees, i.e. by "tree" we will mean an element of $\mathtt{T}$.

Let $\mathtt{t}$ be a tree. We say that a vertex u on $\mathtt{t}$ is a leaf if it has no children, i.e. if $c_u(\mathtt{t}) = 0$. Define $\#\mathtt{t}$ as the total number of vertices of $\mathtt{t}$ and $\#_{\mathcal{L}}\mathtt{t}$ as its number of leaves. For any positive integer n, let $\mathtt{T}_n$ and $\mathtt{T}_n^{\mathcal{L}}$ be the sets of finite trees with n vertices and n leaves respectively. Moreover, write $\mathtt{T}_\infty$ for the set of infinite trees.

We will use the following operations on trees:

- Let $\mathtt{t}_1, \dots, \mathtt{t}_d$ be trees; their *concatenation* is the tree $[\![\mathtt{t}_1, \dots, \mathtt{t}_d]\!]$ obtained by attaching each of their respective roots to a new common root, see Figure 2,
- Let $\mathtt{t}$ and $\mathtt{s}$ be two trees and u be a vertex of $\mathtt{t}$; set $\mathtt{t} \otimes (u, \mathtt{s})$ the grafting of $\mathtt{s}$ on $\mathtt{t}$ at u, i.e. the tree obtained by glueing the root of $\mathtt{s}$ on u, see Figure 3,
- Fix $\mathtt{t}$ a tree, a non-repeating family $(u_i)_{i\in\mathcal{I}}$ of vertices of $\mathtt{t}$, and a family of trees $(\mathtt{s}_i)_{i\in\mathcal{I}}$; let $\mathtt{t} \bigotimes_{i\in\mathcal{I}} (u_i, \mathtt{s}_i)$ be the tree obtained by grafting $\mathtt{s}_i$ on $\mathtt{t}$ at u_i for each i in $\mathcal{I}$.

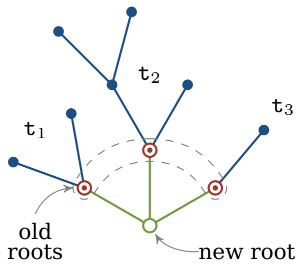

Figure 2: The tree $[\![\mathtt{t}_1, \mathtt{t}_2, \mathtt{t}_3]\!]$.

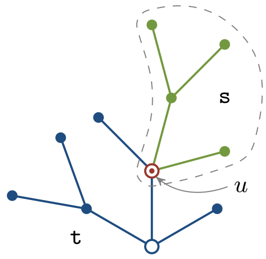

Figure 3: The tree $\mathtt{t} \otimes (u, \mathtt{s})$.

For all $n \geq 0$, let $\mathtt{b}_n$ be the *branch* of length n, i.e. the tree with $n+1$ vertices among which a single leaf. Similarly, define the infinite branch $\mathtt{b}_\infty$ and let $(\mathtt{v}_n)_{n\geq 0}$ be its vertices where $\mathtt{v}_0$ is its root and for all $n \geq 0$, $\mathtt{v}_n = \mathrm{pr}(\mathtt{v}_{n+1})$.

The local limit topology If $\mathtt{t}$ is a tree, we may endow it with the *graph distance* $\mathrm{d_{gr}}$ where for all u and v in $\mathtt{t}$, $\mathrm{d_{gr}}(u, v)$ is defined as the number of edges in the shortest path between u and v. For any non-negative integer R, we will write $\mathtt{t}|_R$ for the closed ball of radius R centred at the root of $\mathtt{t}$, that is the tree $\mathtt{t}|_R := \{u \in \mathtt{t} : \mathrm{d_{gr}}(\varnothing, u) \leq R\}$.

The local distance between two given trees $\mathtt{t}$ and $\mathtt{s}$ is defined as

$$\mathrm{d_{loc}}(\mathtt{t}, \mathtt{s}) := \exp\big[-\inf\{R \geq 0 : \mathtt{t}|_R \neq \mathtt{s}|_R\}\big].$$

The function $\mathrm{d_{loc}}$ is an ultra-metric on $\mathtt{T}$ and the resulting metric space $(\mathtt{T}, \mathrm{d_{loc}})$ is Polish. The following well-known criterion for convergence in distribution with respect to the local limit topology will be useful. See for instance [2, Section 2.2] for a proof (which relies on [16, Theorem 2.3] and the fact that $\mathrm{d_{loc}}$ is an ultra-metric).

Lemma 2.1. *Let T_n, $n \geq 1$ and T be $\mathtt{T}$-valued random variables. Then, $T_n \to T$ in distribution with respect to $\mathrm{d_{loc}}$ iff for all $\mathtt{t} \in \mathtt{T}$ and $R \geq 0$, $\mathbb{P}[T_n|_R = \mathtt{t}|_R] \to \mathbb{P}[T|_R = \mathtt{t}|_R]$ as n tends to infinity.*

2.1.2 Partitions of integers

As discussed in the Introduction, Markov branching trees are closely related to “partitions of integers”. This section thus aims to introduce a few notions on these objects which will be useful for our forthcoming purposes.

Set $\mathcal{P}_0 := \{\varnothing\}$, $\mathcal{P}_1 := \{\varnothing, (1)\}$ and for $n \geq 2$, let $\mathcal{P}_n$ be the set of *partitions* of n, i.e. of finite non-increasing integer sequences with sum n. More precisely, set

$$\mathcal{P}_n := \Big\{\lambda = (\lambda_1, \dots, \lambda_p) \in \mathbb{N}^p \,:\, p \geq 1,\ \lambda_1 \geq \dots \geq \lambda_p > 0 \text{ and } \lambda_1 + \dots + \lambda_p = n\Big\}.$$

Similarly, let $\mathcal{P}_\infty$ be the set of finite non-increasing $\mathbb{N} \cup \{\infty\}$-valued sequences with infinite sum (and therefore at least one infinite part). In other words, define

$$\mathcal{P}_\infty := \Big\{\lambda = (\lambda_1, \dots, \lambda_p) \in \big(\mathbb{N} \cup \{\infty\}\big)^p \,:\, p \geq 1 \text{ and } \infty = \lambda_1 \geq \dots \geq \lambda_p > 0\Big\}.$$

Set $\mathcal{P}_{<\infty} := \bigcup_{n\geq 0} \mathcal{P}_n$ and $\mathcal{P} := \mathcal{P}_{<\infty} \cup \mathcal{P}_\infty$.

Let $\lambda = (\lambda_1, \dots, \lambda_p)$ be in $\mathcal{P}$. We will use the following notations:

- Let $p(\lambda) := p$ be its *length* and $\|\lambda\| = \lambda_1 + \dots + \lambda_p$ its sum (with the conventions $p(\varnothing) = \|\varnothing\| = 0$).

- For $k \in \mathbb{N} \cup \{\infty\}$, let $m_k(\lambda) := \sum_i \mathbb{1}_{\lambda_i = k}$ be the number of occurrences of k in the partition λ.
- For a non-negative integer K, set $\lambda \wedge K := (\lambda_1 \wedge K, \ldots, \lambda_p \wedge K)$. This finite partition will be called the *truncation* of λ at level K.

We endow $\mathcal{P}$ with an ultra-metric distance defined similarly to $\mathrm{d}_{\mathrm{loc}}$. For all λ and μ in $\mathcal{P}$, let

$$\mathrm{d}_{\mathcal{P}}(\lambda, \mu) := \exp\big[-\inf\{K \geq 0 \,:\, \lambda \wedge K \neq \mu \wedge K\}\big].$$

Lemma 2.2. (i) *The function* $\mathrm{d}_{\mathcal{P}}$ *is an ultra-metric distance,*

(ii) *The metric space* $(\mathcal{P}, \mathrm{d}_{\mathcal{P}})$ *is Polish.*

Remark 2.3. For all λ and μ in $\mathcal{P}$ and $K \geq 0$, $\lambda \wedge K = \mu \wedge K$ iff $\mathrm{d}_{\mathcal{P}}(\lambda, \mu) < \mathrm{e}^{-K}$. In particular, $\mathrm{d}_{\mathcal{P}}(\lambda, \mu) = 1$ iff $\lambda \wedge 0 \neq \mu \wedge 0$ in which case $p(\lambda) \neq p(\mu)$.

Proof. (i) Clearly, $\mathrm{d}_{\mathcal{P}}$ is symmetric and $\mathrm{d}_{\mathcal{P}}(\lambda, \mu) = 0$ iff $\lambda = \mu$. Hence, we only need to prove that $\mathrm{d}_{\mathcal{P}}$ satisfies the ultra-metric triangular inequality. Let λ, μ and ν be in $\mathcal{P}$ and assume that $\mathrm{d}_{\mathcal{P}}(\lambda, \nu) > \mathrm{d}_{\mathcal{P}}(\lambda, \mu) \vee \mathrm{d}_{\mathcal{P}}(\mu, \nu)$. Then, there exists $K \geq 0$ such that $\lambda \wedge K = \mu \wedge K = \nu \wedge K$ and $\lambda \wedge K \neq \nu \wedge K$, which is absurd. Consequently, $\mathrm{d}_{\mathcal{P}}(\lambda, \nu) \leq \mathrm{d}_{\mathcal{P}}(\lambda, \mu) \vee \mathrm{d}_{\mathcal{P}}(\mu, \nu)$.

(ii) Observe that $\mathcal{P} \subset \bigcup_{n \geq 0} (\mathbb{N} \cup \{\infty\})^n$ and is as a result both countable and separable. Therefore, it only remains to show that it is complete.

Let $(\lambda_n)_n$ be a Cauchy sequence with respect to $\mathrm{d}_{\mathcal{P}}$. By assumption, there exists an increasing sequence $(n_K)_K$ such that for all $K \geq 0$, $\lambda_n \wedge K = \lambda_m \wedge K$ when $n, m \geq n_K$. In particular, there exists a constant $p \geq 0$ such that $p(\lambda_{n_K}) = p$ for all K. Furthermore, notice that for all $i = 1, \ldots, p$, the sequence $[\lambda_{n_K}(i) \wedge K]_K$ is non-decreasing. For each $i = 1, \ldots, p$, set $\lambda(i) := \sup_K \lambda_{n_K}(i) \wedge K \leq \infty$. Clearly, $\lambda := [\lambda(1), \ldots, \lambda(p)]$ is in $\mathcal{P}$ and is such that $\mathrm{d}_{\mathcal{P}}(\lambda_n, \lambda) \to 0$ when $n \to \infty$. This proves that $(\mathcal{P}, \mathrm{d}_{\mathcal{P}})$ is indeed complete. □

Lemma 2.4. *Let* $(\Lambda_n)_{n \geq 1}$ *and* Λ *be* $\mathcal{P}$*-valued random variables. Then,* Λ_n *converges to* Λ *in distribution with respect to* $\mathrm{d}_{\mathcal{P}}$ *iff for all* λ *in* $\mathcal{P}_{<\infty}$ *and all* $K \geq 0$*, we have* $\mathbb{P}[\Lambda_n \wedge K = \lambda \wedge K] \to \mathbb{P}[\Lambda \wedge K = \lambda \wedge K]$ *as* $n \to \infty$.

Proof. Uses the same arguments as the proof of Lemma 2.1 (recall that $\mathrm{d}_{\mathcal{P}}$ is an ultra-metric and use [16, Theorem 2.3]). □

Remark 2.5. Elements of $\mathcal{P}_{<\infty}$ are closely related to elements of $\mathtt{T}$. Indeed, if $\mathtt{t}$ is a finite tree which can be written as the concatenation of p trees $\mathtt{t}_1, \ldots, \mathtt{t}_p$, i.e. $\mathtt{t} = [\![\mathtt{t}_1, \ldots, \mathtt{t}_p]\!]$, then the decreasing rearrangement of $\#\mathtt{t}_1, \ldots, \#\mathtt{t}_p$ is a partition of n when $\mathtt{t}$ has $n+1$ vertices (the root plus n descendants). We will write $\Lambda(\mathtt{t}) := (\#\mathtt{t}_1, \ldots, \#\mathtt{t}_p)^{\downarrow}$, where $(x_1, \ldots, x_k)^{\downarrow}$ stands for the decreasing rearrangement of $(x_1, \ldots, x_p)$, and call $\Lambda(\mathtt{t})$ the *partition at the root* or *first split* of $\mathtt{t}$.

Similarly, if we consider leaves instead of vertices, then $\Lambda^{\mathcal{L}}(\mathtt{t}) := (\#_{\mathcal{L}}\mathtt{t}_1, \ldots, \#_{\mathcal{L}}\mathtt{t}_p)^{\downarrow}$ is a partition of n when $\mathtt{t}$ has n leaves.

In this article, we will often have to consider sequences of random partitions $\Lambda_n \in \mathcal{P}_n$ that will weakly converge to a limit partition $\Lambda_\infty \in \mathcal{P}_\infty$ such that, $m_\infty(\Lambda_\infty) = 1$ a.s.. In this particular setting, the weak convergence can be characterised as follows.

Lemma 2.6. *For all* $1 \leq n \leq \infty$*, let* q_n *be a probability measure on* $\mathcal{P}_n$ *and assume that* $q_\infty(m_\infty = 1) = 1$*. Then,* $q_n \Rightarrow q_\infty$ *with respect to* $\mathrm{d}_{\mathcal{P}}$ *iff for all* λ *in* $\mathcal{P}_{<\infty}$ *we have* $q_n(n - \|\lambda\|, \lambda) \to q_\infty(\infty, \lambda)$ *as* $n \to \infty$.

Proof. $\Rightarrow$ Let $\lambda = (\lambda_1, \ldots, \lambda_p)$ be in $\mathcal{P}_{<\infty}$ and $K > \lambda_1$. In light of Lemma 2.4,

$$\begin{aligned} q_n(n - \|\lambda\|, \lambda) &= q_n\big(\mu \in \mathcal{P}_n : \mu \wedge K = (K, \lambda) \wedge K\big) \\ &\xrightarrow[n\to\infty]{} q_\infty\big(\mu \in \mathcal{P}_\infty : \mu \wedge K = (K, \lambda) \wedge K\big) = q_\infty(\infty, \lambda). \end{aligned}$$

$\Leftarrow$ For fixed $K \geq 0$ and λ in $\mathcal{P}_{<\infty}$, Fatou's lemma ensures that

$$\begin{aligned} \liminf_{n\to\infty}\ q_n\big(\mu \in \mathcal{P}_n : \mu \wedge K = \lambda \wedge K\big) &= \liminf_{n\to\infty}\ \textstyle\sum_{\nu\in\mathcal{P}_{<\infty}} \mathbb{1}_{(\infty,\nu)\wedge K = \lambda\wedge K}\, q_n(n - \|\nu\|, \nu) \\ &\geq \textstyle\sum_{\nu\in\mathcal{P}_{<\infty}} \mathbb{1}_{(\infty,\nu)\wedge K = \lambda\wedge K}\, q_\infty(\infty, \nu) = q_\infty\big(\mu \in \mathcal{P}_\infty : \mu \wedge K = \lambda \wedge K\big). \end{aligned}$$

Similarly,

$$\liminf_{n\to\infty}\ q_n\big(\mu \in \mathcal{P}_n : \mu \wedge K \neq \lambda \wedge K\big) \geq q_\infty\big(\mu \in \mathcal{P}_\infty : \mu \wedge K \neq \lambda \wedge K\big).$$

As a result and thanks to Lemma 2.4, we get that $q_n \Rightarrow q_\infty$. □

2.2 The Markov-branching property

2.2.1 Finite Markov branching trees

We will now follow [30, Section 1.2] and define two types of family of probability measures on the set of finite unordered rooted trees, satisfying the Markov branching property discussed in the Introduction.

Informally, for a given sequence $q = (q_n)$ of probability measures respectively supported by $\mathcal{P}_n$ (referred to as "first-split distributions" in the Introduction), we want to define a sequence $\mathrm{MB}^q = (\mathrm{MB}^q_n)_n$ of probability measures on the set of finite trees where

- For all n, MB^q_n is supported by the set of trees with n vertices,
- A tree T with distribution MB^q_n is such that
  - The decreasing rearrangement $\Lambda(T)$ of the sizes of the sub-trees above its root is distributed according to q_{n-1},
  - Conditionally on $\Lambda(T) = (\lambda_1, \ldots, \lambda_p)$, the p sub-trees of T above its root are independent with respective distributions $\mathrm{MB}^q_{\lambda_i}$.

Similarly, if $q = (q_n)_n$ is a sequence of probability measures respectively on $\mathcal{P}_n$, we will define a sequence $\mathrm{MB}^{\mathcal{L},q}$ satisfying the same Markov branching property where we count leaves instead of vertices to measure the size of a tree.

Markov branching tree with n vertices First of all, set $\mathcal{N}$ an infinite subset of $\mathbb{N}$ with $1 \in \mathcal{N}$. This set will index the possible number of vertices of the trees we want to generate. Let $q = (q_{n-1})_{n\in\mathcal{N}}$ be a sequence of probability measures such that $q_0(\varnothing) = 1$, $q_1[(1)] = 1$ (if $2 \in \mathcal{N}$), and for all n in $\mathcal{N}$, $n \geq 2$, q_{n-1} is supported by the set $\{\lambda \in \mathcal{P}_{n-1} : \lambda_i \in \mathcal{N}, i = 1, \ldots, p(\lambda)\}$.

Remark 2.7. This last condition comes from the fact that if T is distributed according to MB^q_n, the blocks of $\Lambda(T)$ need to be in $\mathcal{N}$ because the distributions of the corresponding sub-trees belong to the family $(\mathrm{MB}^q_k)_{k\in\mathcal{N}}$.

We now detail a recursive construction for MB^q. Let $\mathrm{MB}^q_1(\{\varnothing\}) = 1$ and for $n \geq 2$, proceed by a decreasing induction as follows:

- Let Λ have distribution q_{n-1},
- Conditionally on $\Lambda = (\lambda_1, \ldots, \lambda_p) \in \mathcal{P}_{n-1}$, let $(T_1, \ldots, T_p)$ be independent random trees such that T_i is distributed according to $\mathrm{MB}^q_{\lambda_i}$ for each $1 \leq i \leq p$,
- Define MB^q_n as the law of the concatenation of these trees, i.e. that of $[\![T_1, \ldots, T_p]\!]$.

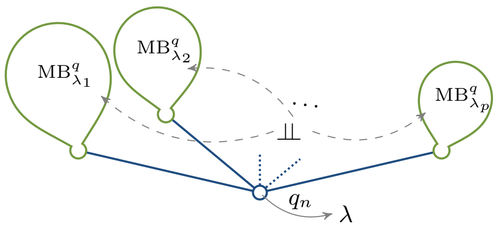

Figure 4: The construction of a tree with distribution MB^q_n.

Markov branching tree with n leaves Similarly, fix an infinite subset $\mathcal{N}$ of $\mathbb{N}$ such that $1 \in \mathcal{N}$ (corresponding to the possible number of leaves of the trees we will generate) and let $q = (q_n)_{n\in\mathcal{N}}$ be such that:

- q_1 is a probability measure on $\{\varnothing, (1)\}$ with $q_1(1) < 1$,
- For all $n > 1$ in $\mathcal{N}$, q_n is a probability measure supported by the set $\{\lambda \in \mathcal{P}_n \,:\, \lambda_i \in \mathcal{N},\, i = 1, \dots, p(\lambda)\}$.

To define $\mathrm{MB}^{\mathcal{L},q}$, we will proceed by the same recursive method used for MB^q: first choose how the size is split between the children sub-trees of the root, and then generate the said sub-trees adequately. However, if for some n in $\mathcal{N}$ we have $q_n(n) = 1$, the recursion will be endless. For this reason, we also require that for all n in $\mathcal{N}$, $q_n(n) < 1$ (i.e. with positive probability, a tree "splits" into smaller trees).

Let $\mathrm{MB}^{\mathcal{L},q}_1$ be the distribution of a branch of geometric length with parameter $1-q_1(1)$, i.e. $\mathrm{MB}^{\mathcal{L},q}_1(\mathsf{b}_k) = q_1(1)^k[1 - q_1(1)]$ for all $k \geq 0$. For $n > 1$, we do as follows:

- Let T_0 be a branch with geometric length with parameter $1 - q_n(n)$ and call U its leaf,
- Let Λ have distribution q_n conditioned on the event $\{m_n = 0\}$,
- Conditionally on $\Lambda = (\lambda_1, \dots, \lambda_p)$, let $(T_1, \dots, T_p)$ be independent random trees respectively distributed according to $\mathrm{MB}^q_{\lambda_i}$ for $1 \leq i \leq p$,
- Graft the concatenation of these trees on the leaf U of T_0, i.e. set $T := T_0 \otimes \big(U, [\![T_1, \dots, T_{p(\Lambda)}]\!]\big)$ and let $\mathrm{MB}^{\mathcal{L},q}_n$ be the distribution of T.

2.2.2 Infinite Markov branching trees

Using the same principle as before (split the number of vertices above the root and generate independent sub-trees with corresponding sizes) we will define a probability measure supported by the set of infinite trees which satisfies a version of the Markov branching property. Let $\mathcal{N}$ and $q = (q_{n-1})_{n\in\mathcal{N}}$ satisfy the conditions exposed in the construction of the sequence MB^q.

In order to lighten notations, for any finite decreasing sequence of integers $\lambda = (\lambda_1, \dots, \lambda_p)$, we define MB^q_λ as the distribution of the concatenation of independent $\mathrm{MB}^q_{\lambda_i}$-distributed trees. More precisely:

- Let $\mathrm{MB}^q_\varnothing$ be the Dirac measure on the tree with a single vertex (its root), namely $\mathrm{MB}^q_\varnothing = \delta_{\{\varnothing\}}$,
- For any $\lambda \in \mathcal{P}_{<\infty}$ with $p = p(\lambda) > 0$ and $\lambda_i \in \mathcal{N}$ for $i = 1, \dots, p$, let $(T_1, \dots, T_p)$ be independent trees with respective distributions $\mathrm{MB}^q_{\lambda_i}$ for all $i = 1, \dots, p$. Set MB^q_λ as the distribution of the concatenation of these trees.

Observe that when $p(\lambda) = 1$, a tree with distribution MB^q_λ is obtained by attaching an edge "under" the root of a $\mathrm{MB}^q_{\lambda_1}$-distributed tree.

Consider q_∞, a probability measure on $\mathcal{P}_\infty$ supported by the set

$$\left\{\lambda \in \mathcal{P}_\infty : \lambda_i \in \mathcal{N} \cup \{\infty\}, i = 1, \ldots, p(\lambda)\right\}$$

and let Λ follow q_∞. Let T° be a Galton-Watson tree with offspring distribution the law of $m_\infty(\Lambda)$. Conditionally on T°, let $(\Lambda_u, T_u)_{u\in T^\circ}$ be independent pairs and such that:

- Λ_u has the same distribution as Λ conditioned on the event $m_\infty(\Lambda) = c_u(T^\circ)$,
- Conditionally on $\Lambda_u = (\infty, \ldots, \infty, \lambda)$ with λ in $\mathcal{P}_{<\infty}$, T_u follows MB^q_λ.

Then, for every vertex u in T°, graft the corresponding tree T_u on T° at u. Let T be the tree hence obtained, i.e. set $T := T^\circ \bigotimes_{u\in T^\circ}(u, T_u)$. Finally, call $\mathrm{MB}^{q,q_\infty}_\infty$ the distribution of T.

Remark 2.8.
- Suppose that $q_\infty(m_\infty = 1) = 1$. In this case, the construction of $\mathrm{MB}^{q,q_\infty}_\infty$ is much simpler: the tree T° is simply the infinite branch and the family $(\Lambda_{\mathtt{v}_n}, T_{\mathtt{v}_n})_{n\geq 0}$ is i.i.d.. In particular, T a.s. has a unique *infinite spine*, i.e. a unique infinite non-backtracking path originating from the root.
- A tree T with distribution $\mathrm{MB}^{q,q_\infty}_\infty$ satisfies the Markov branching property: conditionally on $\Lambda(T)$, the sub-trees of T above its root are independent and their respective distributions are either $\mathrm{MB}^{q,q_\infty}_\infty$ or in the family $(\mathrm{MB}^q_n)_{n\in\mathcal{N}}$, depending on their sizes.
- The same exact construction can be used to define a measure $\mathrm{MB}^{\mathcal{L},q,q_\infty}_\infty$.

2.3 Local limits of Markov-branching trees

Let q be the sequence of first-split distributions associated to a Markov-branching family MB^q (respectively $\mathrm{MB}^{\mathcal{L},q}$). Suppose q_∞ is a probability measure on $\mathcal{P}_\infty$ supported by the set of sequences λ such that for all $i = 1, \ldots, p(\lambda)$, λ_i is either infinite or in $\mathcal{N}$. The aim of this section is to expose suitable conditions on q and q_∞ such that MB^q_n converges weakly to $\mathrm{MB}^{q,q_\infty}_\infty$ (or $\mathrm{MB}^{\mathcal{L},q}_n \Rightarrow \mathrm{MB}^{\mathcal{L},q,q_\infty}_\infty$) for the local limit topology.

Theorem 2.9. *Suppose that when n goes to infinity, q_n converges weakly to q_∞ with respect to the topology induced by $\mathrm{d}_\mathcal{P}$. Then, with respect to $\mathrm{d}_{\mathrm{loc}}$, $\mathrm{MB}^q_n \Rightarrow \mathrm{MB}^{q,q_\infty}_\infty$ (respectively $\mathrm{MB}^{\mathcal{L},q}_n \Rightarrow \mathrm{MB}^{\mathcal{L},q,q_\infty}_\infty$).*

In many cases, the infinite trees we will consider will have a unique infinite spine, which corresponds to $q_\infty(m_\infty = 1) = 1$ and the particular construction mentioned in Remark 2.8. In this situation, we may use Theorem 2.9 alongside Lemma 2.6 to get the following corollary.

Corollary 2.10. *Assume that q_∞ is such that $q_\infty(m_\infty = 1) = 1$ and suppose that for any finite partition λ in $\mathcal{P}_\infty$ we have $q_n(n - \|\lambda\|, \lambda) \to q_\infty(\infty, \lambda)$. Then, $\mathrm{MB}^q_n \Rightarrow \mathrm{MB}^{q,q_\infty}_\infty$ (or $\mathrm{MB}^{\mathcal{L},q}_n \Rightarrow \mathrm{MB}^{\mathcal{L},q,,q_\infty}_\infty$) with respect to the local limit topology.*

Proof of Theorem 2.9. For all n in $\mathcal{N} \cup \{\infty\}$, let T_n follow MB^q_n and Λ_{n-1} follow q_{n-1}. To prove this theorem, we will use Lemma 2.1 and proceed by induction on R. First, it clearly holds that for every tree $\mathtt{t}$, $\mathtt{t}|_0 = \{\varnothing\} = T_n|_0 = T_\infty|_0$ a.s..

Let R be a non-negative integer and suppose that for any $\mathtt{s} \in \mathtt{T}$, $\mathbb{P}[T_n|_R = \mathtt{s}|_R] \to \mathbb{P}[T_\infty|_R = \mathtt{s}|_R]$ as $R \to \infty$. Fix $\mathtt{t} \in \mathtt{T}$ and set $d := c_\varnothing(\mathtt{t})$, the number of children of its root. We may write $\mathtt{t}|_{R+1} = [\![\mathtt{t}_1, \ldots, \mathtt{t}_d]\!]$ for some $\mathtt{t}_1, \ldots, \mathtt{t}_d$ in $\mathtt{T}$ with height R or less. When $n > 1$, we can similarly write T_n as the concatenation of its sub-trees: let

$T_n = [\![T_n^{(1)}, \dots, T_n^{(p)}]\!]$ where $p = c_\varnothing(T_n)$. With these notations, for all $n > 1$ in $\mathcal{N} \cup \{\infty\}$, we have

$$\mathbb{P}\big[T_n|_{R+1} = \mathtt{t}|_{R+1}\big] = \mathbb{P}\Big[\big(c_\varnothing(T_n) = d\big) \cap \big(\exists \sigma \in \mathfrak{S}_d : T_n^{(i)}|_R = \mathtt{t}_{\sigma\cdot i}, i = 1, \dots, d\big)\Big],$$

where $\mathfrak{S}_d$ denotes the set of permutations of $\{1, \dots, d\}$. There exists a subset S of $\mathfrak{S}_d$ such that for any $\sigma \in \mathfrak{S}_d$ there is a unique $\tau \in S$ satisfying $\mathtt{t}_{\sigma\cdot i} = \mathtt{t}_{\tau\cdot i}$ as elements of $\mathtt{T}$ for all $i = 1, \dots, d$. Observe that S only depends on $\mathtt{t}$ and the (arbitrary) labelling of its sub-trees. Then,

$$\begin{aligned}\mathbb{P}\big[T_n|_{R+1} = \mathtt{t}|_{R+1}\big] &= \sum_{\sigma \in S} \mathbb{P}\Big[\big(c_\varnothing(T_n) = d\big) \cap \big(T_n^{(i)}|_R = \mathtt{t}_{\sigma\cdot i}, i = 1, \dots, d\big)\Big] \\ &= \sum_{\sigma \in S} \mathbb{E}\Big[\textstyle\prod_{i=1}^d \mathbb{P}\Big[T_{\Lambda_n(i)}|_R = \mathtt{t}_{\sigma\cdot i} \,\Big|\, \Lambda_{n-1}\Big] \mathbb{1}_{p(\Lambda_{n-1})=d}\Big] \\ &= \sum_{\sigma \in S} \int_{\mathcal{P}} \textstyle\prod_{i=1}^d \mathbb{P}\big[T_{\lambda_i}|_R = \mathtt{t}_{\sigma\cdot i}\big] \mathbb{1}_{p(\lambda)=d}\, q_{n-1}(\mathrm{d}\lambda),\end{aligned}$$

where we have used the Markov branching property. Our induction assumption implies in particular that for all $i = 1, \dots, d$ and $\mathtt{s}$ in $\mathtt{T}$ with height R or less, the function $\mathcal{P} \to [0, 1]$, $\lambda \mapsto \mathbb{P}[T_{\lambda_i}|_R = \mathtt{s}]\, \mathbb{1}_{p(\lambda)=d}$ is continuous. As a result, $\mathbb{P}[T_n|_{R+1} = \mathtt{t}|_{R+1}]$ may be expressed as the integral against q_{n-1} of a finite sum of continuous functions. Therefore, since $q_n \Rightarrow q_\infty$,

$$\mathbb{P}\big[T_n|_{R+1} = \mathtt{t}|_{R+1}\big] \xrightarrow[n\to\infty]{} \mathbb{P}\big[T_\infty|_{R+1} = \mathtt{t}|_{R+1}\big].$$

We proceed in the same way to prove the claim on $\mathrm{MB}^{\mathcal{L},q}$ trees. $\square$

In the next proposition, we prove that the condition "$q_n \Rightarrow q_\infty$" in Theorem 2.9 is optimal for MB^q trees.

Proposition 2.11. *Let $q = (q_{n-1})_{n\in\mathcal{N}}$ be the sequence of first split distributions associated to a family MB^q of Markov branching trees with given number of vertices. If there exists a probability measure q_∞ on $\mathcal{P}_\infty$ such that MB_n^q converges weakly to $\mathrm{MB}_\infty^{q,q_\infty}$ for the local limit topology, then $q_{n-1} \Rightarrow q_\infty$ in the sense of the $\mathrm{d}_\mathcal{P}$ topology.*

Proof. Observe that for all $K \geq 0$ and $\mathtt{t}, \mathtt{s} \in \mathtt{T}$, if $\mathtt{t}|_K = \mathtt{s}|_K$ then $\Lambda(\mathtt{t}) \wedge K = \Lambda(\mathtt{s}) \wedge K$. As a result, $\mathrm{d}_\mathcal{P}[\Lambda(\mathtt{t}), \Lambda(\mathtt{s})] \leq \mathrm{d}_{\mathrm{loc}}(\mathtt{t}, \mathtt{s})$ which proves in particular that $\Lambda : \mathtt{T} \to \mathcal{P}$ is a continuous function. Consequently, since for all possibly infinite n, $\Lambda(T_n)$ has distribution q_{n-1}, in the sense of the $\mathrm{d}_\mathcal{P}$ topology we have $q_{n-1} \Rightarrow q_\infty$ when $n \to \infty$. $\square$

3 Background on scaling limits

In this section, we will introduce the framework needed to consider the scaling limits of both finite and infinite Markov branching trees as well as the corresponding limiting objects: self-similar fragmentation trees with or without immigration. Afterwards, we will also give a few useful results on point processes related to our models of trees.

3.1 $\mathbb{R}$-trees and the GHP topology

To talk about scaling limits of discrete trees, we need to introduce a continuous analogue. We use the framework of $\mathbb{R}$-trees. An *$\mathbb{R}$-tree* (or *real tree*) is a metric space (T, d) such that for all x and y in T:

- There exists a unique isometry $\varphi : [0, d(x, y)] \to T$ such that $\varphi(0) = x$ and $\varphi[(d(x, y)] = y$,

- If $\gamma : [0,1] \to T$ is a continuous injection with $\gamma(0) = x$ and $\gamma(1) = y$, then the image of γ is the same as that of φ, i.e. $\operatorname{Im}\gamma = \operatorname{Im}\varphi =: [\![x, y]\!]$.

This roughly means that any two points in an $\mathbb{R}$-tree can be continuously joined by a single path, up to its reparametrisation, which is akin to the acyclic nature of discrete trees.

To compare two such objects, we will use the Gromov-Hausdorff-Prokhorov distance. More precisely, we will follow the definition from [4] and extend it in a way similar to that of [3].

For any metric space (X, d) let $\mathcal{M}_f(X)$ be the set of all finite non-negative Borel measures on X and $\mathcal{M}(X)$ be the set of all non-negative and *boundedly finite* Borel measures on X, i.e. non-negative Borel measures μ on X such that $\mu(A) < \infty$ for all measurable bounded $A \subset X$.

A *pointed* metric space is a 3-tuple (X, d, ρ) where (X, d) is a metric space and $\rho \in X$ is a fixed point, which we will call its *root*. For any $x \in X$, set $|x| := \mathrm{d}(\rho, x)$ the *height* of x in (X, d, ρ), and let $|X| := \sup_{x\in X} |x|$ be the height of X.

We will call *pointed weighted metric space* any 4-tuple $\mathbf{X} = (X, \mathrm{d}, \rho, \mu)$ where (X, d) is a metric space, $\rho \in X$ is its root and μ is a boundedly finite Borel measure on X.

Remark 3.1. If $\mathbf{X}$ is a pointed weighted metric space, we will implicitly write $\mathbf{X} = (X, \mathrm{d}_X, \rho_X, \mu_X)$ unless otherwise stated.

Two pointed weighted metric spaces $\mathbf{X}$ and $\mathbf{Y}$ will be called *GHP-isometric* if there exists a bijective isometry $\Phi : X \to Y$ such that $\Phi(\rho_X) = \rho_Y$ and $\mu_X \circ \Phi^{-1} = \mu_Y$. Let $\mathbb{K}$ be the set of GHP-isometry classes of *compact* pointed weighted metric spaces.

3.1.1 Comparing compact metric spaces

Let $\mathbf{X}$ and $\mathbf{Y}$ be two pointed weighted compact metric spaces. A *correspondence* between $\mathbf{X}$ and $\mathbf{Y}$ is a measurable subset C of $X \times Y$ which contains (ρ_X, ρ_Y) such that for any $x \in X$ there exists $y \in Y$ with $(x, y) \in C$ and conversely, for any $y \in Y$ there is $x \in X$ such that $(x, y) \in C$. We will denote by $\mathrm{C}(\mathbf{X}, \mathbf{Y})$ (or $\mathrm{C}(X, Y)$ with a slight abuse of notation) the set of all pointed correspondences between $\mathbf{X}$ and $\mathbf{Y}$. For any $C \in \mathrm{C}(\mathbf{X}, \mathbf{Y})$, let its *distortion* be defined as follows:

$$\operatorname{dis}_{\mathbf{X},\mathbf{Y}} C := \sup\left\{|\mathrm{d}_X(x, x') - \mathrm{d}_Y(y, y')| \;:\; (x, y), (x', y') \in C\right\}.$$

When the setting is clear, we will simply write $\operatorname{dis} C := \operatorname{dis}_{\mathbf{X},\mathbf{Y}} C$. Observe that $\operatorname{dis} C \leq 2\left(|X| \vee |Y|\right) < \infty$ and that $\operatorname{dis} C \geq \big||X| - |Y|\big|$.

For any finite Borel measure π on $X \times Y$, we define its *discrepancy* with respect to μ_X and μ_Y as:

$$\mathrm{D}(\pi; \mu_X, \mu_Y) := \|\mu_X - \pi \circ p_X^{-1}\|_{\mathrm{TV}} + \|\mu_Y - \pi \circ p_Y^{-1}\|_{\mathrm{TV}}$$

where $\|\cdot\|_{\mathrm{TV}}$ is the total variation norm, and $p_X : (x, y) \in X \times Y \mapsto x$, $p_Y : (x, y) \in X \times Y \mapsto y$ are the canonical projections from $X \times Y$ to X and Y respectively. The definition of the total variation norm and the triangular inequality give $\mathrm{D}(\pi; \mu_X, \mu_Y) \geq |\mu_X(X) - \mu_Y(Y)|$.

Following [4, Section 2.1], we define the *Gromov-Hausdorff-Prokhorov* distance (or GHP distance for short) between two pointed weighted compact metric spaces $\mathbf{X}$ and $\mathbf{Y}$ as:

$$\mathrm{d}_{\mathrm{GHP}}(\mathbf{X}, \mathbf{Y}) := \inf\left\{\frac{1}{2}\operatorname{dis} C \vee \mathrm{D}(\pi; \mu_X, \mu_Y) \vee \pi(C^c) \;:\; C \in \mathrm{C}(X, Y), \pi \in \mathcal{M}(X \times Y)\right\}$$

where $C^c = X \times Y \setminus C$.

Remark 3.2. Observe that $\mathrm{d}_{\mathrm{GHP}}(\mathbf{X},\mathbf{Y}) \le \big(|X| \vee |Y|\big) \vee \big(\mu_X(X)+\mu_Y(Y)\big)$ and is consequently finite. Moreover, $\mathrm{d}_{\mathrm{GHP}}(\mathbf{X},\mathbf{Y}) \ge \big(1/2 \cdot \big||X|-|Y|\big|\big) \vee \big|\mu_X(X)-\mu_Y(Y)\big|$. Therefore, the functions $\mathbb{K} \to \mathbb{R}_+$, $\mathbf{X} \mapsto ||X$ and $\mathbf{X} \mapsto \mu_X(X)$ are both continuous with respect to $\mathrm{d}_{\mathrm{GHP}}$.

As was mentioned in [4, Section 2.1], $\mathrm{d}_{\mathrm{GHP}}$ is a well-defined distance on $\mathbb{K}$ and $(\mathbb{K}, \mathrm{d}_{\mathrm{GHP}})$ is both complete and separable and thus, Polish. Furthermore, it was also noted that $\mathrm{d}_{\mathrm{GHP}}$ gives rise to the same topology as the GHP distance defined in [3].

Rescaling compact metric spaces For all $m \ge 0$, let $\mathbf{0}^{(m)} := \big(\{\varnothing\}, d, \varnothing, m\delta_\varnothing\big) \in \mathbb{K}$ be the degenerate metric space only made out of its root on which a mass m is put. For a pointed weighted metric space $\mathbf{X}$ and any non-negative real numbers a and b, we will write $(aX, b\mu_X) := (X, a\mathrm{d}_X, \rho_X, b\mu_X)$. When $\mathbf{X}$ is in $\mathbb{K}$ and $\mu_X(X)=m$, we will use the convention $(0X, \mu_X) = \mathbf{0}^{(m)}$ (which makes sense since $(\varepsilon X, \mu_X)$ converges to $\mathbf{0}^{(m)}$ as ε goes to 0 with respect to $\mathrm{d}_{\mathrm{GHP}}$).

Lemma 3.3. *Let $\mathbf{X}$ and $\mathbf{Y}$ be two elements of $\mathbb{K}$. For any non-negative real numbers a, b, c and d:*

$$(i) \quad \mathrm{d}_{\mathrm{GHP}}\big((aX, b\mu_X), (cX, d\mu_X)\big) \le \big(|a-c|\,|X|\big) \vee \big(|b-d|\,\mu_X(X)\big),$$

and

$$(ii) \quad \mathrm{d}_{\mathrm{GHP}}\big((aX, b\mu_X), (aY, b\mu_Y)\big) \le (a \vee b)\,\mathrm{d}_{\mathrm{GHP}}(\mathbf{X},\mathbf{Y}).$$

Proof. (i) Let $C = \{(x,x) : x \in X\} \in \mathrm{C}(X,X)$. We have

$$\mathrm{dis}_{(aX,b\mu_X),\,(cX,d\mu_X)}\, C = \sup\big\{|a\,\mathrm{d}_X(x,y) - c\,\mathrm{d}_X(x,y)| : x,y \in X\big\} \le 2|a-c|\,|X|.$$

Let $\pi \in \mathcal{M}(X \times X)$ be defined for all measurable $A \subset X \times X$ by

$$\pi(A) := \textstyle\int_X \mathbb{1}_A\big((x,x)\big)\, b\mu_X(\mathrm{d}x).$$

Then $\mathrm{D}(\pi; b\mu_X, d\mu_X) = |b-d|\mu_X(X)$ and $\pi(C^c) = 0$.

(ii) For every correspondence $C \in \mathrm{C}(X,Y)$, we clearly have $\mathrm{dis}_{(aX,b\mu_X),\,(aY,b\mu_Y)}\, C = a\,\mathrm{dis}_{\mathbf{X},\mathbf{Y}}\, C$. No less clearly, for any finite measure π on $X \times Y$, $\mathrm{D}(b\pi; b\mu_X, b\mu_Y) = b\mathrm{D}(\pi; \mu_X, \mu_Y)$. □

Corollary 3.4. *The function $\mathbb{K} \times \mathbb{R}_+ \times \mathbb{R}_+ \longrightarrow \mathbb{K}$ defined by $(\mathbf{X}, a, b) \longmapsto (aX, b\mu_X)$ is continuous for the product topology.*

Concatenated compact metric spaces Let $(\mathbf{X}_i)_{i\in\mathcal{I}}$ be a countable family of pointed weighted metric spaces with $\mathbf{X}_i = (X_i, \mathrm{d}_i, \rho_i, \mu_i)$. Let $(X, \mathrm{d}, \rho, \mu)$ where:

- $X = \{\rho\} \sqcup \bigsqcup_{i\in\mathcal{I}} X_i$,
- d is defined by:
  - For all $i,j \in \mathcal{I}$, $\mathrm{d}(\rho,\rho_i) := \mathrm{d}(\rho_i,\rho_j) = 0$,
  - For all $i \in \mathcal{I}$, and $x,y \in X_i$, $\mathrm{d}(x,y) := \mathrm{d}_i(x,y)$,
  - For all $i \ne j$ and $x \in X_i$, $y \in X_j$, $\mathrm{d}(x,y) := \mathrm{d}_i(x,\rho_i) + \mathrm{d}_j(y,\rho_j)$,
- For any Borel subset A of X, $\mu(A) = \sum_{i\in\mathcal{I}} \mu_i(A \cap X_i)$.

With a slight abuse of notation, we will consider (X, d) to be the quotient metric space $X/\sim_{\mathrm{d}}$ where $x \sim_{\mathrm{d}} y$ iff $\mathrm{d}(x,y) = 0$. For each i in $\mathcal{I}$, we will also identify X_i with its image in X by the quotient map. Write $\mathbf{X} =: \langle \mathbf{X}_i\,;\, i \in \mathcal{I}\rangle$.

Remark 3.5. If $(\mathbf{T}_i)_{i\in\mathcal{I}}$ is a countable family of weighted $\mathbb{R}$-trees, then $\langle \mathbf{T}_i\,;\, i \in \mathcal{I}\rangle$ is clearly an $\mathbb{R}$-tree itself.

Lemma 3.6. *For all $i \geq 1$, let $\mathbf{X}_i = (X_i, \mathrm{d}_i, \rho_i, \mu_i)$ be in $\mathbb{K}$. Their concatenation $\langle \mathbf{X}_i\,;\, i \geq 1\rangle$ is an element of $\mathbb{K}$ iff the height $|X_i|$ of X_i goes to 0 as i goes to infinity and $\sum_{i\geq 1} \mu_i(X_i)$ is finite.*

Proof. Set $\mathbf{X} := \langle \mathbf{X}_i\,;\, i \geq 1\rangle$ and for all x in X and positive r, denote the open ball of X centred at x with radius r by $\mathrm{B}_X(x, r) := \{y \in X : \mathrm{d}_X(x, y) < r\}$. Similarly, for all $i \geq 1$ and $x \in X_i$, write $\mathrm{B}_i(x, r) := \{y \in X_i : \mathrm{d}_i(x, y) < r\}$. Clearly, the measure μ_X is finite iff the sum $\sum_{i\geq 1} \mu_i(X_i)$ is.

If $|X_i| \to 0$, then in particular, for all positive ε, there exists a integer n such that $\bigcup_{i>n} X_i \subset \mathrm{B}_X(\rho_X, \varepsilon)$. Moreover, since X_i is compact for all $i = 1, \dots, n$, we can find a finite ε-cover of X_i, i.e. a finite subset A_i of X_i such that $X_i \subset \bigcup_{x\in A_i} \mathrm{B}_i(x, \varepsilon)$. Set $A := \{\rho_X\} \cup A_1 \cup \dots \cup A_n$. Observe that it is finite and that $X \subset \bigcup_{x\in A} \mathrm{B}_X(x, \varepsilon)$. Since this holds for all positive ε, it follows that X is compact.

If $\limsup |X_i| > 0$, then there exists a positive ε such that $|X_i| > \varepsilon$ for infinitely many indices i. As a result, X cannot have a finite ε-cover, which implies that it is not compact. □

Lemma 3.7. *Let $\mathbf{X}_i$, $\mathbf{Y}_i$, $i \geq 1$ be in $\mathbb{K}$ and such that $\mathbf{X} := \langle \mathbf{X}_i\,;\, i \geq 1\rangle$ and $\mathbf{Y} := \langle \mathbf{Y}_i\,;\, i \geq 1\rangle$ both belong to $\mathbb{K}$. We have*

$$\mathrm{d}_{\mathrm{GHP}}\Big(\langle \mathbf{X}_i\,;\, i \geq 1\rangle, \langle \mathbf{Y}_i\,;\, i \geq 1\rangle\Big) \leq \textstyle\sum_{i\geq 1} \mathrm{d}_{\mathrm{GHP}}(\mathbf{X}_i, \mathbf{Y}_i).$$

Proof. Set $\mathbf{X} := \langle \mathbf{X}_i\,;\, i \geq 1\rangle$ and $\mathbf{Y} := \langle \mathbf{Y}_i\,;\, i \geq 1\rangle$. For all positive ε and $i \geq 1$, there exists a correspondence C_i in $\mathrm{C}(\mathbf{X}_i, \mathbf{Y}_i)$ and a finite Borel measure π_i on $X_i \times Y_i$ such that

$$\frac{1}{2}\,\mathrm{dis}\ C_i \vee \mathrm{D}(\pi_i; \mu_{X_i}, \mu_{Y_i}) \vee \pi_i(C_i^c) < \mathrm{d}_{\mathrm{GHP}}(\mathbf{X}_i, \mathbf{Y}_i) + 2^{-i}\varepsilon.$$

Set $C := \bigcup_{i\geq 1} C_i$, which is a correspondence between $\mathbf{X}$ and $\mathbf{Y}$. Let (x, y) and (x', y') be in C. If both (x, y) and (x', y') are in C_i for some i, then clearly, $|\mathrm{d}_X(x, x') - \mathrm{d}_Y(y, y')| \leq \mathrm{dis}\ C_i$. Otherwise, if $(x, y) \in C_i$ and $(x', y') \in C_j$ with $i \neq j$, then using the definition of d_X and d_Y as well as the triangular inequality, we get $|\mathrm{d}_X(x, x') - \mathrm{d}_Y(y, y')| \leq \mathrm{dis}\ C_i + \mathrm{dis}\ C_j$. Therefore, $1/2 \cdot \mathrm{dis}\ C \leq \sum_{i\geq 1} \mathrm{d}_{\mathrm{GHP}}(\mathbf{X}_i, \mathbf{Y}_i) + \varepsilon$.

For all $n \geq 0$, define the finite Borel measure $\pi^{(n)}$ on $X \times Y$ by $\pi^{(n)}(A) := \sum_{i=1}^n \pi_i\big[A \cap (X_i \times Y_i)\big]$ for any Borel set A. By definition,

$$\pi^{(n)}(C^c) = \textstyle\sum_{i=1}^n \pi_i\big[C^c \cap (X_i \times Y_i)\big] = \sum_{i=1}^n \pi_i[C_i^c] \leq \sum_{i\geq 1} \mathrm{d}_{\mathrm{GHP}}(\mathbf{X}_i, \mathbf{Y}_i) + \varepsilon.$$

Moreover, the discrepancy of $\pi^{(n)}$ with respect to μ_X and μ_Y satisfies

$$\begin{aligned}
\mathrm{D}(\pi^{(n)}; \mu_X, \mu_Y) &\leq \textstyle\sum_{i=1}^n \|\mu_{X_i} - \pi_i \circ p_{X_i}^{-1}\|_{\mathrm{TV}} + \|\mu_{Y_i} - \pi_i \circ p_{Y_i}^{-1}\|_{\mathrm{TV}} \\
&\qquad + \textstyle\sum_{j>n} \big(\|\mu_{X_j}\|_{\mathrm{TV}} + \|\mu_{Y_j}\|_{\mathrm{TV}}\big) \\
&\leq \textstyle\sum_{i=1}^n \mathrm{D}(\pi_i; \mu_{X_i}, \mu_{Y_i}) + \sum_{j>n} \big(\mu_{X_j}(X_j) + \mu_{Y_j}(Y_j)\big) \\
&\leq \textstyle\sum_{i\geq 1}^n \mathrm{d}_{\mathrm{GHP}}(\mathbf{X}_i, \mathbf{Y}_i) + \varepsilon + \sum_{j>n} \big(\mu_{X_j}(X_j) + \mu_{Y_j}(Y_j)\big).
\end{aligned}$$

In light of Lemma 3.6, there exists n such that $\sum_{i>n} \mu_{X_i}(X_i) + \mu_{Y_i}(Y_i) < \varepsilon$. As a result, $\mathrm{d}_{\mathrm{GHP}}(\mathbf{X}, \mathbf{Y}) \leq \sum_{i\geq 1} \mathrm{d}_{\mathrm{GHP}}(\mathbf{X}_i, \mathbf{Y}_i) + 2\varepsilon$ which holds for all positive ε. □

3.1.2 Extension to locally compact $\mathbb{R}$-trees

Let $\mathbf{X} = (X, \mathrm{d}_X, \rho_X, \mu_X)$ be a locally compact pointed weighted metric space such that μ_X is a boundedly finite measure. For all $r > 0$, let $\mathbf{X}|_r := \big(X|_r, \mathrm{d}_X, \rho_X, \mu_X|_r\big)$ where $X|_r := \{x \in X : |x| \leq r\}$ is the closed ball with radius r centred at ρ_X and $\mu_X|_r := \mathbb{1}_{X|_r}\mu_X$

is the restriction of μ_X to $X|_r$. Observe that if $r \leq R$, clearly $(\mathbf{X}|_R)|_r = (\mathbf{X}|_r)|_R = \mathbf{X}|_r$. We also define $\partial_r X := \{x \in X : |x| = r\}$.

For any two locally compact pointed weighted metric spaces $\mathbf{X}$ and $\mathbf{Y}$, we define the extended Gromov-Hausdorff-Prokhorov distance between them as:

$$\mathrm{D}_{\mathrm{GHP}}(\mathbf{X}, \mathbf{Y}) := \int_0^\infty \mathrm{e}^{-r} \Big[1 \wedge \mathrm{d}_{\mathrm{GHP}}\big(\mathbf{X}|_r, \mathbf{Y}|_r\big) \Big] \mathrm{d}r.$$

This definition closely resembles that of the GHP distance on locally compact metric spaces defined and studied in [3].

Remark 3.8. Let $\mathbf{X}$ and $\mathbf{Y}$ be two weighted locally compact pointed metric spaces. For all $R \geq 0$,

$$\begin{aligned} &\big|\mathrm{D}_{\mathrm{GHP}}(\mathbf{X}, \mathbf{Y}) - \mathrm{D}_{\mathrm{GHP}}(\mathbf{X}|_R, \mathbf{Y}|_R)\big| \\ &\qquad \leq \int_R^\infty \mathrm{e}^{-r} \underbrace{\big|1 \wedge \mathrm{d}_{\mathrm{GHP}}(\mathbf{X}|_r, \mathbf{Y}|_r) - 1 \wedge \mathrm{d}_{\mathrm{GHP}}(\mathbf{X}|_R, \mathbf{Y}|_R)\big|}_{\leq 1} \,\mathrm{d}r \leq \mathrm{e}^{-R}. \end{aligned}$$

Let $\mathbb{T}$ be the set of GHP-isometry classes of locally compact rooted $\mathbb{R}$-trees endowed with a boundedly finite Borel measure and $\mathbb{T}_c$, be that of compact weighted and rooted $\mathbb{R}$-trees (i.e. $\mathbb{T}_c = \mathbb{K} \cap \mathbb{T}$).

Proposition 3.9. (i) $\mathrm{D}_{\mathrm{GHP}}$ *is a metric on* $\mathbb{T}$*,*

(ii) *If* $\mathbf{T}_n$*,* $n \geq 1$ *and* $\mathbf{T}$ *belong to* $\mathbb{T}$*, then* $\mathrm{D}_{\mathrm{GHP}}(\mathbf{T}_n, \mathbf{T}) \to 0$ *iff* $\mathrm{d}_{\mathrm{GHP}}(\mathbf{T}_n|_r, \mathbf{T}|_r) \to 0$ *for all* $r \geq 0$ *with* $\mu_T(\partial_r T) = 0$*,*

(iii) $(\mathbb{T}, \mathrm{D}_{\mathrm{GHP}})$ *is a Polish metric space,*

(iv) $\mathrm{d}_{\mathrm{GHP}}$ *and* $\mathrm{D}_{\mathrm{GHP}}$ *induce the same topology on* $\mathbb{T}_c$*.*

Proof. (i) Since $\mathrm{d}_{\mathrm{GHP}}$ is a metric, $\mathrm{D}_{\mathrm{GHP}}$ is symmetric and clearly satisfies the triangular inequality. Moreover, if $\mathbf{T}$ and $\mathbf{T}'$ are two elements of $\mathbb{T}$ such that $\mathrm{D}_{\mathrm{GHP}}(\mathbf{T}, \mathbf{T}') = 0$, then for almost every $r \geq 0$, $\mathbf{T}|_r = \mathbf{T}'|_r$. In this case, $\mathbf{T}$ and $\mathbf{T}'$ are GHP-isometric (see [3, Proposition 5.3] for a similar proof).

(ii) Suppose $\mathrm{d}_{\mathrm{GHP}}(\mathbf{T}_n|_r, \mathbf{T}|_r) \to 0$ for all $r \geq 0$ with $\mu_T(\partial_r T) = 0$. Since μ_T is a locally finite measure, the set $\{r > 0 : \mu_T(\partial_r T) > 0\}$ is at most countable. As a result, the sequence $\big(r \mapsto 1 \wedge \mathrm{d}_{\mathrm{GHP}}(\mathbf{T}_n|_r, \mathbf{T}|_r)\big)_{n \geq 1}$ converges to $r \mapsto 0$ almost everywhere in $[0, \infty)$. Lebesgue's dominated convergence theorem then ensures that $\mathrm{D}_{\mathrm{GHP}}(\mathbf{T}_n, \mathbf{T}) \to 0$.

Assume $\mathrm{D}_{\mathrm{GHP}}(\mathbf{T}_n, \mathbf{T}) \to 0$ and let $r > 0$ be such that $\mu_T(\partial_r T) = 0$. For every subsequence $(n_k)_k$, there exists a sub-subsequence $(k_\ell)_\ell$ such that $1 \wedge \mathrm{d}_{\mathrm{GHP}}(\mathbf{T}_{n_{k_\ell}}|_t, \mathbf{T}|_t) \to 0$ for almost every $t \geq 0$ as $\ell \to \infty$. In particular, there exists $R > r$ such that $\mathrm{d}_{\mathrm{GHP}}(\mathbf{T}_{n_{k_\ell}}|_R, \mathbf{T}|_R) \to 0$.

Recall that $\mathrm{d}_{\mathrm{GHP}}$ is topologically equivalent to the metric on $\mathbb{K}$ studied in [3]. Therefore, in light of the proof of [3, Proposition 2.10], if τ_n, $n \geq 1$ and τ are compact $\mathbb{R}$-trees such that $\mathrm{d}_{\mathrm{GHP}}(\tau_n, \tau) \to 0$, then for all $r > 0$ such that $\mu_\tau(\partial_r \tau) = 0$, $\mathrm{d}_{\mathrm{GHP}}(\tau_n|_r, \tau|_r) \to 0$.

As a result, $\mathrm{d}_{\mathrm{GHP}}(\mathbf{T}_{n_{k_\ell}}|_r, \mathbf{T}|_r) \to 0$. From every subsequence $(n_k)_k$ we can thus extract a sub-subsequence $(k_\ell)_\ell$ such that $\mathrm{d}_{\mathrm{GHP}}(\mathbf{T}_{n_{k_\ell}}|_r, \mathbf{T}|_r) \to 0$, which is equivalent to saying that $\mathrm{d}_{\mathrm{GHP}}(\mathbf{T}_n|_r, \mathbf{T}|_r) \to 0$ as $n \to \infty$.

(iii) Since a criterion similar to (ii) holds for the metric studied in [3], this metric is topologically equivalent to $\mathrm{D}_{\mathrm{GHP}}$. As a result and thanks to Theorem 2.9 and Corollary 3.2 in [3], it follows that $(\mathbb{T}, \mathrm{D}_{\mathrm{GHP}})$ is completely metrisable and separable, i.e. it is Polish.

(iv) See Proposition 2.10 in [3]. □

Continuum grafting Let $\{(u_i,\tau_i) : i \in \mathcal{I}\}$ be a family of elements of $\mathbb{R}_+ \times \mathbb{T}_c$ such that $\mathcal{I}$ is at most countable. We define the $\mathbb{R}$-tree $\mathbf{G}\big(\{(u_i,\tau_i) : i \in \mathcal{I}\}\big)$ as

$$\mathbf{G}\Big(\big\{(u_i,\tau_i) : i \in \mathcal{I}\big\}\Big) := \Big(\mathbb{R}_+ \sqcup \bigsqcup_{i\in\mathcal{I}}\tau_i,\, \mathrm{d},\, 0,\, \mu\Big)$$

where the metric d is defined by:

- $\mathrm{d}[u,v] = |u-v|$ for all u and v in $\mathbb{R}_+$,
- $\mathrm{d}[x,y] = \mathrm{d}_{\tau_i}(x,y)$ for all $i \in \mathcal{I}$, x and y in τ_i,
- $\mathrm{d}[x,v] = \mathrm{d}_{\tau_i}(x,\rho_{\tau_i}) + |u_i - v|$ for all $i \in \mathcal{I}$, $x \in \tau_i$ and v in $\mathbb{R}_+$,
- $\mathrm{d}[x,y] = \mathrm{d}_{\tau_i}(x,\rho_{\tau_i}) + \mathrm{d}_{\tau_j}(y,\rho_{\tau_j}) + |u_i - u_j|$ for all $i \neq j \in \mathcal{I}$, $x \in \tau_i$ and $y \in \tau_j$,

and μ is the measure defined for all Borel set A by $\mu(A) := \sum_{i\in\mathcal{I}} \mu_{\tau_i}(A\cap\tau_i)$. The function $\mathbf{G}$ grafts the trees τ_i at height u_i for each $i \in \mathcal{I}$ on $\mathbb{R}_+$ which can be thought of as an infinite (continuous) branch. It is quite obvious that the weighted pointed metric space $\mathbf{G}\big(\{(u_i,\tau_i) : i \in \mathcal{I}\}\big)$ is an $\mathbb{R}$-tree.

Lemma 3.10. *Let $(\tau_i,\mathrm{d}_i,\rho_i,\mu_i)_{i\geq 1}$ be a sequence of compact weighted $\mathbb{R}$-trees and $(u_i)_{i\geq 1}$ be a sequence of non-negative real numbers. Then the weighted $\mathbb{R}$-tree $\mathbf{T} := \mathbf{G}\big(\{(u_i,\tau_i) : i \geq 1\}\big)$ is an element of $\mathbb{T}$ iff for all $K \geq 0$ and $\varepsilon > 0$ the set $\{i \geq 1 : u_i \leq K$ and $|\tau_i| \geq \varepsilon\}$ is finite and $\sum_{i\geq 1} \mathbb{1}_{u_i\leq K}\mu_{\tau_i}(\tau_i) < \infty$.*

Proof. For all x in T and positive r, denote by $\mathrm{B}_T(x,r) := \{y \in T : \mathrm{d}_T(x,y) < r\}$ the open ball of T centred at x with radius r and similarly for all $i \geq 1$ and $x \in \tau_i$, write $\mathrm{B}_i(x,r) := \{y \in \tau_i : \mathrm{d}_i(x,y) < r\}$.

$\Leftarrow$ Assume that for all $K \geq 0$, $\sum_{i\geq 1} \mathbb{1}_{u_i\leq K}\mu_{\tau_i}(\tau_i) < \infty$ and for all positive ε, that the set $\{i \geq 1 : u_i \leq K, |\tau_i| \geq \varepsilon\}$ is finite. Observe that for all non-negative K, $\mu_T(T|_K) \leq \sum_{i\geq 1} \mu_{\tau_i}(\tau_i)\mathbb{1}_{u_i\leq K}$. Therefore, the measure μ_T is boundedly finite and we only need to prove that T is locally compact.

Fix $K \geq 0$ and let ε be positive. For all $i \geq 1$, because τ_i is compact, there exists a finite subset A_i of τ_i such that $\tau_i \subset \bigcup_{x\in A_i} B_i(x,\varepsilon)$. To build an ε-cover of $T|_K$, first observe that if i is such that $u_i \leq K$ and $|\tau_i| < \varepsilon/2$, then τ_i is contained in some open ball with radius ε centred at some $n\varepsilon$ for $0 \leq n \leq K/\varepsilon$. Moreover, by assumption, there are only finitely many indices i with $u_i \leq K$ and $|\tau_i| \geq \varepsilon/2$. Therefore, if we let $A := \{n\varepsilon; 0 \leq n \leq K/\varepsilon\} \cup \{x \in A_i; i \geq 1, u_i \leq K, |\tau_i| \geq \varepsilon/2\}$, then A is finite and $T|_K$ is contained in $\bigcup_{x\in A} B_T(x,\varepsilon)$. As a result, $T|_K$ has a finite ε-cover for all positive ε which means that it is compact.

$\Rightarrow$ Suppose the set $\{i \geq 1 : u_i \leq K, |\tau_i| \geq \varepsilon\}$ is infinite for some $K \geq 0$ and positive ε. In particular, we can find an increasing sequence $(i_n)_n$ with $u_{i_n} \leq K$ and $|\tau_{i_n}| \geq \varepsilon$ for all n. For each $n \geq 1$, let x_n be in τ_{i_n} and such that $\varepsilon/2 < \mathrm{d}_{i_n}(\rho_{i_n}, x_n) \leq \varepsilon$. If $n \neq m$, the definition of the metric on T gives $\mathrm{d}_T(x_n,x_m) > \varepsilon$. Therefore, $(x_n)_n$ has no Cauchy subsequence which implies that $T|_{K+\varepsilon}$ isn't compact and that $\mathbf{T} \notin \mathbb{T}$.

Assume that $\{i \geq 1 : u_i \leq K, |\tau_i| \geq \varepsilon\}$ is finite for all $K \geq 0$ and $\varepsilon > 0$, and that $\sum_{i\geq 1} \mathbb{1}_{u_i\leq K_0}\, \mu_{\tau_i}(\tau_i)$ is infinite for some finite K_0. By assumption, $\{|\tau_i| : u_i \leq K_0\}$ is bounded by a finite constant R. Therefore, $\mu_T(T|_{K_0+R}) \geq \sum_{i\geq 1} \mathbb{1}_{u_i\leq K_0}\mu_{\tau_i}(\tau_i) = \infty$. Consequently, μ_T isn't boundedly finite and $\mathbf{T} \notin \mathbb{T}$. □

Remark 3.11. In the following, when we consider discrete trees, we will see them as $\mathbb{R}$-trees by replacing their edges by segments of length 1.

3.2 Fragmentation trees

In this section, we will present a few results on certain classes of $\mathbb{T}_c$- and $\mathbb{T}$-valued random variables: self-similar fragmentation trees (introduced in [29]) and self-similar fragmentation trees with immigration (see [27]).

3.2.1 Self-similar fragmentation trees

Let $\mathcal{S}^{\downarrow} := \{\mathbf{s} = (s_n)_{n\geq 1} \in \ell_1 \,:\, s_1 \geq s_2 \geq \cdots \geq 0\}$ and endow it with the ℓ_1 norm, i.e. for all s and $\mathbf{r}$ in $\mathcal{S}^{\downarrow}$, say that the distance between s and $\mathbf{r}$ is $\|\mathbf{s}-\mathbf{r}\| = \sum_{i\geq 1} |s_i - r_i|$. Moreover, set $\mathbf{0} := (0,0,\dots)$, $\mathbf{1} := (1,0,0,\dots)$ and $\mathcal{S}^{\downarrow}_{\leq 1} := \{\mathbf{s} \in \mathcal{S}^{\downarrow} \,:\, \|\|s \leq 1\}$.

A *self-similar fragmentation process* is an $\mathcal{S}^{\downarrow}_{\leq 1}$-valued Markovian process $(\mathbf{X}(t); t\geq 0)$ which is continuous in probability, and satisfies $\mathbf{X}(0) = \mathbf{1}$ as well as the following so-called *fragmentation property*. There exists $\alpha \in \mathbb{R}$ such that for all $t_0 \geq 0$, conditionally to $\mathbf{X}(t_0) = \mathbf{s}$, $\big(\mathbf{X}(t_0+t), t\geq 0\big)$ has the same distribution as

$$\Big(\big(s_i\,\mathbf{X}^{(i)}(s_i^{\alpha} t),\, i\geq 1\big)^{\downarrow};\, t\geq 0\Big)$$

where $(\mathbf{X}^{(i)})_{i\geq 1}$ are i.i.d. copies of $\mathbf{X}$. The constant α is called the *self-similarity index* of the process $\mathbf{X}$.

These processes can be seen as the evolution of the fragmentation of an object of mass 1 into smaller objects which will each, in turn, split themselves apart independently from one another, at a rate proportional to their mass to the power α.

It was shown in [8, 9] that the distribution of a self-similar fragmentation process is characterised by a 3-tuple (α, c, ν) where α is the aforementioned self-similarity index, $c \geq 0$ is a so-called erosion coefficient which accounts for a continuous decay in the mass of each particle and ν is a *dislocation measure* on $\mathcal{S}^{\downarrow}_{\leq 1}$, i.e. a σ-finite measure such that $\int (1-s_1)\,\nu(\mathrm{ds}) < \infty$ and $\nu(\{\mathbf{1}\}) = 0$. Informally, at any given time, each particle with mass say x will, independently from the other particles, split into smaller fragments of respective masses $xs_1, xs_2, \dots$ at rate $x^{\alpha}\nu(\mathrm{ds})$.

We will be interested in fragmentation processes with negative self-similarity index $-\gamma < 0$ with no erosion, i.e. with $c = 0$. Furthermore, we will require the dislocation measure ν to be non-trivial, i.e. $\nu(\mathcal{S}^{\downarrow}_{\leq 1}) > 0$, and *conservative*, that is to satisfy $\nu(\|\|s < 1) = 0$. Therefore, the fragmentation processes we will consider will be characterised by a *fragmentation pair* (γ, ν) and we will refer to them as (γ,ν)-fragmentation processes.

Under these assumptions, each particle will split into smaller ones which will in turn break down faster, thus speeding up the global fragmentation rate. Let $\mathbf{X}$ be a (γ,ν)-fragmentation process and set $\tau_{\mathbf{0}} := \inf\{t \geq 0 : \mathbf{X}(t) = \mathbf{0}\}$ the first time at which all the mass has been turned to dust. It was shown in [10, Proposition 2] that $\tau_{\mathbf{0}}$ is a.s. finite and in [25, Section 5.3] that it has exponential moments, i.e. that there exists $a > 0$ such that $\mathbb{E}\big[\exp(a\tau_{\mathbf{0}})\big] < \infty$.

Furthermore, a $\mathbb{T}_c$-valued random variable that encodes the genealogy of the fragmentation of the initial object was defined in [29]. This random $\mathbb{R}$-tree $(\mathcal{T}, \mathrm{d}, \rho, \mu)$ is such that $\mu(\mathcal{T}) = 1$ and if for all $t \geq 0$, $\{\mathcal{T}_i(t) : i \geq 1\}$ is the (possibly empty) set of the closures of the connected components of $\mathcal{T} \setminus (\mathcal{T}|_t)$, then

$$\Big(\big(\mu[\mathcal{T}_i(t)]\,;\, i\geq 1\big)^{\downarrow}\,;\; t\geq 0\Big)$$

is a (γ,ν)-fragmentation process. We will denote the distribution of $(\mathcal{T}, \mathrm{d}, \rho, \mu)$ by $\mathscr{T}_{\gamma,\nu}$.

Remark 3.12.
- More general self-similar fragmentation trees, where both the assumptions "$c = 0$" and "ν is conservative" are dropped, were defined and studied in [49].
- Let $\mathcal{T}$ be a (γ,ν)-self-similar fragmentation tree and $m > 0$. The tree $(m^{\gamma}\mathcal{T}, m\,\mu_{\mathcal{T}})$ encodes the genealogy of a (γ,ν)-self-similar fragmentation process started from a single object with mass m.

Classical examples It was observed in [9] that the Brownian tree, which was introduced in [5], may be described as a self-similar fragmentation tree with parameters $(1/2, \nu_B)$ where ν_B is called the *Brownian dislocation measure* and is defined for all measurable $f : \mathcal{S}^{\downarrow}_{\leq 1} \to \mathbb{R}_+$ by

$$\int f \,\mathrm{d}\nu_B = \int_{1/2}^{1} \left(\frac{2}{\pi\, x^3\,(1-x)^3}\right)^{1/2} f(x, 1-x, 0, 0, \dots)\,\mathrm{d}x.$$

Another important example of fragmentation trees is the family of α-stable trees from [23], where α belongs to $(1,2)$. Indeed, a result from [43] states that the α-stable tree is a $(1-1/\alpha, \nu_\alpha)$-self-similar fragmentation tree with ν_α defined as follows: let $(\Sigma_t; t \geq 0)$ be a $1/\alpha$-stable subordinator with Laplace exponent $\lambda \mapsto -\log \mathbb{E}[\exp(-\lambda\Sigma_1)] = \lambda^{1/\alpha}$ and Lévy measure $\Pi_{1/\alpha}(\mathrm{d}t) := [\alpha\,\Gamma(1-1/\alpha)]^{-1}\, t^{-1-1/\alpha}\, \mathbb{1}_{t>0}\,\mathrm{d}t$, denote the decreasing rearrangement of its jumps on $[0,1]$ by Δ and for all measurable $f : \mathcal{S}^{\downarrow} \to \mathbb{R}_+$, let

$$\int_{\mathcal{S}^{\downarrow}} f\,\mathrm{d}\nu_\alpha = \frac{\Gamma(1-1/\alpha)}{k_\alpha}\mathbb{E}\Big[\Sigma_1\, f\big(\Delta/\Sigma_1\big)\Big]$$

where $k_\alpha := \Gamma(2-\alpha)/[\alpha\,(\alpha-1)]$. Observe that the random point measure $\sum_{i\geq 1}\delta_{\Delta_i}$ on $(0,\infty)$ with atoms $(\Delta_i, i\geq 1)$ is a Poisson Point Process with intensity measure $\Pi_{1/\alpha}$.

Scaling limits of Markov branching trees Self-similar fragmentation trees bear a close relationship with Markov branching trees. Let $\iota : \mathcal{P}_{<\infty} \to \mathcal{S}^{\downarrow}_1$ be such that if $\lambda = (\lambda_1, \dots, \lambda_p)$ is in $\mathcal{P}_n$, then $\iota(\lambda) := (\lambda_1/n, \dots, \lambda_p/n, 0, 0, \dots)$.

Theorem 3.13 ([30], Theorems 5 and 6)**.**
- *Let $(q_n)_{n\in\mathcal{N}}$ be the sequence of first-split distributions of a Markov branching family $\mathrm{MB}^{\mathcal{L},q}$ and for all adequate $n\geq 1$, set $\bar{q}_n := q_n \circ \iota^{-1}$. Suppose there exists a fragmentation pair (γ,ν) and a slowly varying function ℓ such that, for the weak convergence of finite measures on $\mathcal{S}^{\downarrow}$,*

$$n^\gamma \ell(n)\,(1-s_1)\,\bar{q}_n(\mathrm{d}s) \xrightarrow[n\to\infty]{} (1-s_1)\,\nu(\mathrm{d}s).$$

  For all $n\in\mathcal{N}$, let T_n have distribution $\mathrm{MB}^{\mathcal{L},q}_n$ and set $\mu_n := \sum_{u\in\mathcal{L}(T_n)}\delta_u$ the counting measure on the leaves of T_n.
- *Let $(q_{n-1})_{n\in\mathcal{N}}$ be the sequence associated to a Markov branching family MB^q. Assume that there exists a fragmentation pair (γ,ν) and a slowly varying function ℓ with either $\gamma<1$ or $\gamma=1$ and $\ell(n)\to 0$ such that $n^\gamma\ell(n)\,(1-s_1)\,\bar{q}_n(\mathrm{d}s) \Rightarrow (1-s_1)\,\nu(\mathrm{d}s)$. For each $n\in\mathcal{N}$, let T_n be a MB^q_n tree and endow it with its counting measure μ_n.*

Under either set of assumptions, with respect to the GHP topology on $\mathbb{T}_c$,

$$\left(\frac{1}{n^\gamma\ell(n)}T_n, \frac{1}{n}\mu_{T_n}\right) \xrightarrow[n\to\infty]{} \mathscr{T}_{\gamma,\nu} \quad \textit{in distribution.}$$

The following useful result on the heights of Markov branching also holds.

Lemma 3.14. *Suppose that $(q_n)_{n\in\mathcal{N}}$ satisfies the assumptions of Theorem 3.13 with respect to a given fragmentation pair (γ,ν) and a slowly varying function ℓ. Then for any $p>0$, there is a finite constant h_p such that*

$$\sup_{n\in\mathcal{N}} \mathbb{E}\left[\left(\frac{|T_n|}{n^\gamma\ell(n)}\right)^p\right] \leq h_p \quad \textit{and} \quad \mathbb{E}\big[||T^p\big] \leq h_p$$

where $\mathcal{T}$ is a (γ,ν)-fragmentation tree and, as in Theorem 3.13, T_n has distribution either MB^q_n or $\mathrm{MB}^{\mathcal{L},q}_n$.

Proof. See [25, Section 5.3] for the continuous setting and [30, Lemma 33] plus [30, Section 4.5] for the discrete one. □

Concatenation of fragmentation trees Fix a fragmentation pair (γ,ν) and let $(\mathcal{T}_i)_{i\geq 1}$ be a sequence of i.i.d. (γ,ν)-fragmentation trees. For all $i\geq 1$, call μ_i the measure of $\mathcal{T}_i$. Fix s in $\mathcal{S}^{\downarrow}$ and set $(\mathcal{T}_{\langle \mathbf{s}\rangle},\mu_{\langle \mathbf{s}\rangle}) := \big\langle (s_i^{\gamma}\mathcal{T}_i, s_i\mu_i)\,;\, i\geq 1\big\rangle$.

Lemma 3.15. *With these notations,* $(\mathcal{T}_{\langle \mathbf{s}\rangle},\mu_{\langle \mathbf{s}\rangle})$ *a.s. belongs to* $\mathbb{T}_c$.

Proof. Clearly $\mathcal{T}_{\langle \mathbf{s}\rangle}$ is an $\mathbb{R}$-tree and its total mass is $\mu_{\langle \mathbf{s}\rangle}(\mathcal{T}_{\langle \mathbf{s}\rangle}) = \sum_{i\geq 1} s_i\mu_i(\mathcal{T}_i) = \|\|\|s$ which is finite. It only remains to show that it is compact or, in light of Lemma 3.6, that $s_i^{\gamma}\,|\mathcal{T}_i|$ a.s. converges to 0 as i grows to infinity. Since s is summable, for any positive ε,

$$\textstyle\sum_{i\geq 1}\mathbb{P}\big[s_i^{\gamma}|\mathcal{T}_i|>\varepsilon\big] \leq \sum_{i\geq 1}\dfrac{s_i}{\varepsilon^{1/\gamma}}\mathbb{E}\big[|\mathcal{T}_1|^{1/\gamma}\big] \leq \dfrac{1}{\varepsilon^{1/\gamma}}\mathbb{E}\big[|\mathcal{T}_1|^{1/\gamma}\big]\,\|\|\|s<\infty$$

where we have used Markov's inequality and the fact that $|\mathcal{T}_i|^{1/\gamma}\in L^1$ (see Lemma 3.14). Borel-Cantelli's lemma then allows us to deduce that $s_i^{\gamma}|\mathcal{T}_i|\to 0$ a.s. as $i\to\infty$. □

Lemma 3.16. *For all fixed* s *in* $\mathcal{S}^{\downarrow}$, $\mathbb{E}\big[\mathrm{d}_{\mathrm{GHP}}(\mathcal{T}_{\langle \mathbf{s}\rangle},\mathcal{T}_{\langle \mathbf{r}\rangle})\big]$ *converges to* 0 *as* $\mathbf{r}\to\mathbf{s}$.

Proof. For all $n\geq 0$, in light of Lemmas 3.3 and 3.7,

$$\mathrm{d}_{\mathrm{GHP}}\big(\mathcal{T}_{\langle \mathbf{s}\rangle},\mathcal{T}_{\langle \mathbf{r}\rangle}\big) \leq \sum_{i=1}^{n}\Big[\big(|s_i^{\gamma}-r_i^{\gamma}|\,|\mathcal{T}_i|\big)\vee|s_i-r_i|\Big]+\sum_{i>n}\big(s_i+r_i\big)+\sup_{i>n}\Big(s_i^{\gamma}|\mathcal{T}_i|\Big)+\sup_{i>n}\Big(r_i^{\gamma}|\mathcal{T}_i|\Big).$$

If $\gamma\leq 1$, $t\mapsto t^{\gamma}$ is concave, hence Jensen's inequality gives

$$\mathbb{E}\Big[\sup_{i>n}\Big(s_i^{\gamma}|\mathcal{T}_i|\Big)\Big] = \mathbb{E}\Big[\Big(\sup_{i>n}\, s_i|\mathcal{T}_i|^{1/\gamma}\Big)^{\gamma}\Big] \leq \Big(\mathbb{E}\Big[\sup_{i>n}\, s_i|\mathcal{T}_i|^{1/\gamma}\Big]\Big)^{\gamma} \leq \mathbb{E}\big[|\mathcal{T}_1|^{1/\gamma}\big]^{\gamma}\Big(\textstyle\sum_{i>n}s_i\Big)^{\gamma},$$

otherwise, if $\gamma>1$, since (s_i) is non-increasing, for all $i>n$, $s_i^{\gamma}\leq s_{n+1}^{\gamma-1}s_i$ which implies

$$\mathbb{E}\Big[\sup_{i>n}\Big(s_i^{\gamma}|\mathcal{T}_i|\Big)\Big] \leq s_{n+1}^{\gamma-1}\,\mathbb{E}\Big[\sup_{i>n}\Big(s_i|\mathcal{T}_i|\Big)\Big] \leq \mathbb{E}\big[|\mathcal{T}_1|\big]\,s_{n+1}^{\gamma-1}\sum_{i>n}s_i \leq \mathbb{E}\big[|\mathcal{T}_1|\big]\,\Big(\textstyle\sum_{i>n}s_i\Big)^{\gamma}.$$

Consequently, there is a non-negative constant C such that for all integer n and s in $\mathcal{S}^{\downarrow}$, $\mathbb{E}\big[\sup_{i>n}s_i^{\gamma}|\mathcal{T}_i|\big]\leq C\big[\sum_{i>n}s_i\big]^{\gamma}$. Hence, for all s and $\mathbf{r}$ in $\mathcal{S}^{\downarrow}$ and any $n\geq 1$

$$\mathbb{E}\Big[\mathrm{d}_{\mathrm{GHP}}(\mathcal{T}_{\langle \mathbf{s}\rangle},\mathcal{T}_{\langle \mathbf{r}\rangle})\Big] \leq \|\mathbf{s}-\mathbf{r}\|+\mathbb{E}\big[|\mathcal{T}_1|\big]\sum_{i=1}^{n}|s_i^{\gamma}-r_i^{\gamma}|+\sum_{i>n}(s_i+r_i)+C\Big[\Big(\sum_{i>n}s_i\Big)^{\gamma}+\Big(\sum_{i>n}r_i\Big)^{\gamma}\Big].$$

As a result,

$$\limsup_{\mathbf{r}\to\mathbf{s}}\ \mathbb{E}\Big[\mathrm{d}_{\mathrm{GHP}}(\mathcal{T}_{\langle \mathbf{s}\rangle},\mathcal{T}_{\langle \mathbf{r}\rangle})\Big] \leq \inf_{n\geq 1}\ 2\textstyle\sum_{i>n}s_i+2\,C\Big(\sum_{i>n}s_i\Big)^{\gamma}=0.$$ □

3.2.2 Fragmentation trees with immigration

We say that a non-negative Borel measure I on $\mathcal{S}^{\downarrow}$ is an *immigration measure* if it satisfies $\int_{\mathcal{S}^{\downarrow}}(1\wedge\|\|\|s)\,I(\mathrm{d}s)<\infty$ and $I(\{\mathbf{0}\})=0$.

Fix an immigration measure I such that $I(\mathcal{S}^{\downarrow})>0$ and let (γ,ν) be a fragmentation pair. Let $\Sigma=\sum_{n\geq 1}\delta_{(u_n,\mathbf{s}_n)}$ be a Poisson point process on $\mathbb{R}_+\times\mathcal{S}^{\downarrow}$ with intensity $\mathrm{d}u\otimes I(\mathrm{ds})$ independent of a family $(\mathbf{X}^{(n,k)},\,n\geq 1,k\geq 1)$ of i.i.d. (γ,ν)-fragmentation processes. Define the $\mathcal{S}^{\downarrow}$-valued process $\mathbf{X}$ as follows:

$$\mathbf{X}=\big(\mathbf{X}(t),\,t\geq 0\big) := \Bigg(\Bigg(s_{n,k}\mathbf{X}^{(n,k)}\big[s_{n,k}^{-\gamma}(t-u_n)\big]\,;\,n\geq 1: u_n\leq t,\,k\geq 1\Bigg)^{\downarrow}\,;\,t\geq 0\Bigg).$$

We call $\mathbf{X}$ a fragmentation process with immigration with parameters (γ,ν,I). It describes the evolution of the masses of a cluster of independently fragmenting objects, where new objects of sizes s_n appear, or immigrate, at time u_n. These processes were introduced in [26].

Like pure fragmentation processes, the genealogy of these immigrations and fragmentations can be encoded as an infinite weighted $\mathbb{R}$-tree (see [27]), say $(\mathcal{T}^{(I)}, \mathrm{d}, \rho, \mu)$, such that if for all $t \geq 0$, we denote the set of the closures of the bounded connected components of $\mathcal{T}^{(I)} \setminus (\mathcal{T}^{(I)}|_t)$ by $\{\mathcal{T}_i(t) : i \geq 1\}$, then

$$\Big(\big(\mu[\mathcal{T}_i(t)]\,;\, i \geq 1\big)^{\downarrow}\,;\ t \geq 0\Big)$$

is a (γ, ν, I)-fragmentation process with immigration. Let $\mathscr{T}^{I}_{\gamma,\nu}$ be the distribution of $(\mathcal{T}^{(I)}, \mathrm{d}, \rho, \mu)$.

Point process construction The construction of (γ,ν)-fragmentation trees with immigration I described in [27] can be expressed using Poisson point processes, concatenated (γ,ν)-fragmentation trees and the continuum grafting function $\mathbf{G}$ from the end of Section 3.1.1. Let $\Sigma = \sum_{i\geq 1} \delta_{(u_i, \mathbf{s}_i)}$ be a Poisson point process on $\mathbb{R}_+ \times \mathcal{S}^{\downarrow}$ with intensity $\mathrm{d}u \otimes I(\mathrm{d}\mathbf{s})$ and $(\mathcal{T}_{i,j}, \mu_{i,j})_{i,j\geq 1}$ be i.i.d. (γ,ν)-fragmentation trees independent of Σ. For all $i \geq 1$, set

$$\mathcal{T}_i := \big\langle (s_{i,j}^{\gamma}\mathcal{T}_{i,j}, s_{i,j}\mu_{i,j}); j \geq 1 \big\rangle,$$

the concatenation of $(\mathcal{T}_{i,j}; j \geq 1)$ with respective masses $s_{i,j}$. Define $\mathcal{T}^{(I)}$ as the tree obtained by grafting $\mathcal{T}_i$ at height u_i on an infinite branch for each $i \geq 1$, i.e.

$$\mathcal{T}^{(I)} := \mathbf{G}\Big(\big\{(u_i, \mathcal{T}_i)\ :\ i \geq 1\big\}\Big).$$

The random tree $\mathcal{T}^{(I)}$ has distribution $\mathscr{T}^{I}_{\gamma,\nu}$.

Observe that for all $K \geq 0$, we can write the total mass grafted on the infinite branch at height less than K as an integral against the point-process Σ:

$$\textstyle\sum_{i\geq 1} \mathbb{1}_{u_i \leq K} \mu_{\mathcal{T}_i}(\mathcal{T}_i) = \sum_{i\geq 1} \mathbb{1}_{u_i\leq K} \|\mathbf{s}_i\| = \int \mathbb{1}_{u\leq K}\, \|\|s\, \Sigma(\mathrm{d}u, \mathrm{d}\mathbf{s}).$$

Since $\int 1 \wedge \big(\mathbb{1}_{u\leq K}\, \|\|s\big)\, \mathrm{d}u\, I(\mathrm{d}\mathbf{s}) = K \int (1 \wedge \|\|s) I(\mathrm{d}\mathbf{s}) < \infty$, we may use Campbell's theorem (see [38, Section 3.2]) and claim that $\int \mathbb{1}_{u\leq K}\, \|\|s\, \Sigma(\mathrm{d}u, \mathrm{d}\mathbf{s}) < \infty$ a.s.. The second condition of Lemma 3.10 is thus met. Moreover, for all $i \geq 1$,

$$\mathbb{E}\Big[|\mathcal{T}_i|^{1/\gamma}\big|\Sigma\Big] = \mathbb{E}\Big[\sup\nolimits_{j\geq 1} s_{i,j} |\mathcal{T}_{i,j}|^{1/\gamma}\big|\Sigma\Big] \leq \textstyle\sum_{j\geq 1} s_{i,j} \mathbb{E}\Big[|\mathcal{T}_{1,1}|^{1/\gamma}\Big] = \mathbb{E}\Big[|\mathcal{T}_{1,1}|^{1/\gamma}\Big]\, \|\mathbf{s}_i\|$$

where we have used the fact that $(\mathcal{T}_{i,j})_{i,j}$ is an i.i.d. family independent of Σ. Markov's inequality therefore implies that

$$\sum_{i\geq 1} \mathbb{1}_{u_i\leq K} \mathbb{P}\big[|\mathcal{T}_i| \geq \varepsilon |\Sigma\big] \leq \sum_{i\geq 1} \mathbb{1}_{u_i\leq K} \varepsilon^{-1/\gamma}\, \mathbb{E}\big[|\mathcal{T}_i|^{1/\gamma}|\Sigma\big] \leq \frac{\mathbb{E}\big[|\mathcal{T}_{1,1}|^{1/\gamma}\big]}{\varepsilon^{1/\gamma}} \sum_{i\geq 1} \mathbb{1}_{u_i\leq K} \|\mathbf{s}_i\|$$

which is, according to Campbell's formula, a.s. finite. Consequently, using Borel-Cantelli's lemma, we deduce that conditionally on Σ, with probability one, there are finitely many indices $i \geq 1$ such that $u_i \leq K$ and $\mathcal{T}_i$ is higher than ε. It follows from Lemma 3.10 that $\mathcal{T}^{(I)}$ is a.s. $\mathbb{T}$-valued.

Self-similar immigration measures We will say that an immigration measure I with $I(\mathcal{S}^{\downarrow}) > 0$ is *self-similar* with positive index γ (or simply, γ-self-similar) if for all $c > 0$ and measurable $F : \mathcal{S}^{\downarrow} \to \mathbb{R}_+$, $c \int F(\mathbf{s})\, I(\mathrm{d}\mathbf{s}) = \int F(c^{1/\gamma}\mathbf{s})\, I(\mathrm{d}\mathbf{s})$.

Proposition 3.17. *An immigration measure I is γ-self-similar iff $\gamma \in (0,1)$ and there exists a positive constant K as well as an $\mathcal{S}_1^{\downarrow}$-valued random variable X such that for all measurable $F : \mathcal{S}^{\downarrow} \to \mathbb{R}_+$*

$$\int F \,\mathrm{d}I = \int_0^{\infty} \frac{K}{t^{1+\gamma}} \mathbb{E}\big[F(t\,X)\big] \,\mathrm{d}t.$$

Proof. Clearly, if X is an $\mathcal{S}_1^{\downarrow}$-valued random variable, $K > 0$ and $\gamma \in (0,1)$, the measure I on $\mathcal{S}^{\downarrow}$ defined for all measurable $F : \mathcal{S}^{\downarrow} \to \mathbb{R}_+$ by

$$\textstyle\int F \,\mathrm{d}I = K \int_0^{\infty} t^{-1-\gamma} \mathbb{E}\big[F(t\,X)\big] \,\mathrm{d}t$$

is an immigration measure. Moreover, for all $c > 0$, a simple change of variable gives $\int F(c^{1/\gamma}\mathbf{s})\, I(\mathrm{d}\mathbf{s}) = c \int F \,\mathrm{d}I$ which means that I is indeed γ-self-similar.

Conversely, suppose I is a γ-self-similar immigration. Define σ, the probability measure on $\mathcal{S}_1^{\downarrow}$ such that for all measurable $f : \mathcal{S}_1^{\downarrow} \to \mathbb{R}_+$

$$\textstyle\int_{\mathcal{S}_1^{\downarrow}} f(\mathbf{s})\, \sigma(\mathrm{d}\mathbf{s}) := Z^{-1} \int_{\mathcal{S}^{\downarrow}} f(\mathbf{s}/\|\|s)\, \mathbb{1}_{\|\|s \geq 1}\, I(\mathrm{d}\mathbf{s}),$$

where $Z := I\big(\|\cdot\| \geq 1\big)$, and let X be a σ-distributed random variable. Now, for any measurable $g : \mathcal{S}_1^{\downarrow} \to \mathbb{R}_+$ and $t > 0$, because I is self-similar, we get that

$$\textstyle\int_{\mathcal{S}^{\downarrow}} g(\mathbf{s}/\|\|s)\, \mathbb{1}_{\|\|s \geq t}\, I(\mathrm{d}\mathbf{s}) = t^{-\gamma}\, Z\, \mathbb{E}\big[g(X)\big] = \gamma\, Z \int_t^{\infty} u^{-1-\gamma}\, \mathbb{E}\big[g(X)\big] \,\mathrm{d}u.$$

Since this identity holds for any $t > 0$ and measurable $g : \mathcal{S}_1^{\downarrow} \to \mathbb{R}_+$ and because $I(\{\mathbf{0}\}) = 0$, it follows that I may be written in the desired way. Finally, because I is an immigration measure, it must integrate $\mathbf{s} \mapsto 1 \wedge \|\|s$, which implies that γ belongs to $(0,1)$. ◻

The point process construction of fragmentation trees with immigration may be used to prove this next proposition.

Proposition 3.18. *Suppose I is a γ-self-similar immigration measure and let ν be a dislocation measure. If $(\mathcal{T}, \mu)$ denotes a (γ, ν, I)-fragmentation tree with immigration, then for any positive m,*

- *$(m^{\gamma}\mathcal{T}, m\,\mu)$ has the same distribution as $(\mathcal{T}, \mu)$,*
- *$(\mathcal{T}, c\mu)$ and $(c^{-\gamma}\mathcal{T}, \mu)$ are $(\gamma, c^{\gamma}\nu, c^{\gamma} I)$-fragmentation trees with immigration.*

Relationship to compact fragmentation trees Let (γ, ν) be a fragmentation pair and I an immigration measure with $I(\mathcal{S}^{\downarrow}) > 0$. Theorem 17 in [27] states that under suitable conditions, if $(\mathcal{T}, \mu_{\mathcal{T}})$ denotes a (γ, ν)-self-similar fragmentation tree, then $(m^{\gamma}\mathcal{T}, m\mu_{\mathcal{T}})$ converges to $\mathscr{T}_{\gamma,\nu}^{I}$ in distribution as $m \to \infty$ with respect to the extended GHP topology.

For instance, Theorem 11 (iii) in [5], states that if $(\mathcal{T}, \mu_{\mathcal{T}})$ is a standard Brownian tree then when $m \to \infty$, $(m^{1/2}\mathcal{T}, m\,\mu_{\mathcal{T}})$ converges in distribution to Aldous' "self-similar CRT". This result was reformulated in terms of fragmentation trees in [27, Section 1.2]: $(m^{1/2}\mathcal{T}, m\,\mu_{\mathcal{T}})$ converges in distribution as $m \to \infty$ to a $(1/2, \nu_B, I_B)$-fragmentation tree with immigration, where ν_B is the Brownian dislocation measure (see Section 3.2.1) and the *Brownian immigration measure* I_B is defined for all measurable $f : \mathcal{S}^{\downarrow} \to \mathbb{R}_+$ by

$$\int F \,\mathrm{d}I_B := \left(\frac{2}{\pi}\right)^{1/2} \int_{[0,\infty)} \frac{f(x,0,0,\dots)}{x^{3/2}} \,\mathrm{d}x.$$

We will call a $(1/2, \nu_B, I_B)$-fragmentation tree with immigration a *Brownian tree with immigration*. As mentioned in the Introduction, this tree will appear in many of our applications.

Set $\alpha \in (1,2)$ and recall the notations used to define ν_α in Section 3.2.1, in particular, that Δ denotes the decreasing rearrangement of the jumps on $[0,1]$ of an $1/\alpha$-stable subordinator with Laplace exponent $\lambda \mapsto -\log \mathbb{E}[\exp(-\lambda \Sigma_1)] = \lambda^{1/\alpha}$ and that $k_\alpha = \Gamma(2-\alpha)/[\alpha\,(\alpha-1)]$. Let $I^{(\alpha)}$ be the immigration measure defined for all measurable $F : \mathcal{S}^{\downarrow} \to \mathbb{R}_+$ by

$$\int_{\mathcal{S}^{\downarrow}} F \,\mathrm{d}I^{(\alpha)} = \frac{1}{k_\alpha} \int_0^\infty \frac{\mathbb{E}\big[F(t^\alpha\,\Delta)\big]}{t^\alpha} \mathrm{d}t.$$

In [27, Section 5.1], it was observed that if $(\mathcal{T}, \mu_{\mathcal{T}})$ is an α-stable tree, then, as $m \to \infty$, $(m^{1-1/\alpha}\mathcal{T}, m\,\mu_{\mathcal{T}})$ converges in distribution to a $(1-1/\alpha, \nu_\alpha, I^{(\alpha)})$-fragmentation tree with immigration. These trees coincide with the α-stable Lévy trees with immigration introduced in [22, Section 1.2].

3.3 Convergence of point processes

With the notations used in Section 3.2.2, let $\Pi := \sum_{i\geq 1} \delta_{(u_i, \mathbf{s}_i, \mathcal{T}_i)}$. It is a Poisson point process on $\mathbb{R}_+ \times \mathcal{S}^{\downarrow} \times \mathbb{T}_c$ with intensity $\mathrm{d}u \otimes \mathscr{I}(\mathrm{d}s, \mathrm{d}\tau)$ where the measure $\mathscr{I}$ on $\mathcal{S}^{\downarrow} \times \mathbb{T}_c$ is defined as follows: let $(\tau_i, \mu_i)_{i\geq 1}$ be a sequence of i.i.d. (γ, ν)-fragmentation trees and for any s in $\mathcal{S}^{\downarrow}$, similarly to Section 3.2.1, set $\tau_{\langle \mathbf{s}\rangle} := \big\langle (s_i^\gamma \tau_i, s_i \mu_i)\,;\, i \geq 1\big\rangle$ and for all $G : \mathcal{S}^{\downarrow} \times \mathbb{T}_c \to \mathbb{R}_+$, let $\int G \,\mathrm{d}\mathscr{I} := \int \mathbb{E}[G(\mathbf{s}, \tau_{\langle \mathbf{s}\rangle})]\, I(\mathrm{d}s)$.

Moreover, recall the construction of Markov branching trees with a unique infinite spine (see Remark 2.8). If q_∞ is such that $q_\infty(m_\infty = 1) = 1$, then a tree T with distribution $\mathrm{MB}_\infty^{q,q_\infty}$ can be built in the following way: consider the infinite branch and for all $n \geq 0$, graft a tree T_n at height n (where the sequence $(T_n)_{n\geq 0}$ is i.i.d.), such that $\Lambda_n := \Lambda(T_n)$ has distribution $q_* = q_\infty(\infty, \cdot\,)$ and conditionally on $\Lambda_n = \lambda$ in $\mathcal{P}_{<\infty}$, T_n has distribution MB_λ^q. As a result, T is characterised by the point process $\sum_{n\geq 0} \delta_{(n, \Lambda_n, T_n)}$ (or simply by $\sum_{n\geq 0} \delta_{(n, T_n)}$).

Therefore, when considering scaling limits of such trees, it seems natural to take a step back and instead consider the convergence of the underlying point processes on $\mathbb{R}_+ \times \mathcal{S}^{\downarrow} \times \mathbb{T}_c$. We will follow the spirit of [27, Section 2.1.2] and introduce a topology on the set of such point measures adequate for our forthcoming purposes.

Let $\mathscr{R}$ be the set of integer-valued Radon measures on $\mathbb{R}_+ \times \mathcal{S}^{\downarrow} \times \mathbb{T}_c$ which integrate the function $(u, \mathrm{s}, \tau) \longmapsto \mathbb{1}_{u \leq K} \|\|\|s$ for all $K \geq 0$ and are such that $\mu(\mathbb{R}_+ \times \{\mathbf{0}\} \times \mathbb{T}_c) = 0$.

Remark 3.19. Recall that as an immigration measure, I integrates the function s $\in \mathcal{S}^{\downarrow} \to 1 \wedge \|\|\|s$. Campbell's theorem (see [38, Section 3.2]) therefore ensures that Π, the Poisson point process associated to a $\mathcal{T}^I_{\gamma,\nu}$ tree, a.s. belong to $\mathscr{R}$.

Let $\mathscr{F}$ be the set of continuous functions $F : \mathbb{R}_+ \times \mathcal{S}^{\downarrow} \times \mathbb{T}_c \longrightarrow \mathbb{R}_+$ such that there is $K \geq 0$ satisfying $F(u, \mathrm{s}, \tau) \leq \mathbb{1}_{u\leq K} \|\|\|s$ for all (u, s, τ). If ζ is a random element of $\mathscr{R}$, we define its *Laplace transform* as the function $L_\zeta : \mathscr{F} \to \mathbb{R}_+$, defined by $L_\zeta(F) := \mathbb{E}\big[\exp\big(-\int F \,\mathrm{d}\zeta\big)\big]$ for all F in $\mathscr{F}$.

If μ_n, $n \geq 1$ and μ are elements of $\mathscr{R}$, we will say that $\mu_n \to \mu$ iff for all $F \in \mathscr{F}$, $\int F \,\mathrm{d}\mu_n \to \int F \,\mathrm{d}\mu$. Appendix A7 of [35] ensures that when endowed with the topology induced by this convergence, $\mathscr{R}$ is a Polish space. Moreover, Theorems 4.2 and 4.9 of [35] give the following criterion for convergence in distribution of elements of $\mathscr{R}$.

Proposition 3.20 ([35])**.** *Let ξ_n, $n \geq 1$ and ξ be $\mathscr{R}$-valued random variables. Then ξ_n converges to ξ in distribution with respect to the topology on $\mathscr{R}$ iff for all $F \in \mathscr{F}$, $L_{\xi_n}(F) \to L_\xi(F)$.*

The following extension of the Portmanteau theorem to finite measures with any mass will be useful.

Lemma 3.21. *Set (M, d) a metric space and let μ_n, $n \geq 1$ and μ be finite Borel measures on M. Then μ_n converges weakly to μ iff for any bounded Lipschitz-continuous function $f : M \to \mathbb{R}$, $\int f \, \mathrm{d}\mu_n$ converges to $\int f \, \mathrm{d}\mu$ as n goes to infinity.*

Proof. Suppose $\int f \, \mathrm{d}\mu_n \to \int f \, \mathrm{d}\mu$ for all Lipschitz-continuous functions $f : M \to \mathbb{R}$. Observe that since constants are Lipschitz-continuous functions, our assumption implies that $\mu_n(M) \to \mu(M)$. Therefore, if $\mu(M) = 0$, we directly get $\mu_n \Rightarrow \mu$.

Otherwise, there exists n_0 such that $\mu_n(M) > 0$ for all $n \geq n_0$. For all such n, let $\tilde{\mu}_n := [\mu_n(M)]^{-1}\mu_n$ and $\tilde{\mu} := [\mu(M)]^{-1}\mu$ which are probability measures. It ensues from the usual Portmanteau theorem and our assumption that $\tilde{\mu}_n \Rightarrow \tilde{\mu}$. As a result, for any bounded continuous function f, as n goes to ∞, $\int f \, \mathrm{d}\mu_n = \mu_n(M) \int f \, \mathrm{d}\tilde{\mu}_n \to \mu(M) \int f \, \mathrm{d}\tilde{\mu} = \int f \, \mathrm{d}\mu$ which is to say that $\mu_n \Rightarrow \mu$. ☐

4 Scaling limits of infinite Markov-branching trees

In this section, we will state and prove our main result on scaling limits of infinite Markov branching trees as well as its corollary on their volume growth.

Let $\mathcal{N}$ be an infinite subset of $\mathbb{N}$ containing 1 and let $q = (q_{n-1})_{n \in \mathcal{N}}$ be a sequence of first-split distributions where for each n, q_{n-1} is supported by $\{\lambda \in \mathcal{P}_{n-1} : \lambda_i \in \mathcal{N}, i = 1, \dots, p(\lambda)\}$. Recall from Section 2.2.1 that the associated Markov branching family MB^q is well defined. Furthermore, let q_∞ be a probability measure on $\mathcal{P}_\infty$ supported by the set $\{(\infty, \lambda) : \lambda \in \mathcal{P}_{<\infty}, \lambda_i \in \mathcal{N}, i = 1, \dots, p(\lambda)\}$. In this way, the probability measure $\mathrm{MB}_\infty^{q,q_\infty}$ on $\mathtt{T}_\infty$ is also well defined and a.s. yields trees with a unique infinite spine. To lighten notations, let $q_* := q_\infty(\infty, \cdot\,)$ which is a probability measure on $\mathcal{P}_{<\infty}$.

In the remainder of this section, we will assume that:

- (S) There exist some $\gamma > 0$ and a dislocation measure ν on $\mathcal{S}^\downarrow$, such that $n^\gamma(1 - s_1)\,\bar{q}_n(\mathrm{d}\mathbf{s}) \Rightarrow (1 - s_1)\,\nu(\mathrm{d}\mathbf{s})$. In particular, Theorem 3.13 and Lemma 3.14 hold.
- (I) There exists an immigration measure I on $\mathcal{S}^\downarrow$ such that if Λ has distribution q_*, for any continuous $F : \mathcal{S}^\downarrow \to \mathbb{R}_+$ with $F(\mathbf{s}) \leq 1 \wedge \|\|\|s$, $R\,\mathbb{E}\big[F(\Lambda / R^{1/\gamma})\big] \to \int F \, \mathrm{d}I$ as $R \to \infty$.

Remark 4.1. Under Assumption (I), for any continuous $F : \mathcal{S}^\downarrow \to \mathbb{R}_+$ such that $F \leq 1 \wedge \| \cdot \|$ and positive c,

$$\begin{aligned} c \int F(\mathbf{s})\, I(\mathrm{d}\mathbf{s}) &= \lim_{R \to \infty} cR\, \mathbb{E}\big[F(\Lambda / R^{1/\gamma})\big] \\ &= \lim_{S \to \infty} S\, \mathbb{E}\big[F(c^{1/\gamma}\Lambda / S^{1/\gamma})\big] = \int F(c^{1/\gamma}\mathbf{s})\, I(\mathrm{d}\mathbf{s}) \end{aligned}$$

where we have taken $S = cR$. As a result, the immigration measure I is γ-self-similar, as defined in Section 3.2.2, and Proposition 3.18 therefore holds for (γ, ν, I)-fragmentation trees with immigration.

Theorem 4.2. *Let T be an infinite Markov branching tree with distribution $\mathrm{MB}_\infty^{q,q_\infty}$ endowed with its counting measure μ_T. Under Assumptions* (S) *and* (I)*, if $\gamma < 1$, with respect to the extended GHP topology,*

$$\left(\frac{T}{R}, \frac{\mu_T}{R^{1/\gamma}}\right) \xrightarrow[R \to \infty]{} \mathscr{T}_{\gamma,\nu}^I$$

in distribution, where $\mathscr{T}_{\gamma,\nu}^I$ denotes the distribution of a (γ, ν, I)-fragmentation tree with immigration.

Let $\mathbf{T}$ be a fixed element of $\mathbb{T}$. We define its *volume growth function* as $V_{\mathbf{T}} : \mathbb{R}_+ \to \mathbb{R}_+$, $R \mapsto \mu_T(T|_R)$. In other words, $V_{\mathbf{T}}(R)$ is the mass or volume of the closed ball $T|_R$. Once Theorem 4.2 is proved, we will be interested in the volume growth processes associated to these trees.

Proposition 4.3. *Suppose the assumptions of Theorem 4.2 are met. Let T be an infinite Markov branching tree with distribution $\mathrm{MB}_\infty^{q,q_\infty}$ and $(\mathcal{T}, \mu_{\mathcal{T}})$ be a (γ, ν, I)-fragmentation tree with immigration. Then, the volume growth function of $(T/R, \mu_T/R^{1/\gamma})$ converges in distribution to that of $(\mathcal{T}, \mu_{\mathcal{T}})$ with respect to the topology of uniform convergence on compacts of $\mathbb{R}_+$. In particular*

$$\frac{\mu_T(T|_R)}{R^{1/\gamma}} \xrightarrow[R\to\infty]{(\mathrm{d})} \mu_{\mathcal{T}}(\mathcal{T}|_1).$$

We may adapt the proofs of Theorem 4.2 and Proposition 4.3 to get the following theorem.

Theorem 4.4. *Let T be an infinite Markov branching tree with distribution $\mathrm{MB}_\infty^{\mathcal{L},q,q_\infty}$ and endow it with the counting measure μ_T on the set of its leaves. If $\gamma < 1$ and if Assumptions* (S) *and* (I) *hold for $(q_n)_n$ and q_∞ respectively, then the conclusions of both Theorem 4.2 and Proposition 4.3 hold.*

Remark 4.5. Instead of Assumption (I), we may assume that

(I′) There exists $\alpha < 1/\gamma$ and an immigration measure I on $\mathcal{S}^\downarrow$ such that if Λ is distributed according to q_*, $R\,\mathbb{E}\big[F(\Lambda/R^\alpha)\big] \to \int F\,\mathrm{d}I$ for any continuous $F : \mathcal{S}^\downarrow \to \mathbb{R}_+$ with $F(\mathbf{s}) \leq 1 \wedge \|\|\|s$.

If T has distribution $\mathrm{MB}_\infty^{q,q_\infty}$ and is endowed with its counting measure μ_T under (S) and (I′), we get that $(T/R, \mu_T/R^\alpha)$ converges in distribution to the infinite branch $\mathbb{R}_+$ endowed with the random measure $\mu = \sum_{i\geq 1} \|\mathbf{s}_i\|\,\delta_{u_i}$, where $\{(u_i, \mathbf{s}_i); i \geq 1\}$ are the atoms of a Poisson point process Σ on $\mathbb{R}_+ \times \mathcal{S}^\downarrow$ with intensity $\mathrm{d}u \otimes I(\mathrm{ds})$. The tree $(\mathbb{R}_+, \mu)$ encodes the genealogy of a pure immigration process. Furthermore, $\mu_T(T|_R)/R^\alpha$ converges in distribution to $\mu([0,1]) = \int_{[0,1]\times\mathcal{S}^\downarrow} \|\|\|s\,\Sigma(\mathrm{d}u, \mathrm{ds})$.

Similarly, if T is distributed according to $\mathrm{MB}_\infty^{\mathcal{L},q,q_\infty}$ and is endowed with the counting measure on its leaves, the same results hold under (S) and (I′).

To prove Theorem 4.2, we will first study the convergence of the underlying point processes in Section 4.1 which will give us more leeway to manipulate the corresponding trees and end the proof in Section 4.2. Section 4.3 will then focus on proving Proposition 4.3.

4.1 Convergence of the associated point processes

Since $(\mathbb{T}_c, \mathrm{d}_{\mathrm{GHP}})$ is Polish, in light of Assumption (S), Theorem 3.13 and Skorokhod's representation theorem, we can find an i.i.d. sequence $[(T_{i,n})_{n\in\mathcal{N}}, \mathcal{T}_i]_{i\geq 1}$, where for each $i \geq 1$, the family $(T_{i,n})_{n\in\mathcal{N}}, \mathcal{T}_i$ of random trees is such that:

- $T_{i,n}$ has distribution MB_n^q,
- $\mathcal{T}_i$ is a (γ, ν) self-similar fragmentation tree,
- $(T_{i,n}/n^\gamma, \mu_{T_{i,n}}/n) =: \overline{T}_{i,n}$ a.s. converges to $\mathcal{T}_i$ as $n \to \infty$.

For $\lambda = (\lambda_1, \ldots, \lambda_p) \in \mathcal{P}_{<\infty}$, let $T_{[\lambda]} := [\![T_{i,\lambda_i}\,; 1 \leq i \leq p]\!]$. For any $\mathrm{s} \in \mathcal{S}^\downarrow$, let $\mathcal{T}_{\langle \mathbf{s}\rangle} := \langle (s_i^\gamma \mathcal{T}_i, s_i \mu_{\mathcal{T}_i}); i \geq 1\rangle$ which is a compact $\mathbb{R}$-tree (see Lemma 3.15).

Finally, let Λ be a random finite partition with distribution q_* independent of the family $[(T_{i,n})_{n\in\mathcal{N}}, \mathcal{T}_i]_{i\geq 1}$, and for any $R \geq 1$, set $q^{(R)}$ as the distribution of $\Lambda/R^{1/\gamma}$. With these notations, Assumption (I) becomes: $R\,(1 \wedge \|\|\|s)\,q^{(R)}(\mathrm{ds}) \Rightarrow (1 \wedge \|\|\|s)\,I(\mathrm{ds})$ as finite measures on $\mathcal{S}^\downarrow$.

Lemma 4.6. *Let $K \subset \mathcal{S}^{\downarrow}$ be compact. Then $\sup_{\mathbf{s}\in K} \sum_{i>n} s_i \to 0$ as n goes to infinity.*

Proof. Assume the contrary, i.e. that there exists a sequence $(\mathbf{s}^{(n)})_{n\geq 1}$ in K and a positive constant c such that $\sum_{i>n} s_i^{(n)} > c$ for all $n \geq 1$. Since K is compact, we can find a subsequence $(\mathbf{s}^{(n_k)})_k$ and $\mathbf{s} \in K$ such that $\|\mathbf{s}^{(n_k)} - \mathbf{s}\| \to 0$ as $k \to \infty$. Consequently, $0 < c \leq \sum_{i>n_k} s_i^{(n_k)} \leq \sum_{i>n_k} s_i + \|\mathbf{s}^{(n_k)} - \mathbf{s}\| \to 0$ as $k \to \infty$, which is a contradiction. $\square$

Fix $G : \mathcal{S}^{\downarrow} \times \mathbb{T}_c \to \mathbb{R}_+$ a 1-Lipschitz function, i.e. such that for all $\mathbf{s}, \mathbf{s}'$ in $\mathcal{S}^{\downarrow}$ and τ, τ' in $\mathbb{T}_c$, $\left|G(\mathbf{s},\tau) - G(\mathbf{s}',\tau')\right| \leq \|\mathbf{s}-\mathbf{s}'\| + \mathrm{d}_{\mathrm{GHP}}(\tau,\tau')$. Further assume that $G(\mathbf{s},\,\cdot\,) \leq 1 \wedge \||\mathbf{s}\||$ for any $\mathbf{s} \in \mathcal{S}^{\downarrow}$. Finally, set $g : \mathcal{S}^{\downarrow} \to \mathbb{R}_+$ the function defined by $g(\mathbf{s}) := \mathbb{E}[G(\mathbf{s}, \mathcal{T}_{\langle \mathbf{s}\rangle})]$.

Lemma 4.7. *We have*

$$R\,\mathbb{E}\Big[G\Big(R^{-1/\gamma}\Lambda, (R^{-1}T_{[\Lambda]}, R^{-1/\gamma}\mu_{T_{[\Lambda]}})\Big)\Big] \xrightarrow[R\to\infty]{} \int_{\mathcal{S}^{\downarrow}} \mathbb{E}\big[G(\mathbf{s}, \mathcal{T}_{\langle\mathbf{s}\rangle})\big]\, I(\mathrm{d}\mathbf{s}).$$

Proof. Clearly, $g(\mathbf{s}) \leq 1 \wedge \||\mathbf{s}\||$. Moreover, for any $\mathbf{s}$ and $\mathbf{r}$ in $\mathcal{S}^{\downarrow}$,

$$\big|g(\mathbf{s}) - g(\mathbf{r})\big| \leq \mathbb{E}\Big[\big|G(\mathbf{s},\mathcal{T}_{\langle\mathbf{s}\rangle}) - G(\mathbf{r},\mathcal{T}_{\langle\mathbf{r}\rangle})\big|\Big] \leq \|\mathbf{s}-\mathbf{r}\| + \mathbb{E}\big[\mathrm{d}_{\mathrm{GHP}}\big(\mathcal{T}_{\langle\mathbf{s}\rangle},\mathcal{T}_{\langle\mathbf{r}\rangle}\big)\big] \xrightarrow[\mathbf{r}\to\mathbf{s}]{} 0$$

where we have used Lemma 3.16. Therefore, g is continuous and Assumption $(\mathtt{I})$ ensures that

$$R\,\mathbb{E}\Big[G\Big(R^{-1/\gamma}\Lambda, \big(R^{-1}\mathcal{T}_{\langle\Lambda\rangle}, R^{-1/\gamma}\mu_{\mathcal{T}_{\langle\Lambda\rangle}}\big)\Big)\Big] = R\,\mathbb{E}\big[g(R^{-1/\gamma}\Lambda)\big] \xrightarrow[R\to\infty]{} \int_{\mathcal{S}^{\downarrow}} g(\mathbf{s})\, I(\mathrm{d}\mathbf{s}).$$

Consequently, noticing that

$$\begin{aligned} &R\,\mathbb{E}\Big[\Big|G\Big(R^{-1/\gamma}\Lambda, (R^{-1}T_{[\Lambda]}, R^{-1/\gamma}\mu_{T_{[\Lambda]}})\Big) - G\Big(R^{-1/\gamma}\Lambda, (R^{-1}\mathcal{T}_{\langle\Lambda\rangle}, R^{-1/\gamma}\mu_{\mathcal{T}_{\langle\Lambda\rangle}})\Big)\Big|\Big] \\ &\quad\leq R\,\mathbb{E}\Big[\big(1\wedge R^{-1/\gamma}\|\Lambda\|\big)\ \wedge\ \mathrm{d}_{\mathrm{GHP}}\Big((R^{-1}T_{[\Lambda]}, R^{-1/\gamma}\mu_{T_{[\Lambda]}}), (R^{-1}\mathcal{T}_{\langle\Lambda\rangle}, R^{-1/\gamma}\mu_{\mathcal{T}_{\langle\Lambda\rangle}})\Big)\Big] =: \Delta_R, \end{aligned}$$

it will be sufficient to prove that $\Delta_R \to 0$ as $R \to \infty$.

For all $n \geq 1$, thanks to Lemma 3.7 and Remark 3.2 we get

$$\begin{aligned} &\mathrm{d}_{\mathrm{GHP}}\Big((R^{-1}T_{[\Lambda]}, R^{-1/\gamma}\mu_{T_{[\Lambda]}}), (R^{-1}\mathcal{T}_{\langle\Lambda\rangle}, R^{-1/\gamma}\mu_{\mathcal{T}_{\langle\Lambda\rangle}})\Big) \\ &\qquad\leq \sum_{i=1}^{n} \mathrm{d}_{\mathrm{GHP}}\Big((R^{-1}T_{i,\Lambda_i}, R^{-1/\gamma}\mu_{T_{i,\Lambda_i}}), (R^{-1}\Lambda_i^{\gamma}\mathcal{T}_i, R^{-1/\gamma}\Lambda_i\mu_{\mathcal{T}_i})\Big) \\ &\qquad\qquad + \sup_{i>n}\left(\frac{\Lambda_i^{\gamma}}{R}|\overline{T}_{i,\Lambda_i}|\right) + \sup_{i>n}\left(\frac{\Lambda_i^{\gamma}}{R}|\mathcal{T}_i|\right) + 2\sum_{i>n}\frac{\Lambda_i}{R^{1/\gamma}}, \end{aligned}$$

and for each $i \geq 1$, Lemma 3.3 gives

$$\mathrm{d}_{\mathrm{GHP}}\Big((R^{-1}T_{i,\Lambda_i}, R^{-1/\gamma}\mu_{T_{i,\Lambda_i}}), (R^{-1}\Lambda_i^{\gamma}\mathcal{T}_i, R^{-1/\gamma}\Lambda_i\mu_{\mathcal{T}_i})\Big) \leq \left(\frac{\Lambda_i^{\gamma}}{R} \vee \frac{\Lambda_i}{R^{1/\gamma}}\right)\mathrm{d}_{\mathrm{GHP}}\big(\overline{T}_{i,\Lambda_i}, \mathcal{T}_i\big).$$

Let $\varepsilon > 0$ be fixed. As a result of Assumption $(\mathtt{I})$, the sequence $R\,(1\wedge\||\mathbf{s}\||)\,q^{(R)}(\mathrm{d}\mathbf{s})$, $R \geq 1$ is tight and so there exists a compact subset K of $\mathcal{S}^{\downarrow}$ such that $\sup_{R\geq 1} R\int(1\wedge\||\mathbf{s}\||)\,\big(1 - \mathbb{1}_K(\mathbf{s})\big)\,q^{(R)}(\mathrm{d}\mathbf{s}) < \varepsilon$. Moreover, as a compact subset, K is bounded, i.e. $\sup_{\mathbf{s}\in K}\||\mathbf{s}\|| = C < \infty$.

For all $n \geq 1$, recall that $\overline{T}_{1,n}$ and $\mathcal{T}_1$ are endowed with probability measures. Remark 3.2 therefore ensures that $\mathrm{d}_{\mathrm{GHP}}(\overline{T}_{1,n}, \mathcal{T}_1) \leq 2 \vee |\overline{T}_{1,n}| \vee |\mathcal{T}_1|$. As a result, thanks to Lemma 3.14,

$$\sup_n \mathbb{E}\Big[\big(\mathrm{d}_{\mathrm{GHP}}(\overline{T}_{1,n}, \mathcal{T}_1)\big)^2\Big] \leq 3\Big(2^2 + \sup_n \mathbb{E}\big[|\overline{T}_{1,n}|^2\big] + \mathbb{E}\big[|\mathcal{T}_1|^2\big]\Big) \leq 12 + 6\,h_2 < \infty,$$

so the sequence $\left[\mathrm{d_{GHP}}(\overline{T}_{1,n},\mathcal{T}_1)\right]_n$ is bounded in L^2. Since by assumption, it converges to 0 a.s., it also does so in L^1. Furthermore, $\sup_n \mathbb{E}[\mathrm{d_{GHP}}(\overline{T}_{1,n},\mathcal{T}_1)] =: D$ is finite. Consequently, and because the sequence of families $\left\{(T_{i,n})_n,\mathcal{T}_i\right\}_{i\geq 1}$ is i.i.d., for any $\eta>0$, there exists N such that for all $i\geq 1$ and $n\geq N$, $\mathbb{E}\left[\mathrm{d_{GHP}}\left(\overline{T}_{i,n},\mathcal{T}_i\right)\right]<\eta$. This gives the rather crude following bound

$$\mathbb{E}\left[\mathrm{d_{GHP}}(\overline{T}_{i,n},\mathcal{T}_i)\right]\leq D\,\mathbb{1}_{n<N}+\eta.$$

For all $\delta>0$, in light of Lemma 4.6, there exists an integer $m_{K,\delta}$ which depends only on K and δ such that $\sup_{\mathbf{s}\in K}\sum_{i>m_{K,\delta}} s_i<\delta$. Then for all $R\geq 1$ and $\lambda\in\mathcal{P}_{<\infty}$ with $\lambda/R^{1/\gamma}\in K$, if $\gamma\leq 1$, Jensen's inequality gives

$$\begin{aligned}\mathbb{E}\left[\sup_{i>m_{K,\delta}}\left(\frac{\lambda_i^\gamma}{R}|\overline{T}_{i,\lambda_i}|\right)\right]&\leq\left(\mathbb{E}\left[\sup_{i>m_{K,\delta}}\frac{\lambda_i}{R^{1/\gamma}}|\overline{T}_{i,\lambda_i}|^{1/\gamma}\right]\right)^\gamma\\&\leq\left(\sum_{i>m_{K,\delta}}\frac{\lambda_i}{R^{1/\gamma}}\mathbb{E}\left[|\overline{T}_{i,\lambda_i}|^{1/\gamma}\right]\right)^\gamma\leq(h_{1/\gamma})^\gamma\,\delta^\gamma\end{aligned}$$

where $h_{1/\gamma}$ is the constant from Lemma 3.14. Otherwise, if $\gamma>1$, since $(\lambda_i)_{i\geq1}$ is a non-increasing sequence,

$$\begin{aligned}\mathbb{E}\left[\sup_{i>m_{K,\delta}}\left(\frac{\lambda_i^\gamma}{R}|\overline{T}_{i,\lambda_i}|\right)\right]&\leq\left(\frac{\lambda_{m_{K,\delta}+1}}{R^{1/\gamma}}\right)^{\gamma-1}\mathbb{E}\left[\sup_{i>m_{K,\delta}}\frac{\lambda_i}{R^{1/\gamma}}|\overline{T}_{i,\lambda_i}|\right]\\&\leq\delta^{\gamma-1}\sum_{i>m_{K,\delta}}\frac{\lambda_i}{R^{1/\gamma}}\mathbb{E}\left[|\overline{T}_{i,\lambda_i}|\right]\leq h_1\,\delta^\gamma\end{aligned}$$

where h_1 is defined as in Lemma 3.14. Similarly,

$$\mathbb{E}\left[\sup_{i>m_{K,\delta}}\left(\frac{\lambda_i^\gamma}{R}|\mathcal{T}_i|\right)\right]\leq\begin{cases}(h_{1/\gamma})^\gamma\,\delta^\gamma & \text{if }\gamma\leq 1,\\ h_1\,\delta^\gamma & \text{if }\gamma>1.\end{cases}$$

In summary, for all λ in $\mathcal{P}_{<\infty}$ such that $\lambda/R^{1/\gamma}$ belongs to K, we get that

$$\sum_{i>m_{K,\delta}}\frac{\lambda_i}{R^{1/\gamma}}\leq\delta\qquad\text{and}\qquad\mathbb{E}\left[\sup_{i>m_{K,\delta}}\left(\frac{\lambda_i^\gamma}{R}|\overline{T}_{i,\lambda_i}|\right)+\sup_{i>m_{K,\delta}}\left(\frac{\lambda_i^\gamma}{R}|\mathcal{T}_i|\right)\right]\leq B\,\delta^\gamma$$

for some finite constant B independent of ε, η, δ and K.

Therefore, for all positive ε, δ, η, and any $R\geq 1$,

$$\begin{aligned}\Delta_R\leq\varepsilon+R\,\mathbb{E}\Bigg[\mathbb{1}_K\left(\frac{\Lambda}{R^{1/\gamma}}\right)\left(1\wedge\frac{\|\Lambda\|}{R^{1/\gamma}}\right)\wedge\Bigg(\sum_{i=1}^{m_{K,\delta}}\left(\frac{\Lambda_i^\gamma}{R}\vee\frac{\Lambda_i}{R^{1/\gamma}}\right)\mathbb{E}\left[\mathrm{d_{GHP}}\left(\overline{T}_{i,\Lambda_i},\mathcal{T}_i\right)\,\middle|\,\Lambda\right]\\+\mathbb{E}\left[\sup_{i>m_{K,\delta}}\frac{\Lambda_i^\gamma}{R}\,|\overline{T}_{i,\Lambda_i}|+\sup_{i>m_{K,\delta}}\frac{\Lambda_i^\gamma}{R}\,|\mathcal{T}_i|+2\sum_{i>m_{K,\delta}}\frac{\Lambda_i}{R^{1/\gamma}}\,\middle|\,\Lambda\right]\Bigg)\Bigg]\\\leq\varepsilon+R\,\mathbb{E}\left[\left(1\wedge\frac{\|\Lambda\|}{R^{1/\gamma}}\right)\wedge\left((C+C^\gamma)\,m_{K,\delta}\eta+\left(\frac{N^\gamma}{R}+\frac{N}{R^{1/\gamma}}\right)m_{K,\delta}D+2\delta+B\delta^\gamma\right)\right].\end{aligned}$$

Let δ be such that $(2\delta+\delta^\gamma)B<\varepsilon$ and set $\eta<\varepsilon/[(C+C^\gamma)m_{K,\delta}]$. Because of Assumption $(\mathtt{I})$, we therefore get that

$$\limsup_{R\to\infty}\Delta_R\leq\varepsilon+\int_{\mathcal{S}^\downarrow}(2\varepsilon)\wedge\|\|s\,I(\mathrm{d}\mathbf{s}).$$

The monotone convergence theorem implies that the right hand side of this last inequality vanishes when ε decreases to 0. This proves that $\Delta_R\to 0$, which concludes this proof. □

Since the conclusion of Lemma 4.7 is met for any Lipschitz continuous function $G : \mathcal{S}^{\downarrow} \times \mathbb{T}_c \to \mathbb{R}_+$ with $G(\mathbf{s}, \cdot\,) \leq 1 \wedge \|\|s$, Lemma 3.21 gives the following corollary:

Corollary 4.8. *The convergence of Lemma 4.7 holds for any continuous G with $G(\mathbf{s}, \cdot\,) \leq 1 \wedge \|\|s$.*

We will now endeavour to prove that the point processes associated to adequately rescaled Markov branching trees with a unique infinite spine converge in distribution to the point process associated to fragmentation trees with immigration. Let Π be a Poisson point process on $\mathbb{R}_+ \times \mathcal{S}^{\downarrow} \times \mathbb{T}_c$ with intensity $\mathrm{d}u \otimes \mathscr{I}(\mathrm{d}\mathbf{s}, \mathrm{d}\tau)$, where $\mathscr{I}$ is the measure defined at the beginning of Section 3.3. Observe that for all $K \geq 0$,

$$\int \mathbb{1}_{u \leq K} \big(1 \wedge \|\|s\big)\, \mathrm{d}u \otimes \mathscr{I}(\mathrm{d}\mathbf{s}, \mathrm{d}\tau) = K \int_{\mathcal{S}^{\downarrow}} \big(1 \wedge \|\|s\big)\, I(\mathrm{d}\mathbf{s}) < \infty.$$

Campbell's theorem (see [38, Section 3.2]) therefore ensures that Π a.s. satisfies the integrability conditions necessary to belong to the set $\mathscr{R}$ of point measures on $\mathbb{R}_+ \times \mathcal{S}^{\downarrow} \times \mathbb{T}_c$ defined in Section 3.3.

Let T have distribution $\mathrm{MB}_{\infty}^{q, q_\infty}$. By construction of Markov branching trees with a unique infinite spine (see Remark 2.8), there exists a sequence $(\Lambda_n, T_n)_{n \geq 0}$ of i.i.d. random variables such that $T = \mathtt{b}_\infty \bigotimes_{n \geq 0} (\mathtt{v}_n, T_n)$, where Λ_n is distributed according to q_* and conditionally on $\Lambda_n = \lambda$, T_n has distribution $\mathrm{MB}_{\lambda}^{q}$. For all $R \geq 1$, let Π_R be the point process associated to $(T/R, \mu_T/R^{1/\gamma})$, i.e. the $\mathscr{R}$-valued random variable defined for all measurable $f : \mathbb{R}_+ \times \mathcal{S}^{\downarrow} \times \mathbb{T}_c \longrightarrow \mathbb{R}_+$ by

$$\int f \, \mathrm{d}\Pi_R := \sum_{n \geq 0} f\big[n/R,\, \Lambda_n/R^{1/\gamma},\, (T_n/R, \mu_{T_n}/R^{1/\gamma})\big].$$

Lemma 4.9. *With respect to the topology on $\mathscr{R}$ introduced in Section 3.3, Π_R converges to Π in distribution as R goes to infinity.*

Proof. In light of Proposition 3.20, it will be enough to prove that for any function F in the set $\mathscr{F}$, the Laplace transform of Π_R evaluated in F converges to that of Π. Fix such F in $\mathscr{F}$ and recall that it is continuous and that there exists $K \geq 0$ such that $0 \leq F(u, \mathbf{s}, \tau) \leq \|\|s\, \mathbb{1}_{u \leq K}$ for all $(u, \mathbf{s}, \tau)$. Campbell's theorem for Poisson point processes gives

$$L_\Pi(F) = \exp\Big(- \int \big[1 - \mathrm{e}^{-F(u, \mathbf{s}, \tau)}\big]\, \mathrm{d}u \otimes \mathscr{I}(\mathrm{d}\mathbf{s}, \mathrm{d}\tau)\Big).$$

For all $R \geq 1$ and $u \geq 0$, set

$$\begin{aligned} \varphi_R(u) &:= R\, \mathbb{E}\Big[1 - \exp\Big(- F\big[u, \Lambda_0/R^{1/\gamma}, (T_0/R, \mu_{T_0}/R^{1/\gamma})\big]\Big)\Big], \\ \text{and} \qquad \varphi(u) &:= \int \mathbb{E}\big[1 - \exp\big(- F[u, \mathbf{s}, \mathcal{T}_{\langle \mathbf{s} \rangle}]\big)\big]\, I(\mathrm{d}\mathbf{s}). \end{aligned}$$

Using these notations, we may write $\log L_\Pi(F) = - \int_0^K \varphi(u)\, \mathrm{d}u$ and thanks to the i.i.d. nature of the sequence $(\Lambda_n, T_n)_{n \geq 0}$, for all $R \geq 1$,

$$\begin{aligned} \log L_{\Pi_R}(F) &= - \sum_{n=0}^{\lfloor KR \rfloor} \log \mathbb{E}\Big[\exp\Big(- F\big[n/R, \Lambda_0/R^{1/\gamma}, (T_0/R, \mu_{T_0}/R^{1/\gamma})\big]\Big)\Big] \\ &= - \sum_{n=0}^{\lfloor KR \rfloor} \log \Big(1 - 1/R \cdot \varphi_R(n/R)\Big). \end{aligned}$$

The functions φ_R, $R \geq 1$ and φ all have support in $[0, K]$ and are continuous (in light of the dominated convergence theorem). Observe that $0 \leq 1 - \mathrm{e}^{-F(u, \mathbf{s}, \tau)} \leq 1 \wedge \|\|s$. From Corollary 4.8, we know that for all fixed $u \geq 0$, $\varphi_R(u) \to \varphi(u)$ as $R \to \infty$ and that furthermore

$$\sup_{R \geq 1} \sup_{u \geq 0} \varphi_R(u) \leq \sup_{R \geq 1} R\, \mathbb{E}\big[1 \wedge (\|\Lambda_0\|/R^{1/\gamma})\big] < \infty,$$

i.e. that the sequence $(\varphi_R)_{R\geq 1}$ is uniformly bounded by a finite constant, say C. Let ε be positive. It also follows from Corollary 4.8 that there exists a compact subset A of $\mathcal{S}^{\downarrow} \times \mathbb{T}_c$ with

$$\sup_{R\geq 1} \; R\,\mathbb{E}\Big[\big(1 \wedge (\|\Lambda_0\|/R^{1/\gamma})\big) \cdot \mathbb{1}_{A^c}\big(\Lambda_0/R^{1/\gamma}, (T_0/R, \mu_{T_0}/R^{1/\gamma})\big)\Big] < \varepsilon.$$

Recall that F is continuous, hence there exists $\delta > 0$ such that for any (u, s, τ) and (u', s', τ') in the compact set $[0, K] \times A$, if $|u - u'| + \|\mathrm{s} - \mathrm{s}'\| + \mathrm{d}_{\mathrm{GHP}}(\tau, \tau') < \delta$, then $|F(u, \mathrm{s}, \tau) - F(u', \mathrm{s}', \tau')| < \varepsilon$. As a result, and because $x \mapsto \mathrm{e}^{-x}$ is 1-Lipschitz continuous on $\mathbb{R}_+$, for all $R \geq 1$ and u, v in $[0, K]$ with $|u - v| < \delta$,

$$\begin{aligned}\big|\varphi_R(u) - \varphi_R(v)\big| &\leq R\,\mathbb{E}\Big[1 \wedge \Big|F\big[u, \Lambda_0/R^{1/\gamma}, (T_0/R, \mu_{T_0}/R^{1/\gamma})\big] \\ &\qquad\qquad - F\big[v, \Lambda_0/R^{1/\gamma}, (T_0/R, \mu_{T_0}/R^{1/\gamma})\big]\Big|\Big] \\ &\leq \varepsilon + R\,\mathbb{E}\Big[\big(\varepsilon \wedge (\|\Lambda_0\|/R^{1/\gamma})\big) \cdot \mathbb{1}_{A}\big(\Lambda_0/R^{1/\gamma}, (T_0/R, \mu_{T_0}/R^{1/\gamma})\big)\Big],\end{aligned}$$

and in light of Corollary 4.8 and the monotone convergence theorem, we get

$$\limsup_{R\to\infty} \big|\varphi_R(u) - \varphi_R(v)\big| \leq \varepsilon + \textstyle\int \varepsilon \wedge \|\|s\, I(\mathrm{ds}) \xrightarrow[\varepsilon\to 0]{} 0.$$

This ensures that the sequence $(\varphi_R)_{R\geq 1}$ is equicontinuous on $[0, K]$. It follows from the Arzelà-Ascoli theorem that φ_R converges uniformly to φ. In turn, we deduce that

$$\left|\frac{1}{R}\sum_{n=0}^{\lfloor KR\rfloor}\varphi_R(n/R) - \frac{1}{R}\sum_{n=0}^{\lfloor KR\rfloor}\varphi(n/R)\right| \leq \frac{KR+1}{R} \sup_{0\leq u\leq K}\big|\varphi_R(u) - \varphi(u)\big| \xrightarrow[R\to\infty]{} 0.$$

Observe that for all $R \geq 1$, we may write

$$\log L_{\Pi_R}(F) - 1/R \cdot \sum_{n=0}^{\lfloor KR\rfloor}\varphi_R(n/R) = \sum_{n=0}^{\lfloor KR\rfloor}\Big(-\log\big[1 - 1/R\cdot\varphi_R(n/R)\big] - 1/R\cdot\varphi_R(n/R)\Big).$$

Recall that $\sup_{R\geq 1, u\geq 0}\varphi_R(u) \leq C$. Therefore, because the function $[0, 1) \to \mathbb{R}_+$, $x \mapsto -\log(1 - x) - x$ increases with x, for any $R \geq C$ and $n \geq 0$, we get

$$\Big|-\log\big[1 - 1/R\cdot\varphi_R(n/R)\big] - 1/R\cdot\varphi_R(n/R)\Big| \leq \big|-C/R - \log(1 - C/R)\big| = o(1/R).$$

Consequently,

$$\Big|\log L_{\Pi_R}(F) - 1/R \cdot \sum_{n=0}^{\lfloor KR\rfloor}\varphi_R(n/R)\Big| \leq (KR + 1)\,\big|-C/R - \log(1 - C/R)\big| \xrightarrow[R\to\infty]{} 0.$$

Finally, as Riemann sums of the continuous function φ,

$$\frac{1}{R}\sum_{n=0}^{\lfloor KR\rfloor}\varphi(n/R) \xrightarrow[R\to\infty]{} \textstyle\int_0^K \varphi(u)\,\mathrm{d}u = \log\, L_\Pi(F).$$

In summary, $\log\, L_{\Pi_R}(F) \to \log\, L_\Pi(F)$ when $R \to \infty$. □

4.2 Proof of Theorem 4.2

Now that we know that the underlying point processes converge, we can prove convergence of the trees themselves.

Recall that the topology we defined on $\mathscr{R}$ in Section 3.3 makes it a Polish topological space. As such, Skorokhod's representation theorem holds for $\mathscr{R}$-valued random variables. In particular, because of Lemma 4.9, there exist:

- A Poisson point process Π with intensity $\mathrm{d}u \otimes \mathscr{I}(\mathrm{d}\mathbf{s}, \mathrm{d}\tau)$,
- A family $\{(\Lambda_n^{(R)}, \tau_n^{(R)})_{n\geq 0}; R \in \mathbb{N}\}$ such that for all fixed $R \geq 1$, $(\Lambda_n^{(R)}, \tau_n^{(R)})_{n\geq 0}$ is an i.i.d. sequence, $\Lambda_n^{(R)}$ follows q_* and conditionally on $\Lambda_n^{(R)} = \lambda$, $\tau_n^{(R)}$ has distribution MB_λ^q and is endowed with the measure $\mu_{\tau_n^{(R)}} := \sum_{u\in\tau_n^{(R)}} \delta_u$,

such that if for any R we let Π_R be the random element of $\mathscr{R}$ defined for all measurable $f : \mathbb{R}_+ \times \mathcal{S}^\downarrow \times \mathbb{T}_c \longrightarrow \mathbb{R}_+$ by $\int f \,\mathrm{d}\Pi_R := \sum_{n\geq 0} f\big[n/R, \Lambda_n^{(R)}/R^{1/\gamma}, (\tau_n^{(R)}/R, \mu_{\tau_n^{(R)}}/R^{1/\gamma})\big]$, then Π_R a.s. converges to Π when $R \to \infty$.

Let $\{(u_i, \mathbf{s}_i, \mathcal{T}_i); i \geq 1\}$ be the atoms of Π and set $\Sigma := \sum_{i\geq 1} \delta_{(u_i,\mathbf{s}_i)}$. By definition of the intensity measure of Π, there exists a family $\{\mathcal{T}_{i,j}\,;\, i, j \geq 1\}$ of i.i.d. (γ, ν)-fragmentation trees independent of Σ such that for all $i \geq 1$, $\mathcal{T}_i := \langle (s_{i,j}^\gamma \mathcal{T}_{i,j}, s_{i,j}\mu_{\mathcal{T}_{i,j}}); j \geq 1\rangle$. Set $\mathcal{T}^{(I)} := \mathbf{G}(\{(u_i, \mathcal{T}_i); i \geq 1\})$ where $\mathbf{G}$ is the continuum grafting function defined in Section 3.1.2 and recall that it is a (γ, ν)-fragmentation tree with immigration I (see Section 3.2.2). For all $\varepsilon > 0$, let

$$\mathcal{T}_\varepsilon^{(I)} := \mathbf{G}\big(\{(u_i, \mathcal{T}_i); i \geq 1, \|\mathbf{s}_i\| \geq \varepsilon\}\big).$$

This tree can be thought of as $\mathcal{T}^{(I)}$ on which all sub-trees grafted on the spine with mass less than ε have been cut away. Observe that because of the definition of the function $\mathbf{G}$, the measure on $\mathcal{T}_\varepsilon^{(I)}$ is simply the restriction of $\mu_{\mathcal{T}^{(I)}}$ to $\mathcal{T}_\varepsilon^{(I)}$.

For all R, set $\tau^{(R)} := \mathtt{b}_\infty \bigotimes_{n\geq 0}(\mathtt{v}_n, \tau_n^{(R)})$ and denote its counting measure by $\mu_{\tau^{(R)}}$. Observe that $\tau^{(R)}$ is distributed according to $\mathrm{MB}_\infty^{q,q_\infty}$. Let $T^{(R)} := (R^{-1}\tau^{(R)}, R^{-1/\gamma}\mu_{\tau^{(R)}})$ be the rescaled infinite Markov branching tree associated to Π_R. Moreover, for all positive ε, let $T_\varepsilon^{(R)}$ be the tree obtained by removing from $T^{(R)}$ all the sub-trees grafted on its spine with mass less than ε, i.e. set

$$T_\varepsilon^{(R)} := \mathbf{G}\Big(\Big\{\big[n/R, (R^{-1}\tau_n^{(R)}, R^{-1/\gamma}\mu_{\tau_n^{(R)}})\big] \;\big|\; n \geq 0 : \|\Lambda_n^{(R)}\| \geq R^{1/\gamma}\varepsilon\Big\}\Big).$$

The tree $T_\varepsilon^{(R)}$ is clearly a subset of $T^{(R)}$ and it is endowed with the restriction of $\mu_{T^{(R)}}$.

In this section we will endeavour to prove Theorem 4.2. In order to do so, we will use the following criterion for convergence in distribution.

Theorem 4.10 ([16], Theorem 3.2)**.** *Let (M, d) be a metric space. If X_n, $X_n^{(k)}$, $X^{(k)}$, $n \geq 1$, $k \geq 1$ and X are M-valued random variables satisfying:*

(i) For all $k \geq 1$, $X_n^{(k)} \Rightarrow X^{(k)}$ as $n \to \infty$,

(ii) $X^{(k)} \Rightarrow X$ as $k \to \infty$,

(iii) For any positive η, $\lim_{k\to\infty} \limsup_{n\to\infty} \mathbb{P}\big[\mathrm{d}(X_n^{(k)}, X_n) > \eta\big] = 0$,

Then X_n converges to X in distribution.

Remark 4.11. Condition (i) is akin to finite-dimensional convergence of X_n to X and Conditions (ii) and (iii) to tightness of $(X_n)_n$.

In our setting, the sequence $(T^{(R)}; R \in \mathbb{N})$ of rescaled $\mathrm{MB}_\infty^{q,q_\infty}$ trees will play the role of $(X_n)_n$ and the limit variable X will be $\mathcal{T}^{(I)}$, a (γ, ν)-fragmentation tree with immigration I. The intermediate family $(X_n^{(k)})_{n,k}$ will be replaced by $(T_\varepsilon^{(R)}; R \geq 1)$ with $\varepsilon \to 0$ along some countable subset of $(0, \infty)$. Similarly, we'll consider $\mathcal{T}_\varepsilon^{(I)}$ trees instead of $(X^{(k)})_k$.

Lemma 4.12. *With these notations, $\mathcal{T}_\varepsilon^{(I)}$ a.s. converges to $\mathcal{T}^{(I)}$ as $\varepsilon \to 0$ with respect to $\mathrm{D}_{\mathrm{GHP}}$.*

Proof. For all $\varepsilon > 0$, let C_ε be the correspondence between $\mathcal{T}^{(I)}$ and $\mathcal{T}_\varepsilon^{(I)}$ defined by $C_\varepsilon := \{(x,x) : x \in \mathcal{T}_\varepsilon^{(I)}\} \cup \bigcup_{i\geq 1 : \|\mathbf{s}_i\| < \varepsilon} \mathcal{T}_i \times \{u_i\}$ and set π_ε, the boundedly finite Borel measure on $\mathcal{T}^{(I)} \times \mathcal{T}_\varepsilon^{(I)}$, such that for all Borel A, $\pi_\varepsilon(A) := \int_{\mathcal{T}_\varepsilon^{(I)}} \mathbb{1}_A(x,x)\, \mu_{\mathcal{T}_\varepsilon^{(I)}}(\mathrm{d}x)$. Let $K \geq 0$ be fixed. Call $\pi_\varepsilon|_K$ the restriction of π_ε to $\mathcal{T}^{(I)}|_K \times \mathcal{T}_\varepsilon^{(I)}|_K$. The monotone convergence theorem yields

$$\mathrm{D}\big(\pi_\varepsilon|_K; \mu_{\mathcal{T}^{(I)}}|_K, \mu_{\mathcal{T}_\varepsilon^{(I)}}|_K\big) = \pi_\varepsilon|_K(C_\varepsilon^c) \leq \int \|\|\|s \mathbb{1}_{\|\|\|s<\varepsilon} \mathbb{1}_{u \leq K} \Sigma(\mathrm{d}u, \mathrm{d}\mathbf{s}) \xrightarrow[\varepsilon\to 0]{\text{a.s.}} 0.$$

Let $C_\varepsilon|_K := C_\varepsilon \cap \big(\mathcal{T}^{(I)}|_K \times \mathcal{T}_\varepsilon^{(I)}|_K\big)$ and observe that it is a correspondence between $\mathcal{T}^{(I)}|_K$ and $\mathcal{T}_\varepsilon^{(I)}|_K$. Its distortion satisfies

$$\operatorname{dis} C_\varepsilon|_K \leq 2 \sup \Big\{ |\mathcal{T}_i| \,:\, i \geq 1, u_i \leq K, \|\mathbf{s}_i\| < \varepsilon \Big\} \xrightarrow[\varepsilon\to 0]{\text{a.s.}} 0.$$

As a result, $\mathrm{d}_{\mathrm{GHP}}\big(\mathcal{T}^{(I)}|_K, \mathcal{T}_\varepsilon^{(I)}|_K\big) \to 0$ a.s. as $\varepsilon \to 0$. Since this holds for all $K \geq 0$, Proposition 3.9 (ii) ensures that $\mathrm{D}_{\mathrm{GHP}}\big(\mathcal{T}^{(I)}, \mathcal{T}_\varepsilon^{(I)}\big)$ a.s. converges to 0 when $\varepsilon \to 0$. □

Lemma 4.13. *For all positive η,*

$$\lim_{\varepsilon\to 0} \limsup_{R\to\infty} \mathbb{P}\Big[\mathrm{D}_{\mathrm{GHP}}\big(T^{(R)}, T_\varepsilon^{(R)}\big) > \eta\Big] = 0.$$

Proof. We will proceed in a way similar to the proof of Lemma 4.12. For all $R \geq 1$ and $\varepsilon > 0$, define the correspondence $C_\varepsilon^{(R)}$ between $T^{(R)}$ and $T_\varepsilon^{(R)}$ as $C_\varepsilon^{(R)} := \{(u,u) : u \in T_\varepsilon^{(R)}\} \cup \{(u, n/R) : n \geq 1, \|\Lambda_n\| < R^{1/\gamma}\varepsilon, u \in \tau_n^{(R)}\}$ and let $\pi_\varepsilon^{(R)}$ be the boundedly finite measure $T^{(R)} \times T_\varepsilon^{(R)}$ defined for all Borel sets A by $\pi_\varepsilon^{(R)}(A) := \int_{T_\varepsilon^{(R)}} \mathbb{1}_A(x,x)\, \mu_{T_\varepsilon^{(R)}}(\mathrm{d}x)$.

For all $K \geq 0$, set $C_\varepsilon^{(R)}\big|_K := C_\varepsilon^{(R)} \cap \big(T^{(R)}|_K \times T_\varepsilon^{(R)}|_K\big)$, which is a correspondence between $T^{(R)}|_K$ and $T_\varepsilon^{(R)}|_K$, and let $\pi_\varepsilon^{(R)}\big|_K$ be the restriction of $\pi_\varepsilon^{(R)}$ to $T^{(R)}|_K \times T_\varepsilon^{(R)}|_K$. Then, for any non-negative K,

$$\operatorname{dis}_{T^{(R)}|_K, T_\varepsilon^{(R)}|_K} C_\varepsilon^{(R)}\big|_K \leq \frac{2}{R} \sup \Big\{ |\tau_n^{(R)}| \,:\, 0 \leq n \leq RK,\ \|\Lambda_n^{(R)}\| < R^{1/\gamma}\varepsilon \Big\}.$$

For all $n \geq 0$ and $R \geq 1$, $|\tau_n^{(R)}| = 1 + \sup\{|\tau_{n,i}^{(R)}| : 1 \leq i \leq p(\Lambda_n^{(R)})\}$. Further observe that thanks to Lemma 3.14, we can find a finite constant h such that for all $n \geq 0$, $R \geq 1$ and $i = 1, \ldots, p(\Lambda_n^{(R)})$, $\mathbb{E}[(1 + |\tau_{n,i}^{(R)}|)^{1/\gamma} | \Lambda_n^{(R)}] \leq h\, \Lambda_n^{(R)}(i)$. Therefore, since the sequence $(\Lambda_n^{(R)}, \tau_n^{(R)})_{n\geq 1}$ is i.i.d.,

$$\begin{aligned}
\mathbb{E}\Big[\Big(\operatorname{dis}_{T^{(R)}|_K, T_\varepsilon^{(R)}|_K} C_\varepsilon^{(R)}\big|_K \Big)^{1/\gamma}\Big] \\
&\leq (KR+1)\, \frac{2^{1/\gamma}}{R^{1/\gamma}}\, \mathbb{E}\Big[\textstyle\sum_{i=1}^{p(\Lambda_0^{(R)})} (1 + |\tau_{0,i}^{(R)}|)^{1/\gamma} \mathbb{1}_{\|\Lambda_0^{(R)}\| < R^{1/\gamma}\varepsilon}\Big] \\
&\leq (KR+1)\, \frac{2^{1/\gamma} h}{R^{1/\gamma}}\, \mathbb{E}\Big[\|\Lambda_0^{(R)}\|\, \mathbb{1}_{\|\Lambda_0^{(R)}\| < R^{1/\gamma}\varepsilon}\Big].
\end{aligned}$$

Similarly,

$$\begin{aligned}
\mathbb{E}\Big[\mathrm{D}\Big(\pi_\varepsilon^{(R)}\big|_K; \mu_{T^{(R)}}\big|_K, \mu_{T_\varepsilon^{(R)}}\big|_K\Big)\Big] &= \mathbb{E}\Big[\pi_\varepsilon^{(R)}\big|_K\big[(C_\varepsilon^{(R)})^c\big]\Big] \\
&= (KR+1)\, \frac{1}{R^{1/\gamma}}\, \mathbb{E}\Big[\|\Lambda_0^{(R)}\|\, \mathbb{1}_{\|\Lambda_0^{(R)}\| < R^{1/\gamma}\varepsilon}\Big].
\end{aligned}$$

In light of Assumption (I),

$$\begin{aligned}
(KR+1)\, \frac{1}{R^{1/\gamma}}\, \mathbb{E}\Big[\|\Lambda_0^{(R)}\|\, \mathbb{1}_{\|\Lambda_0^{(R)}\| < R^{1/\gamma}\varepsilon}\Big] \\
&\leq (KR+1)\, \mathbb{E}\Big[\varepsilon \wedge \frac{\|\Lambda_0^{(R)}\|}{R^{1/\gamma}}\Big] \xrightarrow[R\to\infty]{} K \int (\varepsilon \wedge \|\|\|s)\, I(\mathrm{d}\mathbf{s}).
\end{aligned}$$

Finally, for any positive η, if $K > -\log(\eta/2)$, using Markov's inequality and the monotone convergence theorem,

$$
\begin{aligned}
\limsup_{R\to\infty} \mathbb{P}\Big[\mathrm{D}_{\mathrm{GHP}}\big(T^{(R)}, T_\varepsilon^{(R)}\big) > \eta\Big] &\le \limsup_{R\to\infty} \mathbb{P}\Big[\mathrm{D}_{\mathrm{GHP}}\big(T^{(R)}|_K, T_\varepsilon^{(R)}|_K\big) > \eta - 2\mathrm{e}^{-K}\Big] \\
&\le \limsup_{R\to\infty} \left(\frac{\mathbb{E}\Big[\big(\operatorname{dis}_{T^{(R)}|_K, T_\varepsilon^{(R)}|_K} C_\varepsilon^{(R)}\big|_K\big)^{1/\gamma}\Big]}{(\eta - 2\mathrm{e}^{-K})^{1/\gamma}} + \frac{\mathbb{E}\Big[\mathrm{D}\big(\pi_\varepsilon^{(R)}\big|_K; \mu_{T^{(R)}}\big|_K, \mu_{T_\varepsilon^{(R)}}\big|_K\big)\Big]}{\eta - 2\mathrm{e}^{-K}} \right) \\
&\le \left(\frac{2^{1/\gamma} K h}{(\eta - 2\mathrm{e}^{-K})^{1/\gamma}} + \frac{K}{\eta - 2\mathrm{e}^{-K}} \right) \int (\varepsilon \wedge \|\!|\!| s)\, I(\mathrm{d}\mathbf{s}) \xrightarrow[\varepsilon\to 0]{} 0.
\end{aligned}
$$

□

The next result is both intuitive and easy to prove. Its proof will therefore be left to the reader.

Lemma 4.14. *Fix n a positive integer and let $\mathbf{G}_n$ be the restriction of $\mathbf{G}$ to $(\mathbb{R}_+ \times \mathbb{T}_c)^n$; $\mathbf{G}_n$ is a continuous function for the product topology.*

Lemma 4.15. *Let $K \ge 0$ and $\varepsilon > 0$ be fixed. Almost surely, for any continuous $F : \mathbb{R}_+ \times \mathcal{S}^\downarrow \times \mathbb{T}_c \to \mathbb{R}_+$ bounded by 1,*

$$
\begin{aligned}
&\limsup_{R\to\infty} \int F(u,\mathbf{s},\tau)\, \mathbb{1}_{u\le K,\, \|\!|\!| s \ge \varepsilon}\, \mathrm{d}\Pi_R(u,\mathbf{s},\tau) \le \int F(u,\mathbf{s},\tau)\, \mathbb{1}_{u\le K,\, \|\!|\!| s \ge \varepsilon}\, \mathrm{d}\Pi(u,\mathbf{s},\tau), \\
\text{and} \quad &\liminf_{R\to\infty} \int F(u,\mathbf{s},\tau)\, \mathbb{1}_{u< K,\, \|\!|\!| s > \varepsilon}\, \mathrm{d}\Pi_R(u,\mathbf{s},\tau) \ge \int F(u,\mathbf{s},\tau)\, \mathbb{1}_{u< K,\, \|\!|\!| s > \varepsilon}\, \mathrm{d}\Pi(u,\mathbf{s},\tau).
\end{aligned}
$$

Proof. Let φ and φ_n, $n \ge 1$ be the functions from $\mathbb{R}_+ \times \mathcal{S}^\downarrow \times \mathbb{T}_c$ to $\mathbb{R}_+$ defined for all $(u,\mathbf{s},\tau)$ by $\varphi(u,\mathbf{s},\tau) := \mathbb{1}_{u\le K}\, \mathbb{1}_{\|\!|\!| s \ge \varepsilon}$ and $\varphi_n(u,\mathbf{s},\tau) := [1 - n(u-K)_+]_+ \times [1 - n(\varepsilon - \|\!|\!| s)_+]_+$ respectively (where $x_+ = x \vee 0$ for any real number x). Observe that for all $n \ge 1$, φ_n is continuous and that for n large enough, $\varepsilon\varphi_n F$ is an element of $\mathscr{F}$. Therefore, everywhere on the event $\{\Pi_R \to \Pi\}$, $\int \varphi_n F \,\mathrm{d}\Pi_R \to \int \varphi_n F \,\mathrm{d}\Pi$ for any fixed $n \ge 1$. Furthermore, $\varphi_n \downarrow_n \varphi$ so the monotone convergence theorem yields $\inf_{n\ge 1} \int \varphi_n F \,\mathrm{d}\Pi = \int \varphi F \,\mathrm{d}\Pi$ and for all $R \ge 1$, $\inf_{n\ge1} \int \varphi_n F \,\mathrm{d}\Pi_R = \int \varphi F \,\mathrm{d}\Pi_R$. As a result, on $\{\Pi_R \to \Pi\}$,

$$
\limsup_{R\to\infty} \int \varphi F \,\mathrm{d}\Pi_R \le \inf_{n\ge1} \left[\limsup_{R\to\infty} \int \varphi_n F \,\mathrm{d}\Pi_R \right] = \int \varphi F \,\mathrm{d}\Pi.
$$

Similarly, if we let $\psi(u,\mathbf{s},\tau) := \mathbb{1}_{u<K}\, \mathbb{1}_{\|\!|\!| s > \varepsilon}$, there exists a sequence $(\psi_n)_n$ of continuous functions such that $\psi_n \uparrow_n \psi$ and for n large enough, $\varepsilon\psi_n F$ is in $\mathscr{F}$. The same kind of arguments lead to

$$
\liminf_{R\to\infty} \int \psi F \,\mathrm{d}\Pi_R \ge \sup_{n\ge1} \left[\liminf_{R\to\infty} \int \psi_n F \,\mathrm{d}\Pi_R \right] = \int \psi F \,\mathrm{d}\Pi
$$

everywhere on $\{\Pi_R \to \Pi\}$. □

Lemma 4.16. *Let ε be positive and such that $\Pi\big((u,\mathbf{s},\tau) : \|\!|\!| s = \varepsilon\big) = 0$ a.s.. Then $T_\varepsilon^{(R)}$ a.s. converges to $\mathcal{T}_\varepsilon^{(I)}$ as $R \to \infty$.*

Proof. Observe that for any $K \ge 0$, $\Pi\big((u,\mathbf{s},\tau) : u = K\big) = 0$ a.s. which implies that with probability 1, for any continuous bounded $F : \mathbb{R}_+ \times \mathcal{S}^\downarrow \times \mathbb{T}_c \to \mathbb{R}_+$,

$$
\int F(u,\mathbf{s},\tau)\, \mathbb{1}_{u\le K,\, \|\!|\!| s \ge \varepsilon}\, \mathrm{d}\Pi(u,\mathbf{s},\tau) = \int F(u,\mathbf{s},\tau)\, \mathbb{1}_{u< K,\, \|\!|\!| s > \varepsilon}\, \mathrm{d}\Pi(u,\mathbf{s},\tau).
$$

Consequently, in light of Lemma 4.15,

$$
\mathbb{1}_{u\le K,\, \|\!|\!| s \ge \varepsilon}\, \Pi_R(\mathrm{d}u, \mathrm{d}\mathbf{s}, \mathrm{d}\tau) \xrightarrow[R\to\infty]{\text{a.s.}} \mathbb{1}_{u\le K,\, \|\!|\!| s \ge \varepsilon}\, \Pi(\mathrm{d}u, \mathrm{d}\mathbf{s}, \mathrm{d}\tau).
$$

Moreover, the (finite) measures $\mathbb{1}_{u\leq K,\, \|\|\|s\geq\varepsilon}\, \Pi(\mathrm{d}u, \mathrm{d}\mathbf{s}, \mathrm{d}\tau)$ and $\mathbb{1}_{u\leq K,\, \|\|\|s\geq\varepsilon}\, \Pi_R(\mathrm{d}u, \mathrm{d}\mathbf{s}, \mathrm{d}\tau)$, $R \geq 1$ may be written as finite sums of Dirac measures. As a result, almost surely, the atoms of $\mathbb{1}_{u\leq K,\, \|\|\|s\geq\varepsilon}\, \Pi_R(\mathrm{d}u, \mathrm{d}\mathbf{s}, \mathrm{d}\tau)$ converge to those of $\mathbb{1}_{u\leq K,\, \|\|\|s\geq\varepsilon}\, \Pi(\mathrm{d}u, \mathrm{d}\mathbf{s}, \mathrm{d}\tau)$ when $R \to \infty$. Lemma 4.14 then ensures that $T_\varepsilon^{(R)}|_K$ a.s. converges to $\mathcal{T}_\varepsilon^{(I)}|_K$. Since this holds for any $K \geq 0$, Proposition 3.9 allows us to conclude. □

Proof of Theorem 4.2. Noticing that the set of all $\varepsilon > 0$ with $\mathbb{P}\big[\Pi\big((u, \mathbf{s}, \tau) : \|\|\|s = \varepsilon\big) = 0\big] < 1$ is at most countable, we may consider a sequence $(\varepsilon_k)_{k\geq 1}$ of positive real numbers which converges to 0 and such that for all k, $\Pi\big((u, \mathbf{s}, \tau) : \|\|\|s = \varepsilon_k\big) = 0$ a.s.. Lemmas 4.12, 4.13 and 4.16 then respectively prove that conditions (ii), (iii) and (i) of Theorem 4.10 are met for $T^{(R)}$, $T_{\varepsilon_k}^{(R)}$, $\mathcal{T}_{\varepsilon_k}^{(I)}$, $R \geq 1$, $k \geq 1$ and $\mathcal{T}^{(I)}$. Therefore, $T^{(R)} \Rightarrow \mathcal{T}^{(I)}$ with respect to $\mathrm{D_{GHP}}$. □

4.3 Volume growth of infinite Markov branching trees

We now turn to the proof of Proposition 4.3. Recall that if $\mathbf{T} \in \mathbb{T}$ is fixed, then $V_{\mathbf{T}}$, the volume growth function of $\mathbf{T}$, is given by

$$V_{\mathbf{T}} : \mathbb{R}_+ \longrightarrow \mathbb{R}_+,\ R \longmapsto \mu_T(T|_R).$$

Notice that $V_{\mathbf{T}}$ is a non-negative, non-decreasing càdlàg function.

Proof of Proposition 4.3. Proposition 3.9 ensures that $(\mathbb{T}, \mathrm{D_{GHP}})$ is a Polish metric space. In light of Skorokhod's representation theorem and since the assumptions of Theorem 4.2 are met, there exist a sequence $(\tau_R)_{R\geq 1}$ of $\mathrm{MB}_\infty^{q,q_\infty}$ trees as well as a (γ, ν, I)-fragmentation tree with immigration $\mathcal{T}^{(I)}$ such that $(R^{-1}\tau_R, R^{-1/\gamma}\mu_{\tau_R}) =: T^{(R)}$ a.s. converges to $\mathcal{T}^{(I)}$.

Proposition 3.9 and Remark 3.2 ensure that a.s., for all $t \geq 0$ such that $\mu_{\mathcal{T}^{(I)}}[\partial_t \mathcal{T}^{(I)}] = 0$, $V_{T^{(R)}}(t)$ converges to $V_{\mathcal{T}^{(I)}}(t)$. Now observe that $\mu_{\mathcal{T}^{(I)}}[\partial_t \mathcal{T}^{(I)}] = 0$ iff $V_{\mathcal{T}^{(I)}}$ is continuous at t. Therefore, if we prove that $V_{\mathcal{T}^{(I)}}$ is a.s. continuous on $\mathbb{R}_+$, since volume growth functions are monotone, we may use the following classical result to conclude this proof:

> *If $(f_n)_n$ is a sequence of monotone functions from a compact interval I to $\mathbb{R}$ such that $f_n \to f$ point-wise for some continuous function f, then $f_n \to f$ uniformly on I.*

Following the construction of fragmentation trees with immigration detailed in Section 3.2.2, there exist a Poisson point process $\Sigma = \sum_{i\geq 1} \delta_{(u_i, \mathbf{s}_i)}$ on $\mathbb{R}_+ \times \mathcal{S}^\downarrow$ with intensity $\mathrm{d}u \otimes I(\mathrm{d}\mathbf{s})$ and a family $\big[\mathcal{T}_{i,j}; i, j \geq 1\big]$ of i.i.d. (γ, ν)-fragmentation trees independent of Σ such that

$$\mathcal{T}^{(I)} = \mathbf{G}\left(\left\{\left(u_i, \left\langle (s_{i,j}^\gamma \mathcal{T}_{i,j}, s_{i,j}\mu_{\mathcal{T}_{i,j}}); j \geq 1 \right\rangle\right) \,:\, i \geq 1\right\}\right).$$

With these notations, we may write $V_{\mathcal{T}^{(I)}} = \sum_{i\geq 1}\sum_{j\geq 1} s_{i,j} V_{\mathcal{T}_{i,j}}\big[(\,\cdot\, - u_i)_+/s_{i,j}^\gamma\big]$. Furthermore, for any non-negative K, since $V_{\mathcal{T}_{i,j}} \leq 1$ for all $i, j \geq 1$,

$$\textstyle\sum_{i\geq 1}\sum_{j\geq 1} s_{i,j}\mathbb{1}_{u_i\leq K} = \int \mathbb{1}_{u\leq K}\, \|\|\|s\, \Sigma(\mathrm{d}u, \mathrm{d}\mathbf{s})$$

which is a.s. finite, as already noticed. As a result and in light of the Weierstrass M-test, the restriction of $V_{\mathcal{T}^{(I)}}$ to the compact interval $[0, K]$ is a series which a.s. converges uniformly on $[0, K]$.

Proposition 1.9 in [11] implies that the volume growth function of (γ, ν)-fragmentation trees is a.s. continuous. In particular, with probability one, $V_{\mathcal{T}_{i,j}}$ is continuous for all i and j. As a uniformly converging series of continuous functions, $V_{\mathcal{T}^{(I)}}|_{[0,K]}$ is a.s. continuous on $[0, K]$. Since this holds for any $K \geq 0$, $V_{\mathcal{T}^{(I)}}$ is a.s. continuous on $\mathbb{R}_+$, which concludes this proof. □

4.4 Unary immigration measures

Before concluding this section, we will state a useful criterion to prove Assumption (I) when the limit immigration measure is unary, i.e. when it is supported by the set $\{(s, 0, 0, \dots) : s > 0\}$. In light of Remark 4.1, we will only study *self-similar* unary immigration measures.

Let $\gamma \in (0, 1)$. Proposition 3.17 ensures that any unary γ-self-similar immigration measure may be written as $c\, I_\gamma^{\mathrm{un}}$ where c is a positive constant and I_γ^{un} is the measure defined by

$$\int_{\mathcal{S}^\downarrow} f \,\mathrm{d}I_\gamma^{\mathrm{un}} = \int_0^\infty f(x, 0, 0, \dots)\, x^{-1-\gamma}\,\mathrm{d}x$$

for any measurable $f : \mathcal{S}^\downarrow \to \mathbb{R}_+$.

Remark 4.17. Recall the immigration measures defined in Section 3.2.2. The Brownian immigration measure I_B is unary and may be written as $I_B = (2/\pi)^{1/2}\, I_{1/2}^{\mathrm{un}}$. On the other hand, for any $\alpha \in (1, 2)$, $I^{(\alpha)}$ isn't unary.

Lemma 4.18. *Let X be an integer valued random variable such that there exist $\gamma \in (0, 1)$ and a positive constant c satisfying $n^{1+\gamma}\mathbb{P}[X = n] \to c$. In this case, for all continuous $f : \mathbb{R}_+ \to \mathbb{R}_+$ with $f(x) \le 1 \wedge x$, $R\,\mathbb{E}[f(X/R^{1/\gamma})] \to \int_0^\infty c\, f(x)\, x^{-1-\gamma}\,\mathrm{d}x$ as R goes to infinity.*

Proof. By assumption, for all $\varepsilon > 0$, there exists an integer N such that for all $n \ge N$, $\left|n^{1+\gamma}\mathbb{P}[X = n] - c\right| < \varepsilon$. As a result

$$R \sum_{n>N} (c-\varepsilon) \frac{1}{n^{1+\gamma}} f\left(\frac{n}{R^{1/\gamma}}\right) \le R\,\mathbb{E}\left[f\left(\frac{X}{R^{1/\gamma}}\right)\right] \le R \sum_{n=1}^{N} \frac{n}{R^{1/\gamma}} + R \sum_{n>N} (c+\varepsilon) \frac{1}{n^{1+\gamma}} f\left(\frac{n}{R^{1/\gamma}}\right).$$

As a Riemann sum, $R \sum_{n>N} n^{-1-\gamma} f(n/R^{1/\gamma})$ converges toward $\int_0^\infty f(x)\, x^{-1-\gamma}\,\mathrm{d}x$ as R goes to infinity. The desired result then follows. $\square$

Proposition 4.19. *Let Λ be a random finite partition such that as $n \to \infty$, $n^{1+\gamma}\mathbb{P}[\|\Lambda\| = n] \to c$ for some $\gamma \in (0, 1)$, $c > 0$ and $n^\gamma \mathbb{P}[\Lambda_1 \ge n]$ converges to c/γ. For all $R \ge 1$, let $q^{(R)}$ be the distribution of $\Lambda/R^{1/\gamma}$. Then, $R\,(1 \wedge \|\|s)\, q^{(R)}(\mathrm{d}s)$ converges weakly to $(1 \wedge \|\|s)\, c\, I_\gamma^{\mathrm{un}}(\mathrm{d}s)$ as $R \to \infty$ in the sense of finite measures on $\mathcal{S}^\downarrow$.*

Proof. The main idea for this proof is to show that the tail of Λ is asymptotically negligible when its first component is large, or more precisely, that $R\,\mathbb{E}\big[1 \wedge \big([\|\Lambda\| - \Lambda_1]/R^{1/\gamma}\big)\big]$ converges to 0 when R goes to infinity. Since $\|\Lambda\|$ fulfils the assumptions of Lemma 4.18,

$$R\,\mathbb{E}\big[1 \wedge (\|\Lambda\|/R^{1/\gamma})\big] \xrightarrow[R\to\infty]{} c \int 1 \wedge \|\|s\, I_\gamma^{\mathrm{un}}(\mathrm{d}s) = c/[\gamma\,(1-\gamma)] =: C_\gamma$$

Furthermore, $\Lambda_1 \le \|\Lambda\|$, so we get that $\limsup_{R\to\infty} R\,\mathbb{E}\big[1 \wedge (\Lambda_1/R^{1/\gamma})\big] \le C_\gamma$. In light of Fatou's lemma and the assumption on the probability tail of Λ_1,

$$\liminf_{R\to\infty} R\,\mathbb{E}\left[1 \wedge \frac{\Lambda_1}{R^{1/\gamma}}\right] = \liminf_{R\to\infty} \int_0^1 R\,\mathbb{P}\big[\Lambda_1 \ge R^{1/\gamma} t\big]\,\mathrm{d}t \ge \int_0^1 c\gamma^{-1} t^{-\gamma}\,\mathrm{d}t = C_\gamma.$$

In summary, when $R \to \infty$, $R\,\mathbb{E}\big[1 \wedge (\Lambda_1/R^{1/\gamma})\big] \to C_\gamma$.

Now observe that if a, b, x and y are four real numbers, then $a \wedge x + b \wedge y \le (a+b) \wedge (x+y)$. In particular, for all $\varepsilon \in (0,1)$, $1 \wedge (\|\Lambda\|/R^{1/\gamma}) \ge (1-\varepsilon) \wedge (\Lambda_1/R^{1/\gamma}) + \varepsilon \wedge \big([\|\Lambda\| - \Lambda_1] \,/\, R^{1/\gamma}\big)$. Moreover,

$$\begin{aligned} \lim_{R\to\infty} R\,\mathbb{E}\left[(1-\varepsilon) \wedge \frac{\Lambda_1}{R^{1/\gamma}}\right] &= \lim_{R\to\infty} (1-\varepsilon)\, R\,\mathbb{E}\left[1 \wedge \frac{\Lambda_1}{[(1-\varepsilon)^\gamma R]^{1/\gamma}}\right] \\ &= (1-\varepsilon)^{1-\gamma} \left(\lim_{S\to\infty} S\,\mathbb{E}\left[1 \wedge \frac{\Lambda_1}{S^{1/\gamma}}\right]\right) = (1-\varepsilon)^{1-\gamma}\, C_\gamma \end{aligned}$$

where we have taken $S = (1-\varepsilon)^\gamma R$. Similarly, for $S = \varepsilon^\gamma R$,

$$\limsup_{R\to\infty} R\,\mathbb{E}\left[\varepsilon \wedge \frac{\|\Lambda\| - \Lambda_1}{R^{1/\gamma}}\right] = \varepsilon^{1-\gamma}\left(\limsup_{S\to\infty} S\,\mathbb{E}\left[1 \wedge \frac{\|\Lambda\| - \Lambda_1}{S^{1/\gamma}}\right]\right).$$

Therefore,

$$\limsup_{R\to\infty} R\,\mathbb{E}\left[1 \wedge \frac{\|\Lambda\| - \Lambda_1}{R^{1/\gamma}}\right] \leq \inf_{\varepsilon\in(0,1)} \frac{C_\gamma - (1-\varepsilon)^{1-\gamma}\, C_\gamma}{\varepsilon^{1-\gamma}} = 0.$$

Let $f : \mathcal{S}^\downarrow \to \mathbb{R}_+$ be a Lipschitz-continuous function bounded by 1 and set $g(x) := f(x,0,0,\dots)$ for all $x \geq 0$. There exists a constant $K \geq 0$ such that for all $\mathbf{x}$ and $\mathbf{y}$ in $\mathcal{S}^\downarrow$, $|f(\mathbf{x}) - f(\mathbf{y})| \leq 1 \wedge (K\,\|\mathbf{x}-\mathbf{y}\|)$. Therefore

$$\left| R\,\mathbb{E}\left[\left(1\wedge\frac{\|\Lambda\|}{R^{1/\gamma}}\right) f\left(\frac{\Lambda}{R^{1/\gamma}}\right) - \left(1\wedge\frac{\|\Lambda\|}{R^{1/\gamma}}\right) g\left(\frac{\|\Lambda\|}{R^{1/\gamma}}\right)\right]\right| \leq R\,\mathbb{E}\left[1\wedge\frac{2K\,(\|\Lambda\|-\Lambda_1)}{R^{1/\gamma}}\right] \xrightarrow[R\to\infty]{} 0.$$

Used jointly with our assumption on $\|\Lambda\|$ and Lemma 4.18, this ensures that $R\,\mathbb{E}\big[(1\wedge \|\Lambda\|/R^{1/\gamma})\, f(\Lambda/R^{1/\gamma})\big]$ converges to $\int (1\wedge \|\|s)\, f(\mathbf{s})\, c\, I_\gamma^{\mathrm{un}}(\mathrm{d}\mathbf{s})$ as $R\to\infty$. Lemma 3.21 concludes this proof. $\square$

5 Applications

In this section, we will develop applications of our three main results (Theorems 2.9, 4.2 and Proposition 4.3) to various models of random trees which satisfy the Markov branching property. With our unified approach, we will recover known results and get new ones.

5.1 Galton-Watson trees

Let ξ be a probability measure on $\mathbb{Z}_+$ with mean 1 and $\xi(1) < 1$ (*critical regime*). We will be interested in *unordered* Galton-Watson trees with offspring ditribution ξ, the law of which we will write GW_ξ. For any finite tree $\mathtt{t}$,

$$\mathrm{GW}_\xi(\mathtt{t}) := \sum_{\mathtt{t}'\in\mathtt{T}^{\mathrm{ord}}\,:\,\mathtt{t}'\sim\mathtt{t}} \prod_{u\in\mathtt{t}'} \xi\big[c_u(\mathtt{t}')\big].$$

For each positive integer n such that $\mathrm{GW}_\xi(\mathtt{T}_n) > 0$, let GW_ξ^n be the measure GW_ξ conditioned on the set $\mathtt{T}_n$ of trees with n vertices. Similarly, if n satisfies $\mathrm{GW}_\xi(\mathtt{T}_{\mathcal{L},n}) > 0$, define $\mathrm{GW}_\xi^{\mathcal{L},n}$ as GW_ξ conditioned on the set $\mathtt{T}_{\mathcal{L},n}$ of trees with n leaves. Moreover, let $d := \gcd\{n-1; \mathrm{GW}_\xi(\mathtt{T}_n) > 0\}$ and $d_{\mathcal{L}} := \gcd\{n-1; \mathrm{GW}_\xi(\mathtt{T}_{\mathcal{L},n}) > 0\}$.

Kesten's tree Let $\hat\xi$ be the *size-biased distribution* of ξ, that is $\hat\xi(k) = k\xi(k)$ for all $k \geq 0$. By assumption, the mean of ξ is 1, so $\hat\xi$ is a probability measure. We define GW_ξ^∞ as the distribution of *Kesten's tree* which is obtained as follows:

- Let $(X_n)_{n\geq 0}$ be a sequence of i.i.d. random variables such that $X_n + 1$ follows $\hat\xi$,
- Independently of this sequence, let $(T_{n,k}; n\geq 0, k\geq 1)$ be i.i.d. GW_ξ trees,
- For each $n \geq 0$, let $T_n := [\![T_{n,1},\dots,T_{n,X_n}]\!]$,
- For all $n \geq 0$, graft T_n on an infinite branch at height n respectively, i.e. set $T := \mathtt{b}_\infty \bigotimes_{n\geq 0} (\mathtt{v}_n, T_n)$ and denote its distribution by GW_ξ^∞.

Remark 5.1. These infinite trees were first indirectly introduced in [37] by Kesten who studied the genealogy of Galton-Watson processes conditioned to hit 0 after a large time. This result entails that if T is a GW_ξ tree, the conditional distribution of T on $|T| \geq n$ converges to GW_ξ^∞ as $n \to \infty$. Kesten's tree can thus be, in a way, considered as a GW_ξ tree conditioned to have infinite height.

This tree also appears as the local limit of conditioned critical Galton-Watson trees under various types of conditioning, see [2]. In particular, it was first proved in [36] (in terms of Galton-Watson processes) and in [7] (in terms of trees) that if ξ is critical and has finite variance, then $\mathrm{GW}_\xi^n \Rightarrow \mathrm{GW}_\xi^\infty$. In [20], it was shown that under the same assumptions, $\mathrm{GW}_\xi^{\mathcal{L},n} \Rightarrow \mathrm{GW}_\xi^\infty$. In both cases, the finite variance assumption may be dropped, see [33] and [2].

The local limits of Galton-Watson trees conditioned on their size with offspring distribution with means less than 1 were studied in [34], [33] and [1]. See also [50] for the study of the local limits of multi-type critical Galton-Watson trees.

Using Theorem 2.9, we will recover the following proposition in Section 5.1.1.

Proposition 5.2. *In the sense of the* $\mathrm{d_{loc}}$ *topology,* GW_ξ^n *and* $\mathrm{GW}_\xi^{\mathcal{L},n}$ *both converge weakly towards* GW_ξ^∞.

Afterwards, we will study scaling limits of Kesten's tree in the spirit of Theorem 4.2. Recall the descriptions of the Brownian tree with immigration and α-stable Lévy trees with immigration from Section 3.2.2.

Proposition 5.3. *Let* T *be a tree with distribution* GW_ξ^∞ *and define* $\mu_T := \sum_{u\in T}\delta_u$ *and* $\mu_T^{\mathcal{L}} := \sum_{u\in\mathcal{L}(T)}\delta_u$ *the counting measures on the set of its vertices and leaves respectively.*

(i) Finite variance: *Suppose* ξ *has finite variance* σ^2 *and that* $d=1$*. Then, with respect to the* $\mathrm{D_{GHP}}$ *topology,*

$$\left(\frac{T}{R}, \frac{\mu_T}{R^2}\right) \xrightarrow[R\to\infty]{\text{(d)}} \left(\mathcal{T}_B, \frac{\sigma^2}{4}\mu_B\right)$$

where $(\mathcal{T}_B, \mu_B)$ *is the Brownian tree with immigration.*

(i') *If* ξ *has finite variance* σ^2 *and if* $d_{\mathcal{L}} = 1$*, then*

$$\left(\frac{T}{R}, \frac{\mu_T^{\mathcal{L}}}{R^2}\right) \xrightarrow[R\to\infty]{\text{(d)}} \left(\mathcal{T}_B, \frac{\sigma^2\,\xi(0)}{4}\mu_B\right).$$

(ii) Stable case: *Suppose that* $\xi(n) \sim c\,n^{-1-\alpha}$ *as* $n\to\infty$ *for some positive constant* c *and* $\alpha\in(1,2)$*. Then,*

$$\left(\frac{T}{R}, \frac{\mu_T}{R^{\alpha/(\alpha-1)}}\right) \xrightarrow[R\to\infty]{\text{(d)}} \left(\mathcal{T}_\alpha, (ck_\alpha)^{1/(\alpha-1)}\mu_\alpha\right)$$

where $(\mathcal{T}_\alpha, \mu_\alpha)$ *is the* α*-stable immigration Lévy tree and* $k_\alpha = \Gamma(2-\alpha)/[\alpha\,(\alpha-1)]$.

Remark 5.4. Both (i) and (ii) were proved in [22] and (i') seems to be a new, if predictable, result.

We also mention that under the assumptions of (ii), $(T/R, \mu_T^{\mathcal{L}}/R^{\alpha/(\alpha-1)})$ should converge in distribution to $\left(\mathcal{T}_\alpha, (ck_\alpha)^{1/(\alpha-1)}\,\xi(0)\mu_\alpha\right)$. We won't prove this statement as Assumption (S) hasn't been proved in this case and to do so would require quite a bit of computation. The scaling limits of Galton-Watson trees with such an offspring distribution conditioned on their number of leaves were however studied in [39].

Section 5.1.2 will focus on the finite variance case, first on (i) and then on (i'). We will prove Proposition 5.3 in the stable case (ii) in Section 5.1.3.

5.1.1 Markov branching property and local limits

Let $\mathcal{N} := \{n \geq 1 \,:\, \mathrm{GW}_\xi(\mathtt{T}_n) > 0\}$. Proposition 37 in [30] states that the sequence of probability measures $(\mathrm{GW}_\xi^n)_{n\in\mathcal{N}}$ satisfies the Markov branching property, i.e. we have $\mathrm{GW}_\xi^n = \mathrm{MB}_n^q$ for all adequate n with q_{n-1} defined for all $\lambda = (\lambda_1, \ldots, \lambda_p)$ in $\mathcal{P}_{n-1}$ by

$$q_{n-1}(\lambda) = \frac{p!\xi(p)}{\prod_{j\geq 1} m_j(\lambda)!} \frac{\prod_{i=1}^p \mathbb{P}[\#T = \lambda_i]}{\mathbb{P}[\#T = n]}$$

where T is a GW_ξ tree.

Similarly, if we let $\mathcal{N}_\mathcal{L} := \{n \geq 1 \,:\, \mathrm{GW}_\xi(\mathtt{T}_{\mathcal{L},n}) > 0\}$, then in light of [46, Lemma 8], the family $(\mathrm{GW}_\xi^{\mathcal{L},n})_{n\in\mathcal{N}_\mathcal{L}}$ of probability measures satisfies the Markov branching property and the associated sequence $q^\mathcal{L}$ of first-split distributions such that $\mathrm{GW}_\xi^{\mathcal{L},n} = \mathrm{MB}_n^{\mathcal{L},q^\mathcal{L}}$ is given for all n in $\mathcal{N}_\mathcal{L}$ and $\lambda = (\lambda_1, \ldots, \lambda_p)$ in $\mathcal{P}_n$ by

$$q_n^\mathcal{L}(\lambda) = \frac{p!\xi(p)}{\prod_{j\geq 1} m_j(\lambda)!} \frac{\prod_{i=1}^p \mathbb{P}[\#_\mathcal{L} T = \lambda_i]}{\mathbb{P}[\#T_\mathcal{L} = n]}$$

where T still denotes a GW_ξ tree.

A Kesten tree with distribution GW_ξ^∞ can be seen as an infinite Markov branching tree with distribution $\mathrm{MB}_\infty^{q,q_\infty}$ where q_∞ is defined for any $\lambda = (\lambda_2, \ldots, \lambda_p)$ in $\mathcal{P}_{<\infty}$ by

$$q_\infty(\infty, \lambda) := \hat{\xi}(p) \, \frac{(p-1)!}{\prod_{j\geq 1} m_j(\lambda)!} \prod_{i=2}^p \mathbb{P}[\#T = \lambda_i].$$

The distribution of Kesten's tree may also be rewritten as $\mathrm{GW}_\xi^\infty = \mathrm{MB}_\infty^{\mathcal{L},q^\mathcal{L},q_\infty^\mathcal{L}}$ where $q_\infty^\mathcal{L}$ is given for all $\lambda \in \mathcal{P}_{<\infty}$ by

$$q_\infty^\mathcal{L}(\infty, \lambda) = \hat{\xi}(p) \, \frac{(p-1)!}{\prod_{j\geq 1} m_j(\lambda)!} \prod_{i=2}^p \mathbb{P}[\#_\mathcal{L} T = \lambda_i].$$

Proposition 5.2 is a direct consequence of the following results from Sections 4.3 and 4.4 in [2] used alongside Theorem 2.9.

Lemma 5.5. *If T is a GW_ξ tree, then*

$$\frac{\mathbb{P}[\#T = (n+1)d + 1]}{\mathbb{P}[\#T = nd + 1]} \xrightarrow[n\to\infty]{} 1 \quad \text{and} \quad \frac{\mathbb{P}[\#_\mathcal{L} T = (n+1)d_\mathcal{L} + 1]}{\mathbb{P}[\#_\mathcal{L} T = nd_\mathcal{L} + 1]} \xrightarrow[n\to\infty]{} 1.$$

Proof of Proposition 5.2. Let $\lambda = (\lambda_2, \ldots, \lambda_p)$ be an element of $\mathcal{P}_{<\infty}$. If there exists $2 \leq i \leq p$ such that $\lambda_i - 1$ isn't divisible by d, then for all $n \in \mathcal{N}$, $q_{n-1}(n - 1 - \|\lambda\|, \lambda) = 0 = q_\infty(\infty, \lambda)$. Otherwise, for $n \in \mathcal{N}$ large enough, in light of Lemma 5.5

$$\begin{aligned} q_{n-1}\big(n - 1 - \|\lambda\|, \lambda\big) &= \frac{p!\xi(p)}{\prod_{j\geq 1} m_j(\lambda)!} \frac{\mathbb{P}[\#T = n - \|\lambda\|]}{\mathbb{P}[\#T = n]} \prod_{i=1}^p \mathbb{P}[\#T = \lambda_i] \\ &\xrightarrow[n\to\infty]{} \hat{\xi}(p) \, \frac{(p-1)!}{\prod_{j\geq 1} m_j(\lambda)!} \prod_{i=2}^p \mathbb{P}[\#T = \lambda_i] = q_\infty(\infty, \lambda). \end{aligned}$$

Similarly, as n goes to infinity, $q_n^\mathcal{L}(n - \|\lambda\|, \lambda) \to q_\infty^\mathcal{L}(\infty, \lambda)$. Since these hold for any λ in $\mathcal{P}_{<\infty}$, we end this proof by using Corollary 2.10. $\square$

5.1.2 Scaling limits, finite variance

In the remainder of this section, $(T_i)_{i\geq 1}$ will denote i.i.d. Galton-Watson trees with offspring distribution ξ, $(Y_n)_{n\geq 1}$, i.i.d. ξ distributed random variables and for all $n \geq 1$,

$S_n := Y_1 + \cdots + Y_n - n$. We will also consider N, a random variable independent of both $(T_i)_i$ and $(Y_n)_n$ and such that $N+1$ follows $\hat{\xi}$.

The following so called *Otter-Dwass' formula* or *cyclic lemma* (see [44, Chapter 6] for instance) will be the cornerstone of many forthcoming computations.

Lemma 5.6 (Otter-Dwass' formula)**.** *With these notations, for all $k \geq 1$ and $n \geq 1$,*

$$\mathbb{P}\big[\#T_1 + \cdots + \#T_k = n\big] = \frac{k}{n}\mathbb{P}\big[S_n = -k\big].$$

Let q_* be the probability distribution on $\mathcal{P}_{<\infty}$ defined by $q_* = q_\infty(\infty, \cdot)$. Let Λ follow q_* and recall that it has the same distribution as $(\#T_1, \ldots, \#T_N)^\downarrow$.

In this paragraph, we'll assume that the variance σ^2 of ξ is finite and that $d = 1$. Recall that the Brownian tree with immigration is a $(1/2, \nu_B, I_B)$-fragmentation tree with immigration. It was proved in [30, Section 5.1] that Assumption (S) of Theorem 4.2 is fulfilled for $\gamma = 1/2$ and $\nu = \sigma/2 \cdot \nu_B$. To prove Proposition 5.3, it will therefore be sufficient to show that Assumption (I) is satisfied for $\gamma = 1/2$ and $I = \sigma/2 \cdot I_B$. For all $R \geq 1$, let $q^{(R)}$ be the distribution of Λ/R^2.

Proposition 5.7. *In the sense of weak convergence of finite measures on $\mathcal{S}^\downarrow$, $R\,(1 \wedge \|\|s)\, q^{(R)}(\mathrm{d}s)$ converges as R goes to infinity toward $(1 \wedge \|\|s)\,\sigma/2 \cdot I_B(\mathrm{d}s)$.*

Since I_B is unary, in order to prove Proposition 5.7, it will be enough to show that Λ satisfies the assumptions of Proposition 4.19. The next two lemmas will prove that both are met.

Lemma 5.8. *When n goes to infinity, $n^{3/2}\mathbb{P}[\|\Lambda\| = n] \to (\sigma^2/2\pi)^{1/2}$.*

Proof. In light of Otter-Dwass' formula, for all $n \geq 1$,

$$\begin{aligned} n^{3/2}\mathbb{P}\big[\|\Lambda\| = n\big] &= n^{3/2}\textstyle\sum_{k\geq 1}\mathbb{P}\big[\#T_1 + \cdots + \#T_k = n \,|\, N = k\big]\,\mathbb{P}\big[N = k\big] \\ &= \textstyle\sum_{k\geq 1} k\hat{\xi}(k+1)\, n^{1/2}\,\mathbb{P}\big[S_n = -k\big]. \end{aligned}$$

Recall the local central limit theorem in the finite variance case:

$$\sup\nolimits_{k\in\mathbb{Z}}\Big| n^{1/2}\,\mathbb{P}[S_n = k] - (2\pi\sigma^2)^{-1/2}\,\mathrm{e}^{-k^2/2n\sigma^2}\Big| \xrightarrow[n\to\infty]{} 0.$$

As a result, there exists a finite constant C such that $n^{1/2}\mathbb{P}[S_n = -k] \leq C$ for all $n \geq 1$ and $k \geq 1$ and if $k \geq 1$ is fixed, $n^{1/2}\mathbb{P}[S_n = -k] \to (2\pi\sigma^2)^{-1/2}$. Furthermore, $\sum_{k\geq 1} k\hat{\xi}(k+1) = \sigma^2$ so Lebesgue's dominated convergence theorem yields

$$\lim_{n\to\infty} n^{3/2}\mathbb{P}[\|\Lambda\| = n] = \sum_{k\geq 1} k\hat{\xi}(k+1)\,\Big(\lim\nolimits_{n\to\infty} n^{1/2}\mathbb{P}[S_n = -k]\Big) = \big(\sigma^2/2\pi\big)^{1/2}. \qquad \square$$

Lemma 5.9. *When $n \to \infty$, $n^{1/2}\mathbb{P}[\Lambda_1 \geq n]$ converges to $(2\sigma^2/\pi)^{1/2}$.*

Proof. Observe that for all $n \geq 0$, the event $\{\Lambda_1 \geq n\}$ has the same probability as $\{N \geq 1, \exists i \leq N : \#T_i \geq n\}$. Therefore $\mathbb{P}[\Lambda_1 \geq n] = \sum_{k\geq 1}\hat{\xi}(k+1)\big(1 - \mathbb{P}[\#T_1 < n]^k\big)$. Let G be the generating function of ξ, i.e. $G(s) = \sum_{k\geq 0}\xi(k)\,s^k$ for all $s \in [0,1]$. This function is twice-differentiable on $[0,1]$ and we may write $\mathbb{P}[\Lambda_1 \geq n] = G'(1) - G'\big(1 - \mathbb{P}[\#T_1 \geq n]\big)$.

For all $n \geq 1$, Otter-Dwass' formula gives $n^{1/2}\mathbb{P}[\#T_1 \geq n] = n^{1/2}\sum_{m\geq n} m^{-1}\,\mathbb{P}[S_m = -1]$. The local central limit theorem ensures that $m^{1/2}\mathbb{P}[S_m = -1] \to (2\pi\sigma^2)^{-1/2}$ as $m \to \infty$. Therefore, for all positive ε and n large enough,

$$n^{1/2}\,\Big|\mathbb{P}[\#T_1 \geq n] - \textstyle\sum_{m\geq n} m^{-3/2}(2\pi\sigma^2)^{-1/2}\Big| \leq n^{1/2}\displaystyle\sum_{m\geq n} m^{-3/2}\,\varepsilon \xrightarrow[n\to\infty]{} 2\varepsilon.$$

Incidentally, $n^{1/2}\mathbb{P}[\#T_1 \geq n]$ and $n^{1/2}\sum_{m\geq n} m^{-3/2}(2\pi\sigma^2)^{-1/2}$ have the same limit when $n \to \infty$ which is to say that $n^{1/2}\mathbb{P}[\#T_1 \geq n] \to (2/\pi\sigma^2)^{1/2}$ as $n \to \infty$. As a result,

$$n^{1/2}\mathbb{P}[\Lambda_1 \geq n] = n^{1/2}\Big[G'(1) - G'\big(1 - \mathbb{P}[\#T_1 \geq n]\big)\Big] \xrightarrow[n\to\infty]{} \left(\frac{2}{\pi\sigma^2}\right)^{1/2} G''(1) = \left(\frac{2\sigma^2}{\pi}\right)^{1/2}.$$

□

Lemmas 5.8, 5.9 and Proposition 4.19 prove Proposition 5.7. Theorem 4.2 therefore implies that $(T/R, \mu_T/R^2)$ converges in distribution to a $(1/2, \sigma/2 \cdot \nu_B, \sigma/2 \cdot I_B)$ fragmentation tree with immigration. Using Proposition 3.18, we may restate this last result as Proposition 5.3 (i). Furthermore, as a result of Proposition 4.3, we get that in particular, $\mu_T(T|_R)/R^2$ converges in distribution to $(\sigma^2/4)\,\mu_{\mathcal{T}_B}(\mathcal{T}_B|_1)$ or equivalently to $\mu_{\mathcal{T}_B}(\mathcal{T}_B|_{\sigma/2})$.

We will now prove Proposition 5.3 (i'). Assume that $d_{\mathcal{L}} = 1$. Theorem 7 in [46] proves that the family $(q_n^{\mathcal{L}})_n$ of first split distributions associated to Galton-Watson trees conditioned on their number of leaves satisfies Assumption $(\mathtt{S})$: $n^{1/2}\,(1-s_1)\,\bar{q}_n^{\mathcal{L}} \Rightarrow \sigma\,\xi(0)^{1/2}/2 \cdot (1-s_1)\,\nu_B(\mathrm{d}\mathbf{s})$. As a result, we only need to prove Assumption $(\mathtt{I})$ for $\gamma = 1/2$ and $I = \sigma\,\xi(0)^{1/2}/2 \cdot I_B$.

Proof of Proposition 5.3 (i'). Theorem 6 in [46] states that there exists a critical probability distribution ζ on $\mathbb{Z}_+$ such that $\#_{\mathcal{L}}T_1$, the number of leaves of T_1, has the same distribution as $\#\tau$, where τ follows GW_ζ. Lemma 6 further states that if ξ has finite variance σ^2, then ζ has variance $\sigma^2/\xi(0)$.

Let $\Lambda^{\mathcal{L}}$ be such that (∞, Λ) is distributed according to $q_\infty^{\mathcal{L}}$. The random partition $\Lambda^{\mathcal{L}}$ is distributed like $(\#_{\mathcal{L}}T_1, \ldots, \#_{\mathcal{L}}T_N)^{\downarrow}$, or equivalently, like $(\#\tau_1, \ldots, \#\tau_N)^{\downarrow}$, where $(\tau_n)_{n\geq 1}$ are i.i.d. GW_ζ trees independent of N. Therefore, if $(V_n)_{n\geq 1}$ is a sequence of i.i.d. ζ-distributed random variables and if $Z_n := V_1 + \cdots + V_n - n$, proceeding as in the proof of Lemma 5.8 gives:

$$n^{3/2}\mathbb{P}[\|\Lambda^{\mathcal{L}}\| = n] = \textstyle\sum_{k\geq 0} k\,\hat{\xi}(k+1) n^{1/2}\mathbb{P}[Z_n = -k] \xrightarrow[n\to\infty]{} \big[\sigma^2\xi(0)/(2\pi)\big]^{1/2}.$$

Similarly, the same kind of computations as in Lemma 5.9 yields

$$n^{1/2}\mathbb{P}[\Lambda_1^{\mathcal{L}} \geq n] = n^{1/2}\Big[G'(1) - G'\big(1 - \mathbb{P}[\#\tau_1 \geq n]\big)\Big] \xrightarrow[n\to\infty]{} \big[2\sigma^2\xi(0)/\pi\big]^{1/2}$$

where G still denotes the generating function of ξ. As a result, because of Theorem 4.2 and Proposition 4.19, when $R \to \infty$, $(T/R, \mu_T^{\mathcal{L}}/R^2)$ converges in distribution to a $(1/2, \sigma\xi(0)^{1/2}/2 \cdot \nu_B, \sigma\xi(0)^{1/2}/2 \cdot I_B)$ fragmentaion tree with immigration. Proposition 3.18 then allows us to conlude. □

5.1.3 Scaling limits, stable case

In this paragraph, we'll suppose that there exist $\alpha \in (1,2)$ and a positive constant c such that $n^{1+\alpha}\xi(n) \to c$ when $n \to \infty$.

Recall that Λ denotes a q_*-distributed variable and has the same distribution as $(\#T_1, \ldots, \#T_N)^{\downarrow}$ where $N+1$ is distributed according to $\hat{\xi}$ and is independent of the sequence $(T_n)_{n\geq 1}$ of i.i.d. GW_ξ trees. Moreover, we will use the notations introduced to define ν_α and $I^{(\alpha)}$ in Sections 3.2.1 and 3.2.2: $(\Sigma_t; t \geq 0)$ will denote a $1/\alpha$-stable subordinator with Laplace exponent $\lambda \mapsto -\log \mathbb{E}[\exp(-\lambda\Sigma_1)] = \lambda^{1/\alpha}$ and Δ will be the decreasing rearrangement of its jumps on $[0,1]$.

It was proved in [30, Section 5.2] that the family $q = (q_n)_{n\in\mathcal{N}}$ of first-split distributions associated to $(\mathrm{GW}_\xi^n)_{n\in\mathcal{N}}$ satisfies Assumption $(\mathtt{S})$ of Theorem 4.2 for $\gamma = 1 - 1/\alpha$ and $\nu = (c\,k_\alpha)^{1/\alpha} \cdot \nu_\alpha$. Proposition 5.3 (ii) will therefore be a consequence of the next proposition. For all $R \geq 1$, write $q^{(R)}$ for the distribution of $R^{-\alpha/(\alpha-1)}\Lambda$.

Proposition 5.10. *When $R \to \infty$, $R\,(1 \wedge \|\|s)\, q^{(R)}(\mathrm{d}\mathbf{s})$ converges weakly to $(c\,k_\alpha)^{1/\alpha}(1 \wedge \|\|s)\, I^{(\alpha)}(\mathrm{d}\mathbf{s})$.*

Proof. As shown in [30, Section 5.2], $n^{1+1/\alpha}\,\mathbb{P}[\#T_1 = n]$ converges to $[(c\,k_\alpha)^{1/\alpha}\alpha\,\Gamma(1-1/\alpha)]^{-1}$. Therefore, $(\#T_n)_{n\geq 1}$ lies in the domain of attraction of a $1/\alpha$-stable distribution. More accurately, in the Skorokhod topology,

$$\left(\frac{\#T_1 + \cdots + \#T_{\lfloor nt\rfloor}}{n^\alpha}\,;\, t \geq 0\right) \xrightarrow[n\to\infty]{\text{(d)}} \frac{1}{c\,k_\alpha}\Big(\Sigma_t\,;\, t \geq 0\Big).$$

This, in conjunction with Skorokhod's representation theorem, implies that there exists a sequence $(X_n)_{n\geq 0}$, where for all $n \geq 1$,

$$X_n \overset{\text{(d)}}{=} \frac{c\,k_\alpha}{n^\alpha}\big(\#T_1, \dots, \#T_n, 0, 0, \dots\big)^{\downarrow},$$

which a.s. converges to (a version of) Δ.

Let $F : \mathcal{S}^{\downarrow} \to \mathbb{R}_+$ be a Lipschitz continuous function such that $F(\mathbf{s}) \leq 1 \wedge \|\|s$ and set $f : \mathbb{R}_+ \to \mathbb{R}_+$, $t \mapsto \mathbb{E}[F(t^\alpha/(c\,k_\alpha)\cdot\Delta)]$. The dominated convergence theorem ensures that the function f is continuous. It is clearly bounded by 1 and

$$f(t) \leq \mathbb{E}\Big[1 \wedge \big(t^\alpha/(c\,k_\alpha)\cdot\|\Delta\|\big)\Big] = \mathbb{E}\Big[1 \wedge \Sigma_{(c\,k_\alpha)^{-1/\alpha}t}\Big] \leq \frac{t}{(c\,k_\alpha)^{1/\alpha}}\int_{\mathbb{R}_+}(1\wedge x)\,\Pi_{1/\alpha}(\mathrm{d}x).$$

Since $n^\alpha\mathbb{P}[N = n] \to c$, Lemma 4.18 ensures that when $R \to \infty$, $R\,\mathbb{E}\big[f(N/R^{1/(\alpha-1)})\big]$ converges to $c\int_0^\infty t^{-\alpha} f(t)\,\mathrm{d}t = (c\,k_\alpha)^{1/\alpha}\int F\,\mathrm{d}I^{(\alpha)}$. Furthermore, because Λ is distributed like $(c\,k_\alpha)^{-1} N^\alpha X_N$,

$$\left|R\,\mathbb{E}\left[F\left(\frac{\Lambda}{R^{\alpha/(\alpha-1)}}\right) - f\left(\frac{N}{R^{1/(\alpha-1)}}\right)\right]\right| \leq R\,\mathbb{E}\left[1 \wedge \left(K\left(\frac{N}{R^{1/(\alpha-1)}}\right)^{\alpha}\|X_N - \Delta\|\right)\right]$$

where $K\cdot(c\,k_\alpha)$ is bigger than the Lipschitz constant of F. We will now endeavour to prove that this last quantity goes to 0 when $R \to \infty$. For all $\mathbf{s}$ in $\mathcal{S}^{\downarrow}$, let $\mathbf{s} \wedge 1$ be the sequence $(s_i \wedge 1)_{i\geq 1}$. Then for any $\mathbf{x}$ and $\mathbf{y}$ in $\mathcal{S}^{\downarrow}$, we may write $\|\mathbf{x} - \mathbf{y}\| = \|\mathbf{x}\wedge 1 - \mathbf{y}\wedge 1\| + \|(\mathbf{x} - \mathbf{x}\wedge 1) - (\mathbf{y} - \mathbf{y}\wedge 1)\|$.

In light of Lemma 4.18, $n\,\mathbb{E}\big[1 \wedge (\#T_1/n^\alpha)\big]$ converges to $[(c\,k_\alpha)^{1/\alpha}\,\Gamma(2-1/\alpha)]^{-1}$. It ensues from the i.i.d. nature of the sequence $(\#T_i)_{i\geq 1}$ that

$$\sup_{n\geq 1}\mathbb{E}\big[\|X_n \wedge 1\|^2\big] = \sup_{n\geq 1}\left(n\,\mathbb{E}\left[\left(\frac{\#T_1}{n^\alpha}\wedge 1\right)^2\right] + n(n-1)\,\mathbb{E}\left[\frac{\#T_1}{n^\alpha}\wedge 1\right]^2\right) < \infty.$$

Fatou's lemma (or classical results on Poisson Point Process, see [38, Section 3.2]) ensures that $\mathbb{E}[\|\Delta\wedge 1\|^2]$ is also finite. As a result, the sequence $(\|X_n\wedge 1 - \Delta\wedge 1\|)_{n\geq 1}$ is bounded in L^2. Since $\|X_n \wedge 1 - \Delta\wedge 1\| \to 0$ a.s., we also have $\mathbb{E}\big[\|X_n\wedge 1 - \Delta\wedge 1\|\big] \to 0$.

If $\beta < 1/\alpha$, then $\mathbb{E}[\|\Delta - \Delta\wedge 1\|^\beta] \leq \mathbb{E}[\|\Delta\|^\beta] = \mathbb{E}[\Sigma_1^\beta] < \infty$. Moreover, since it converges, the sequence $\big(m^{1+1/\alpha}\mathbb{P}[\#T_1 = m]\big)_m$ is bounded by a finite constant, say Q. Consequently,

$$\begin{aligned}\mathbb{E}\big[\|X_n - X_n\wedge 1\|^\beta\big] = n\,\mathbb{E}\left[\left(\frac{\#T_1}{n^\alpha} - 1\right)_+^\beta\right] &\leq Q\,n\sum_{k>n^\alpha}\frac{k^\beta}{n^{\alpha\beta}}\,\frac{1}{k^{1+1/\alpha}}\\ &\xrightarrow[n\to\infty]{} Q\int_1^\infty\frac{\mathrm{d}t}{t^{1+1/\alpha-\beta}} = \frac{\alpha\,Q}{1-\alpha\beta}\end{aligned}$$

which proves that the sequence $\left(\mathbb{E}[\|X_n - X_n \wedge 1\|^{\beta}]\right)_{n\geq 1}$ is bounded. Since this holds for all $\beta < 1/\alpha$, if ε is positive and such that $(1+\varepsilon)\beta =: \beta' < 1/\alpha$, then

$$\sup_{n\geq 1} \mathbb{E}\Big[\big(\|(X_n - X_n\wedge 1) - (\Delta - \Delta\wedge 1)\|^{\beta}\big)^{1+\varepsilon}\Big] \leq \sup_{n\geq 1} \mathbb{E}\Big[\|X_n - X_n\wedge 1\|^{\beta'} + \|\Delta - \Delta\wedge 1\|^{\beta'}\Big] < \infty.$$

Hence, the sequence $\left(\|(X_n - X_n \wedge 1) - (\Delta - \Delta \wedge 1)\|^{\beta}\right)_{n\geq 1}$ is bounded in $L^{1+\varepsilon}$. Because it converges to 0 almost surely, its mean also goes to 0 as n tends to infinity.

For all $\beta < 1/\alpha$ and $\varepsilon > 0$, there exist a finite constant C and a finite integer n_ε such that for all $n \geq 1$

$$\mathbb{E}\Big[\|X_n \wedge 1 - \Delta \wedge 1\|\Big] \vee \mathbb{E}\Big[\|(X_n - X_n \wedge 1) - (\Delta - \Delta \wedge 1)\|^{\beta}\Big] \leq \varepsilon + C\mathbb{1}_{n<n_\varepsilon}.$$

Using the same arguments as in the proof of Lemma 4.18 it is easy to prove that for any $\kappa > \alpha - 1$,

$$R\,\mathbb{E}\big[1 \wedge (N/R^{1/(\alpha-1)})^{\kappa}\big] \xrightarrow[R\to\infty]{} c\int_0^{\infty} \frac{1\wedge t^{\kappa}}{t^{\alpha}}\,\mathrm{d}t = \frac{c}{\kappa - (\alpha-1)} + \frac{c}{\alpha - 1}.$$

Consequently, if $\beta \in (1 - 1/\alpha, 1/\alpha)$, we get

$$\begin{aligned}
&\limsup_{R\to\infty} R\,\mathbb{E}\left[1 \wedge \left(K\left(\frac{N}{R^{1/(\alpha-1)}}\right)^{\alpha} \|X_N - \Delta\|\right)\right] \\
&\quad\leq \limsup_{R\to\infty} R\,\mathbb{E}\left[1 \wedge \left(K\left(\frac{N}{R^{1/(\alpha-1)}}\right)^{\alpha} \mathbb{E}\Big[\|X_N \wedge 1 - \Delta \wedge 1\| \;\Big|\; N\Big]\right)\right] \\
&\qquad\qquad + R\,\mathbb{E}\left[1 \wedge \left(K^{\beta}\left(\frac{N}{R^{1/(\alpha-1)}}\right)^{\alpha\beta} \mathbb{E}\Big[\|(X_N - X_N \wedge 1) - (\Delta - \Delta\wedge 1)\|^{\beta} \;\Big|\; N\Big]\right)\right] \\
&\quad\leq \limsup_{R\to\infty} R\,\mathbb{E}\left[1 \wedge \left(K\,\frac{N^{\alpha}}{R^{\alpha/(\alpha-1)}}\left(\varepsilon + C\mathbb{1}_{N<n_\varepsilon}\right)\right)\right] \\
&\qquad\qquad + R\,\mathbb{E}\left[1 \wedge \left(K^{\beta}\frac{N^{\alpha\beta}}{R^{\alpha\beta/(\alpha-1)}}\left(\varepsilon + C\mathbb{1}_{N<n_\varepsilon}\right)\right)\right] \\
&\quad\leq \limsup_{R\to\infty} \;\frac{K\,C\,n_\varepsilon^{\alpha}}{R^{\alpha/(\alpha-1)-1}} + \frac{K^{\beta}C\,n_\varepsilon^{\alpha\beta}}{R^{\alpha\beta/(\alpha-1)-1}} + K^{\alpha/(\alpha-1)}\varepsilon^{\alpha/(\alpha-1)}\,R\,\mathbb{E}\left[1 \wedge \frac{N^{\alpha}}{R^{\alpha/(\alpha-1)}}\right] \\
&\qquad\qquad + K^{\alpha/(\alpha-1)}\varepsilon^{[\alpha/(\alpha-1)]/\beta}\,R\,\mathbb{E}\left[1 \wedge \frac{N^{\alpha\beta}}{R^{\alpha\beta/(\alpha-1)}}\right] \\
&\quad= O(\varepsilon^{\alpha/(\alpha-1)}).
\end{aligned}$$

Since this holds for any positive ε, it follows that

$$R\,\mathbb{E}\left[1 \wedge \left(K\left(\frac{N}{R^{1/(\alpha-1)}}\right)^{\alpha} \|X_N - \Delta\|\right)\right] \xrightarrow[R\to\infty]{} 0,$$

which in turn proves that $R\,\mathbb{E}\big[F(\Lambda/R^{\alpha/(\alpha-1)})\big]$ indeed converges to $(c\,k_\alpha)^{1/\alpha}\int_{\mathcal{S}^{\downarrow}} F\,\mathrm{d}I^{(\alpha)}$. We conclude with Lemma 3.21. $\square$

5.2 Cut-trees

Let τ be a finite labelled tree. If τ is made out of a single vertex, let its cut-tree $\mathrm{Cut}\,(\tau)$ be the tree with a single vertex. Otherwise, define the cut-tree of τ as the (unordered) binary tree $\mathrm{Cut}\,(\tau)$ obtained by the following recursive process:

- Pick $a \to b$ uniformly at random among the edges of τ and remove that edge,
- Let τ_1 and τ_2 be the two sub-trees of τ formerly connected by $a \to b$,
- Define the cut-tree of τ as the concatenation of the cut-trees of τ_1 and τ_2, i.e. set $\mathrm{Cut}(\tau) := [\![\mathrm{Cut}(\tau_1), \mathrm{Cut}(\tau_2)]\!]$.

With this definition, if τ has n vertices, then $\mathrm{Cut}(\tau)$ has n leaves. The cut-tree of τ represents the genealogy of its dismantling when we remove edge after edge, until all have been deleted.

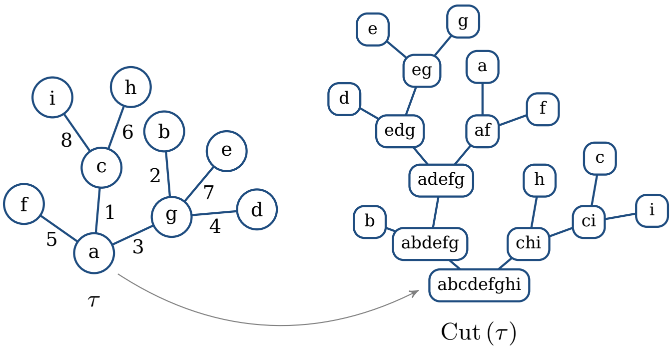

Figure 5: A labelled tree τ and its cut-tree
(the edges of τ are labelled in the order they are removed).

Cut-trees were introduced in [12] as a means of generalising the study of the number of cuts necessary to isolate a marked vertex or a finite number of marked vertices. In this section, we will study the local and scaling limits of two models of cut-trees, studied in [12] and [14], which both satisfy the Markov branching property. Also see [15] and [21] for the study of the cut-trees of conditioned Galton-Watson trees.

5.2.1 Cut-trees of Cayley trees

A *Cayley tree* of size $n \geq 1$ is a labelled tree τ_n chosen uniformly at random in the set of trees with n labelled vertices (for convenience, with labels 1 through n). It is well-known that, viewed as an unlabelled tree, τ_n has the same distribution as an unordered Galton-Watson tree with offspring law $\mathrm{Poisson}(1)$ conditioned to have n vertices. For all $n \geq 1$, let $T_n := \mathrm{Cut}(\tau_n)$ be the cut-tree of a Cayley tree with size n.

Let $(\vartheta_n)_{n\geq 0}$ be a sequence of i.i.d. unconditioned $\mathrm{GW}_{\mathrm{Poisson}(1)}$ trees. Let T_∞ be the tree obtained by attaching for each $n \geq 0$ the cut-tree of ϑ_n to the vertex of an infinite branch at height n by an edge. In other words, set $T_\infty := \mathtt{b}_\infty \bigotimes_{n\geq 0} \left(\mathtt{v}_n, [\![\mathrm{Cut}(\vartheta_n)]\!]\right)$.

The aim of this section will be to prove the next two results.

Proposition 5.11. *When $n \to \infty$, T_n converges to T_∞ in distribution with respect to the local limit topology.*

Proposition 5.12. *Endow T_∞ with counting measure on its leaves μ_∞. Then, as R goes to infinity, $(T_\infty/R, \mu_\infty/R^2)$ converges to $(\mathcal{T}_B, 1/2 \cdot \mu_B)$ in distribution with respect to the $\mathrm{D_{GHP}}$ topology, where $(\mathcal{T}_B, \mu_B)$ denotes the Brownian tree with immigration.*

Markov branching property It was stated in [12] that (T_n) satisfies the Markov branching property and more specifically, that the distribution of T_n is $\mathrm{MB}_n^{\mathcal{L},q}$ where the associated first-split distributions are given by $q_1(1)=1$, for all $n\geq 2$, $q_n(p\neq 2)=0$ and if $1\leq k<n/2$,

$$q_n(n-k,k)=\frac{(n-k)^{n-k-1}}{(n-k)!}\,\frac{k^{k-1}}{k!}\,\frac{(n-2)!}{n^{n-3}}.$$

The tree T_∞ can be described as an infinite Markov branching tree with distribution $\mathrm{MB}_\infty^{\mathcal{L},q,q_\infty}$ where the probability measure q_∞ is defined by $q_\infty(p\neq 2)=q_\infty(m_\infty\neq 1)=0$ and for all positive k, $q_\infty(\infty,k)=\mathbb{P}[\#\vartheta=k]$ where ϑ is a $\mathrm{GW}_{\mathrm{Poisson}\,(1)}$ tree. Recall that the size of ϑ has Borel distribution with parameter 1, therefore, for any positive k, $q_\infty(\infty,k)=k^{k-1}\,\mathrm{e}^{-k}/k!$.

Local limits For any $k\geq 1$, when $n\to\infty$, Stirling's approximation gives

$$q_n(n-k,k)\sim\frac{k^{k-1}\,\mathrm{e}^{2-k}}{k!}(1-2/n)^n\xrightarrow[n\to\infty]{}\frac{k^{k-1}\,\mathrm{e}^{-k}}{k!}=q_\infty(\infty,k).$$

We may then use Corollary 2.10 and thus prove Proposition 5.12.

Scaling limits Section 2.1 in [12] proves that $n^{1/2}\,(1-s_1)\,\bar{q}_n(\mathrm{d}s)$ converges weakly to $(1-s_1)\,1/2\cdot\nu_B(\mathrm{d}s)$ in the sense of finite measures on $\mathcal{S}^{\downarrow}_{\leq 1}$.

Moreover, q_∞ is a.s. binary, and Stirling's approximation ensures that $n^{3/2}q_\infty(\infty,n)$ converges to $(2\pi)^{-1/2}$ when n goes to infinity. Therefore, if Λ is such that (∞,Λ) follows q_∞ and if $q^{(R)}$ is the distribution of Λ/R^2, then Proposition 4.19 implies that $R\,(1\wedge\|\!|s)\,q^{(R)}(\mathrm{d}s)$ weakly converges to $(1\wedge\|\!|s)\,1/2\cdot I_B(\mathrm{d}s)$ as $R\to\infty$. In other words, Assumption $(\mathtt{I})$ is also satisfied.

Consequently, Theorem 4.2 ensures that when $R\to\infty$, $(T_\infty/R,\mu_\infty/R^2)$ converges in distribution to a $(1/2,1/2\cdot\nu_B,1/2\cdot I_B)$ fragmentation tree with immigration with respect to the topology induced by $\mathrm{D}_{\mathrm{GHP}}$. Proposition 3.18 then concludes the proof of Proposition 5.12.

5.2.2 Cut-trees of uniform recursive trees

A *recursive tree* with n vertices is a labelled tree (with labels 1 through n) such that the labels on the shortest path from 1 to any given leaf are increasing. For all $n\geq 1$, let τ_n denote a labelled tree chosen uniformly at random among the set of recursive trees with n vertices and call T_n its cut-tree.

Define a probability measure π on $\mathbb{N}$ by $\pi(n)=1/[n(n+1)]$ and let $(X_n,\vartheta_n)_{n\geq 0}$ be a sequence of i.i.d. variables, where for each n, X_n follows π and conditionally on $X_n=\ell$, ϑ_n is a recursive tree with ℓ vertices. Define T_∞ as the tree obtained by attaching the cut-tree of ϑ_n by an edge to an infinite branch at height n, i.e. set $T_\infty:=\mathtt{b}_\infty\bigotimes_{n\geq 0}\left(\mathtt{v}_n,[\![\mathrm{Cut}\,(\vartheta_n)]\!]\right)$.

Proposition 5.13. *In the sense of the local limit topology, T_n converges in distribution to T_∞ when $n\to\infty$.*

It was observed in [13] and [14] that the sequence $(T_n)_{n\geq 1}$ is Markov branching. Moreover, we may deduce from [13, Section 2] the expression of the respective distributions q_n of $\Lambda^{\mathcal{L}}(T_n)$. Clearly, $q_1(1)=1$, and for $n\geq 2$, if X denotes a random variable with distribution π, then for all $k\leq n/2$, $q_n(n-k,k)=\mathbb{P}[X=k|X<n]+\mathbb{P}[X=n-k|X<n]\,\mathbb{1}_{k\neq n/2}$. In particular,

$$q_n(n-k,k)=\begin{cases}\dfrac{n}{n-1}\left(\dfrac{1}{k(k+1)}+\dfrac{1}{(n-k)(n-k+1)}\right) & \text{if } k<n/2,\\ \dfrac{4}{(n-1)(n+2)} & \text{if } k=n/2.\end{cases}$$

The tree T_∞ may also be described as an infinite Markov branching tree with distribution $\mathrm{MB}_\infty^{\mathcal{L},q,q_\infty}$ where the measure q_∞ is given by $q_\infty(p \neq 2) = q_\infty(m_\infty \neq 1) = 0$ and for all $k \geq 1$, $q_\infty(\infty, k) = \pi(k)$.

If k is a fixed integer, then $q_n(n-k, k)$ clearly converges to $q_\infty(\infty, k)$. We conclude the proof of Proposition 5.13 with Corollary 2.10.

Remark 5.14. It was shown in [14] that $(n/\log n)^{-1} T_n$ converges to the real interval $[0,1]$ rooted at 0 and endowed with the Lebesgue measure. However, Assumption (S) doesn't hold.

5.3 The α-γ model

In this section, we will study trees generated according to the algorithm of the α-γ model described in [19]. This algorithm was introduced as an interpolation between various models of sequentially growing trees such as Rémy's algorithm [45], used to generate uniform binary trees with any number of leaves, Marchal's [42], which gives the n-dimensional marginal of Duquesne-Le Gall's stable trees (the discrete tree spanned by n leaves chosen uniformly at random in a stable tree), and Ford's α-model [24], used for instance in phylogeny.

Let $0 \leq \gamma \leq \alpha \leq 1$. Start with $T_1 := \{\varnothing\}$, the trivial tree, and $T_2 := \{\varnothing, (1), (2)\}$, a tree with two leaves attached to its root. Then for $n \geq 3$, conditionally on the tree T_{n-1}:

- Assign to each edge of T_{n-1} (considered as a planted tree, i.e. a tree in which a phantom edge has been attached under the root) the weight $1-\alpha$ if the edge ends with a leaf or γ otherwise,
- Also assign to each non-leaf vertex u the weight $[c_u(T_{n-1}) - 1]\alpha - \gamma$,
- Pick an edge or a vertex in T_{n-1} with probability proportional to these weights,
  - If an edge was picked, place a new vertex at its middle and attach a new leaf to it,
  - If a vertex was selected, attach a new leaf to it,

and let T_n be the tree thus obtained. We will also call $\mathrm{AG}^n_{\alpha,\gamma}$ its distribution for all $n \geq 1$ and $0 \leq \gamma \leq \alpha \leq 1$.

Remark 5.15. As mentioned at the beginning of this section, some particular choices of parameters give previously studied algorithms:

- When $\alpha = \gamma = 1/2$, we get Rémy's algorithm [45],
- If $\beta \in (1,2)$, taking $\alpha = 1/\beta$ and $\gamma = 1-\alpha$ gives Marchal's algorithm [42],
- When $\alpha = \gamma$, this algorithm coincides with that of Ford's α-model [24].

The Beta geometric distribution Fix θ in $(0,1)$. Let Π be a Beta random variable with parameters $(1-\theta, \theta)$, and conditionally on Π, let X have geometric distribution with parameter $1-\Pi$, meaning that $\mathbb{P}[X = n \,|\, \Pi] = \Pi^n(1-\Pi)$ for every integer $n \geq 0$. We say that X is a *beta geometric* variable of parameters $(\theta, 1-\theta)$. For all integers $n \geq 0$,

$$\mathbb{P}[X = n] = \mathbb{E}\big[\Pi^n(1-\Pi)\big] = \frac{1}{\mathrm{B}(1-\theta,\theta)} \int_0^1 x^{n-\theta}(1-x)^\theta \mathrm{d}x = \frac{\theta\,\Gamma(n+1-\theta)}{\Gamma(1-\theta)\,(n+1)!}.$$

We will also use the convention $X = 0$ a.s. if $\theta = 1$ and $X = \infty$ a.s. if $\theta = 0$.

Infinite α-γ tree Assume that $0 < \gamma \leq \alpha \leq 1$. Let $(X_n)_{n\geq 0}$ be a sequence of i.i.d. beta geometric random variables with parameters $(\gamma/\alpha, 1-\gamma/\alpha)$. Let $(Y_{n,k}, \tau_{n,k})$ be a sequence of i.i.d. variables independent of $(X_n)_n$ such that $Y_{n,k}$ is a $(\alpha, 1-\alpha)$ beta geometric variable and conditionally on $Y_{n,k} = \ell$, $\tau_{n,k}$ is an α-γ tree with $\ell+1$ leaves, i.e. $\tau_{n,k}$ follows $\mathrm{AG}^{\ell+1}_{\alpha,\gamma}$.

Finally, conditionally on $(X_n, Y_{n,k}, \tau_{n,k}; n \geq 0, k \geq 0)$, define T_∞ as the tree obtained by grafting for each $n \geq 0$ the concatenation of $\tau_{n,i}$, $0 \leq i \leq X_n$ at height n on an infinite branch. In other words,

$$T_\infty := \mathtt{b}_\infty \bigotimes\nolimits_{n\geq 0} \big(\mathtt{v}_n, [\![\tau_{n,0}, \dots \tau_{n,X_n}]\!]\big)$$

and denote by $\mathrm{AG}^\infty_{\alpha,\gamma}$ its distribution.

Remark 5.16. In Ford's α-model, i.e. when $\alpha = \gamma > 0$, $X_n = 0$ a.s. for all n, so a single tree is grafted at each height. Similarly, when $\alpha = 1$ and $0 < \gamma \leq \alpha$, $Y_{n,k} = 0$ a.s..

We will start our study of the α-γ model by proving this next proposition with the help of Theorem 2.9. Similar results for $\alpha = \gamma$ were already proved in [47] and in [18, Lemma 3.8] for any $0 < \gamma \leq \alpha \leq 1$.

Proposition 5.17. *For any $0 < \gamma \leq \alpha \leq 1$, the probability measure $\mathrm{AG}^n_{\alpha,\gamma}$ converges weakly to $\mathrm{AG}^\infty_{\alpha,\gamma}$ as n grows to ∞ in the sense of the local limit topology.*

We will then study the scaling limits of these infinite trees: Section 5.3.2 will focus on the case $0 < \gamma < \alpha < 1$ and Section 5.3.3, on $\alpha = \gamma$.

5.3.1 Markov branching property and local limits

Proposition 1 in [19] states that the sequence $(\mathrm{AG}^n_{\alpha,\gamma})_n$ satisfies the Markov branching property. Moreover, the sequence $q = (q_n)_n$ associated to the first split distributions of T_n, i.e. such that q_n is the law of $\Lambda^{\mathcal{L}}(T_n)$ for all $n \geq 1$, is given by $q_1(\varnothing) = 1$, and for any $n \geq 2$, for all $\lambda = (\lambda_1, \dots, \lambda_p) \in \mathcal{P}_n$,

$$q_n(\lambda) = \frac{1}{\prod_{j\geq 1} m_j(\lambda)} \left(\gamma + \frac{1-\alpha-\gamma}{n(n-1)} \sum_{i\neq j} \lambda_i \lambda_j\right) \times \frac{\Gamma(1-\alpha)\, n!}{\Gamma(n-\alpha)} \, \frac{\alpha^{p-2}\, \Gamma(p-1-\gamma/\alpha)}{\Gamma(1-\gamma/\alpha)} \prod_{i=1}^{p} \frac{\Gamma(\lambda_i - \alpha)}{\Gamma(1-\alpha)\, \lambda_i!},$$

with the conventions $\Gamma(0) = \infty$ and $\Gamma(0)/\Gamma(0) = 1$ (which will be used throughout this section).

We can also write $\mathrm{AG}^\infty_{\alpha,\gamma} = \mathrm{MB}^{\mathcal{L},q,q_\infty}_\infty$ where q_∞ is the measure on $\mathcal{P}_\infty$ given by

$$q_\infty(\infty, \lambda) = \frac{\gamma/\alpha\, \Gamma(p-\gamma/\alpha)}{\Gamma(1-\gamma/\alpha)\, p!} \, \frac{p!}{\prod_{j\geq 1} m_j(\lambda)!} \prod_{i=1}^{p} \frac{\alpha\, \Gamma(\lambda_i - \alpha)}{\Gamma(1-\alpha)\, \lambda_i!}$$

for all $\lambda = (\lambda_1, \dots, \lambda_p)$ in $\mathcal{P}_{<\infty}$ and $q_\infty(\mu) = 0$ for all μ in $\mathcal{P}_\infty$ with either $p(\mu) = 1$ or $m_\infty(\mu) > 1$.

If X has beta geometric distribution with parameters $(\gamma/\alpha, 1-\gamma/\alpha)$ and is independent of the i.i.d. sequence $(Y_i)_{i\geq 0}$ of beta geometric variables with parameters $(\alpha, 1-\alpha)$, for any $\lambda = (\lambda_1, \dots, \lambda_p)$ in $\mathcal{P}_{<\infty}$, we get that

$$q_\infty(\infty, \lambda) = \mathbb{P}\big[X = p-1,\, (Y_1+1, \dots, Y_{p(\lambda)}+1)^\downarrow = \lambda\big]$$

which ensures that q_∞ is a probability measure on $\mathcal{P}_\infty$.

Proof of Proposition 5.17. Let $\lambda = (\lambda_1, \ldots, \lambda_p)$ be in $\mathcal{P}_{<\infty}$. Then, for n large enough, in light of Stirling's approximation,

$$q_n(n - \|\lambda\|, \lambda) = \frac{1}{\prod_{j\geq 1} m_j(\lambda)!} \left(\gamma + \overbrace{\frac{1-\alpha-\gamma}{n(n-1)} \sum_{i\neq j} \lambda_i \lambda_j}^{\substack{0 \\ \mathrm{dis}\, playstyle \uparrow\ n\to\infty}}\right) \overbrace{\frac{\Gamma(n - \|\lambda\| - \alpha)\, n!}{\Gamma(n-\alpha)\,(n - \|\lambda\|)!}}^{\substack{1 \\ \mathrm{dis}\, playstyle \uparrow\ n\to\infty}}$$
$$\times \frac{\alpha^{p-1}\,\Gamma(p - \gamma/\alpha)}{\Gamma(1-\gamma/\alpha)} \prod_{i=1}^{p} \frac{\Gamma(\lambda_i - \alpha)}{\Gamma(1-\alpha)\,\lambda_i!}$$
$$\xrightarrow[n\to\infty]{} \frac{\gamma/\alpha\,\Gamma(p-\gamma/\alpha)}{\Gamma(1-\gamma/\alpha)\,p!} \frac{p!}{\prod_{j\geq 1} m_j(\lambda)!} \prod_{i12}^{p} \frac{\alpha\,\Gamma(\lambda_i - \alpha)}{\Gamma(1-\alpha)\,\lambda_i!} = q_\infty(\infty, \lambda).$$

We conclude with Corollary 2.10. □

5.3.2 Scaling limits

In this paragraph, we will assume that $0 < \gamma < \alpha < 1$. Let Σ be an α-stable subordinator with Laplace exponent $\lambda \mapsto \lambda^\alpha$ and Lévy measure $\Pi_\alpha(\mathrm{d}t) = \alpha/\Gamma(1-\alpha)\, t^{-1-\alpha}\, \mathbb{1}_{t>0}\, \mathrm{d}t$. Define Δ as the decreasing rearrangement of its jumps on $[0,1]$. We define the dislocation measure $\nu_{\alpha,\gamma}$ for all measurable functions $f : \mathcal{S}^{\downarrow}_{\leq 1} \to \mathbb{R}_+$ by

$$\int_{\mathcal{S}^{\downarrow}_{\leq 1}} f\, \mathrm{d}\nu_{\alpha,\gamma} := \frac{\Gamma(1-\alpha)}{\alpha\,\Gamma(1-\gamma/\alpha)} \mathbb{E}\Big[\Sigma_1^{\alpha+\gamma} \left(\gamma + (1-\alpha-\gamma) \textstyle\sum_{i\neq j} \Delta_i \Delta_j\right) f\big(\Delta/\Sigma_1\big)\Big].$$

Results from [19] and [31] ensure that the family q satisfies Assumption (S): when $n \to \infty$, $n^\gamma (1 - s_1)\, \bar{q}_n(\mathrm{d}s)$ converges weakly towards $(1 - s_1)\, \nu_{\alpha,\gamma}(\mathrm{d}s)$.

We also define the immigration measure $I_{\alpha,\gamma}$ for all measurable functions $F : \mathcal{S}^{\downarrow} \to \mathbb{R}_+$ by

$$\int_{\mathcal{S}^{\downarrow}} F\, \mathrm{d}I_{\alpha,\gamma} := \frac{\gamma/\alpha}{\Gamma(1-\gamma/\alpha)} \int_0^\infty \frac{\mathbb{E}\big[F(t^{1/\alpha}\,\Delta)\big]}{t^{1+\gamma/\alpha}}\, \mathrm{d}t.$$

Proposition 5.18. *Let T be distributed according to $\mathrm{AG}^\infty_{\alpha,\gamma}$ and endow it with μ_T, the counting measure on the set of its leaves. With respect to the $\mathrm{D_{GHP}}$ topology, $(T/R, \mu_T/R^{1/\gamma})$ converges in distribution to a $(\gamma, \nu_{\alpha,\gamma}, I_{\alpha,\gamma})$ fragmentation tree with immigration.*

Proof. Let Λ be such that (∞, Λ) follows q_∞. For all $R \geq 1$, set $q^{(R)}$ as the distribution of $R^{-1/\gamma}\Lambda$. In light of Theorem 4.2, it is sufficient to prove that $R\,(1 \wedge \|\|s)\, q^{(R)}(\mathrm{d}s) \Rightarrow (1 \wedge \|\|s)\, I_{\alpha,\gamma}(\mathrm{d}s)$ when $R \to \infty$.

To prove this claim, we may proceed as in the proof of Proposition 5.10. The only significant difference is that the constant β used near the end of that proof must now belong to the open interval (γ, α). □

Remark 5.19. Let β be in $(1,2)$ and set $\alpha = 1/\beta$, $\gamma = 1 - \alpha$. It was proved in [42] that the distribution $\mathrm{AG}^n_{1/\beta, 1-1/\beta}$ coincides with $\mathrm{GW}^{\mathcal{L},n}_\xi$, where the generating function of ξ is given by $s \mapsto s + \beta^{-1}(1-s)^\beta$. The results of Propositions 5.17 and 5.18 are then consistent with those of Proposition 5.2 and Remark 5.4.

5.3.3 Ford's α-model

When $\alpha = \gamma$, no weight is ever assigned to vertices. Consequently, the trees generated by this algorithm are a.s. binary (i.e. each vertex has either two children or none).

Furthermore, the sequence $(q_n)_n$ of associated first split distributions is much simpler: $q_1(\varnothing)$ still equals 1, and for $n \geq 2$, if $\alpha < 1$, for all $1 \leq k \leq n/2$,

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

finally if $\alpha = 1$, $q_n(n-1,1) = 1$.

Moreover, if α is positive, for all $n \geq 1$, $q_\infty(\infty, n) = \alpha\,\Gamma(n-\alpha)/[\Gamma(1-\alpha)\,n!]$ and $q_\infty(\lambda) = 0$ if $p(\lambda) \neq 2$ or $m_\infty(\lambda) \neq 1$. As a result, a tree with distribution $\mathrm{AG}^\infty_{\alpha,\alpha}$ is obtained by grafting at each height of an infinite spine a single tree with distribution $\mathrm{AG}^{N+1}_{\alpha,\alpha}$ where N, its number of leaves minus 1, has beta geometric distribution of parameters $(\alpha, 1-\alpha)$.

Scaling limits of Ford's α model Let $\alpha \in (0,1)$. Results from [31, Section 5.2] ensure that $(T_n)_n$ satisfies Assumption (S): when $n \to \infty$, $n^\alpha\,(1-s_1)\,\bar{q}_n(\mathrm{d}\mathbf{s}) \Rightarrow (1-s_1)\,\nu^{(\mathrm{F})}_\alpha(\mathrm{d}\mathbf{s})$ where $\nu^{(\mathrm{F})}_\alpha$ is the binary dislocation measure defined for all measurable $f : \mathcal{S}^{\downarrow}_{\leq 1} \to \mathbb{R}_+$ by

$$\int f\,\mathrm{d}\nu^{(\mathrm{F})}_\alpha = \frac{1}{\Gamma(1-\alpha)} \int_{1/2}^1 \left(\frac{\alpha}{[x(1-x)]^{1+\alpha}} + \frac{2-4\alpha}{[x(1-x)]^\alpha} \right) f(x, 1-x, 0, 0, \dots)\,\mathrm{d}x.$$

Furthermore, q_∞ is a.s. binary and Stirling's approximation ensures that $q_\infty(\infty, n)$ is equivalent to $[\alpha/\Gamma(1-\alpha)]\,n^{-1-\alpha}$ when $n \to \infty$. Consequently, if Λ is such that (∞, Λ) follows q_∞ and $q^{(R)}$ denotes the distribution of $\Lambda/R^{1/\alpha}$, Proposition 4.19 proves that $R\,(1 \wedge \|\|\|s)\,q^{(R)}(\mathrm{d}\mathbf{s}) \Rightarrow (1 \wedge \|\|\|s)\,[\alpha/\Gamma(1-\alpha)]\,I^{\mathrm{un}}_\alpha(\mathrm{d}\mathbf{s})$ as $R \to \infty$. Therefore, if we set $I^{(\mathrm{F})}_\alpha := \alpha/\Gamma(1-\alpha) \cdot I^{\mathrm{un}}_\alpha$, we may use Theorem 4.2 and Proposition 4.3 to get the following result:

Proposition 5.20. *Let T be an $\mathrm{AG}^\infty_{\alpha,\alpha}$ tree with α in $(0,1)$ and endow it with the counting measure on the set of its leaves. Then, $(T/R, \mu^{\mathcal{L}}_T/R^{1/\alpha})$ converges in distribution to a $(\alpha, \nu^{(\mathrm{F})}_\alpha, I^{(\mathrm{F})}_\alpha)$-fragmentation tree with immigration with respect to the topology induced by $\mathrm{D}_{\mathrm{GHP}}$.*

Remark 5.21. When $\alpha = 1/2$, i.e. in Rémy's algorithm, these results coincide with Proposition 5.2 and Proposition 5.3 (i') for $\xi(0) = \xi(2) = 1/2$.

When $\alpha = 1$ In this case, the algorithm's output is deterministic: for each $n \geq 2$, a tree T_n with distribution $\mathrm{AG}^n_{1,1}$ is simply equal to a branch of length $n-1$ upon which a single leaf has been grafted at each non-leaf vertex (a "comb" of length n). Similarly, an infinite tree with distribution $\mathrm{AG}^\infty_{1,1}$ is the "infinite comb", obtained by attaching a single leaf to all the vertices of the infinite branch.

As a result, if T has distribution $\mathrm{AG}^\infty_{1,1}$ and μ_T denotes the counting measure on the set of its leaves, then clearly, $(T/R, \mu_T/R)$ converges as $R \to \infty$ to the metric space $\mathbb{R}_+$ rooted at 0 and endowed with the usual Lebesgue measure.

When $\alpha = 0$ Observe that $q_n(n-k,k) = (2 - \mathbb{1}_{k=n/2})/(n-1)$. Then for all $K \geq 1$ and n large enough,

$$\mathbb{P}[\Lambda^{\mathcal{L}}(T_n) \wedge K = \infty_2 \wedge K] = 1 - \frac{K-1}{n-1} \xrightarrow[n\to\infty]{} 1,$$

which implies $\Lambda^{\mathcal{L}}(T_n) \to (\infty, \infty)$ a.s. when $n \to \infty$. Theorem 2.9 then ensures that T_n converges in distribution to the complete infinite binary tree (in which every vertex has 2 children). Moreover, since $T_n \subset T_{n+1}$ a.s., this convergence happens almost surely.

With Assumptions (S) **and** (I′) Let us give an application of the result from Remark 4.5.

For any α in $(0,1)$ and $n \in \mathbb{N} \cup \{\infty\}$, denote by $q_n^{(\alpha)}$ the first-split distribution (with respect to the number of leaves) associated to a tree with distribution $\mathrm{AG}^n_{\alpha,\alpha}$. Now, fix $0 < \alpha < \beta < 1$, and consider a tree T with distribution $\mathrm{MB}_\infty^{\mathcal{L},q^{(\alpha)},q_\infty^{(\beta)}}$ endowed with μ_T, the counting measure on the set of its leaves. We may deduce from previous results that $(q_n^{(\alpha)})_{n\geq 1}$ and $q_\infty^{(\beta)}$ satisfy Assumptions (S) and (I′).

As a result, $(T/R, \mu_T/R^{1/\beta})$ converges in distribution to the metric space $\mathbb{R}_+$ rooted at 0 and endowed with a random measure $\mu = \sum_{i\geq 1} \|\mathbf{s}_i\| \, \delta_{u_i}$, where $\sum_{i\geq 1} \delta_{(u_i,\mathbf{s}_i)}$ is a Poisson point process on $\mathbb{R}_+ \times \mathcal{S}^\downarrow$ with intensity measure $\mathrm{d}u \otimes I_\beta^{(\mathrm{F})}(\mathrm{d}\mathbf{s})$.

5.4 Aldous' β-splitting model

This section will focus on the study a model of binary random trees introduced in [6, Section 4] as a Markov branching model. Let $\beta > -2$ be fixed. Set $q_1(\varnothing) := 1$ and for all $n \geq 2$ and $1 \leq k \leq n/2$,

$$q_n(n-k,k) := \frac{2 - \mathbb{1}_{2k=n}}{Z_n} \frac{\Gamma(n-k+1+\beta)}{(n-k)!} \frac{\Gamma(k+1+\beta)}{k!}$$

where Z_n is a normalising constant. For all $n \geq 1$, let T_n be a random tree with distribution $\mathrm{MB}_n^{\mathcal{L},q}$.

Remark 5.22.
- The constant Z_n is given by

$$Z_n := \sum_{k=1}^{n-1} \frac{\Gamma(n-k+1+\beta)}{(n-k)!} \frac{\Gamma(k+1+\beta)}{k!}.$$

  When $\beta > -1$, it simplifies to $Z_n = [\mathrm{B}(1+\beta, 1+\beta) - 2\ \mathrm{B}(n+1+\beta, 1+\beta)] \cdot \Gamma(n+2+2\beta)/n!$ (where B denotes the usual Beta function) and when $\beta = -1$, it becomes $Z_n = 2/n \cdot \sum_{k=1}^{n-1} k^{-1}$.
- When $\beta = -3/2$, observe that the sequence $(q_n)_n$ is the same as that of the α-model with $\alpha = 1/2$ (see Section 5.3.3). Therefore, like Rémy's algorithm, this model generates uniform binary trees with any given number of leaves.

There are three regimes in this model, respectively $\beta > -1$, $\beta = -1$ and $\beta \in (-2,-1)$. The asymptotic behaviour of q_n were studied in [6, Section 5] in these three regimes.

5.4.1 Local limits

In this paragraph, we will focus on the study of the local limits of T_n. We will once again rely on the Markov branching nature of the model and on Theorem 2.9.

Proposition 5.23. $\boldsymbol{\beta \geq -1}$ **:** *In the sense of the local limit topology, T_n converges in distribution to the infinite binary tree.*

$\boldsymbol{\beta \in (-2,-1)}$ **:** *Let X follow the beta geometric distribution with parameters $(2+\beta, -1-\beta)$ (see Section 5.3). Define q_∞, a probability measure on $\mathcal{P}_\infty$, by $q_\infty(\infty, k) = \mathbb{P}[X = k-1]$ for any $k \geq 1$ and $q_\infty(\lambda) = 0$ if $p(\lambda) \neq 2$ or $m_\infty(\lambda) \neq 1$. With these notations, T_n converges in distribution to $\mathrm{MB}_\infty^{\mathcal{L},q,q_\infty}$ with respect to the local limit topology.*

Remark 5.24. Suppose $\beta \in (-2,-1)$ and let $(X_n, \tau_n)_{n\geq 0}$ be an i.i.d. sequence such that for each n, X_n has beta geometric distribution with parameters $(2+\beta, -1-\beta)$ and conditionally on $X_n = k-1$, τ_n is distributed like T_k. Finally, denote by T_∞ the tree obtained by attaching by a single edge the tree τ_n respectively at each height n of an infinite branch, i.e. $T_\infty := \mathtt{b}_\infty \bigotimes_{n\geq 0} (\mathtt{v}_n, [\![\tau_n]\!])$. The tree T_∞ hence obtained has distribution $\mathrm{MB}_\infty^{\mathcal{L},q,q_\infty}$.

Proof. Observe that in light of Stirling's approximation, $\Gamma(n+1+\beta)/n! \sim n^{\beta}$ when $n \to \infty$.

$\beta \geq -1$: When $\beta > -1$, using Stirling's approximation once again, we get that $Z_n \sim \mathrm{B}(1+\beta, 1+\beta)\, n^{-1-2\beta}$ so if $k \geq 1$ is a fixed integer, $q_n(n-k,k) = O(n^{1+\beta})$ when $n \to \infty$.

When $\beta = -1$, $Z_n \sim 2/n \cdot \log n$ hence, for any fixed $k \geq 1$, $q_n(n-k,k) \sim 1/(k \log n)$ as $n \to \infty$.

Therefore, for any $\beta \geq -1$, if $K \geq 1$,

$$q_n\big[\mu \in \mathcal{P}_n : \mu \wedge K = (K,K)\big] = 1 - \textstyle\sum_{k=1}^{K} q_n(n-k,k) \xrightarrow[n\to\infty]{} 1.$$

Lemma 2.4 then ensures that $q_n \Rightarrow \delta_{(\infty,\infty)}$. It follows from Theorem 2.9 that T_n converges in distribution to the (deterministic) infinite binary tree.

$\beta \in (-2,-1)$: Let $\beta \in (-2,-1)$. In light of Stirling's formula, we know that the sequence $\big(i^{-\beta}\,\Gamma(i+1+\beta)/i!\big)_{i\geq 1}$ is bounded by a finite constant. As a result, the dominated convergence theorem ensures that

$$\begin{aligned}\frac{Z_n}{n^{\beta}} &= \sum_{k\geq 1} \frac{\Gamma(k+1+\beta)}{k!}\,\frac{\Gamma(n-k+1+\beta)}{(n-k)^{\beta}\,(n-k)!}\,\frac{(n-k)^{\beta}}{n^{\beta}}\,(2-\mathbb{1}_{2k=n})\,\mathbb{1}_{2k\leq n} \\ &\xrightarrow[n\to\infty]{} 2\sum_{k\geq 1}\frac{\Gamma(k+1+\beta)}{k!} = 2\frac{\Gamma(2+\beta+)}{-1-\beta}\end{aligned}$$

where we have used the definition of the beta geometric distribution with parameters $(2+\beta, -1-\beta)$ as introduced in Section 5.3.

Consequently, for any fixed positive integer k,

$$\begin{aligned}\lim_{n\to\infty} q_n(n-k,k) &= \lim_{n\to\infty} 2\,\frac{\Gamma(k+1+\beta)}{k!}\,\frac{\Gamma(n-k+1+\beta)}{Z_n\,(n-k)!} \\ &= \frac{(-1-\beta)\,\Gamma(k+1+\beta)}{\Gamma(2+\beta)\,k!} = q_\infty(\infty,k).\end{aligned}$$

We may then conclude with Corollary 2.10. □

5.4.2 Scaling limits

We will now study the scaling limits of the β-splitting model when $\beta \in (-2,-1)$ with the help of Theorem 4.2.

Let $\nu_\beta^{(\mathrm{B})}$ be the dislocation measure such that for all measurable $f : \mathcal{S}^{\downarrow}_{\leq 1} \to \mathbb{R}_+$,

$$\int f \,\mathrm{d}\nu_\beta^{(\mathrm{B})} := \frac{-1-\beta}{\Gamma(2+\beta)} \int_0^{1/2} t^{\beta}\,(1-t)^{\beta}\, f(1-t,t,0,0,\dots)\,\mathrm{d}t.$$

It follows from Section 5.1 in [31] that $(q_n)_{n\geq 1}$ satisfies Assumption (S) for $\gamma = -1-\beta$ and $\nu = \nu_\beta^{(\mathrm{B})}$. More precisely, $n^{-1-\beta}\,(1-s_1)\,\bar{q}_n(\mathrm{d}s)$ converges weakly to $(1-s_1)\,\nu_\beta^{(\mathrm{B})}(\mathrm{d}s)$ as finite measures on $\mathcal{S}^{\downarrow}_{\leq 1}$.

Let Λ denote a random integer such that (∞,Λ) has distribution q_∞ and for all $R \geq 1$, set $q^{(R)}$ as the distribution of $\Lambda/R^{1/(-1-\beta)}$. Just like in Section 5.3.3, Stirling's approximation and Proposition 4.19 ensure that Assumption (I) is met for $\gamma = -1-\beta$ and the immigration measure $I_\beta^{(\mathrm{B})} := (-1-\beta)/\Gamma(2+\beta)\cdot I^{\mathrm{un}}_{-1-\beta}$. As a result,

Proposition 5.25. *Fix $\beta \in (-2,-1)$. Let T be a $\mathrm{MB}^{\mathcal{L},q,q_\infty}_{\infty}$ tree and endow it with μ_T, the counting measure on the set of its leaves. In the topology induced by $\mathrm{D_{GHP}}$, $(T/R, \mu_T^{\mathcal{L}}/R^{1/(-1-\beta)})$ converges in distribution to a $(-1-\beta, \nu_\beta^{(\mathrm{B})}, I_\beta^{(\mathrm{B})})$-fragmentation tree with immigration.*

5.5 k-ary growing trees

Let $k \geq 2$ be an integer. In this section, we will study a model of k-ary trees, i.e. trees in which vertices have either 0 or k children, described in [32]. This model is yet another generalisation of Rémy's algorithm [45] (which corresponds to $k = 2$).

The following algorithm allows us to get a sequence $(T_n)_{n\geq 0}$ of k-ary trees such that for all n, T_n has n internal vertices (vertices that aren't leaves) or, equivalently, $kn + 1$ vertices or $(k-1)n + 1$ leaves. First, let T_0 be the trivial tree $\{\varnothing\}$ and for $n \geq 1$, conditionally on T_{n-1}:

- Pick an edge of T_{n-1} (considered as a planted tree) uniformly at random,
- Place a new vertex on that edge and attach $k-1$ new leaves to it,

and call T_n the resulting tree. We will denote the distribution of T_n by GT_k^n.

The negative Dirichlet multinomial distribution Let Π be a $(k-1)$-dimensional Dirichlet variable with k parameters $(1/k, \dots, 1/k)$, i.e. Π takes its values in the $(k-1)$-dimensional simplex $\{\boldsymbol{x} \in (0,\infty)^k : x_1 + \dots + x_k = 1\}$. Conditionally on Π, let $X = (X_1, \dots, X_{k-1})$ have negative multinomial distribution of parameters $(1; \Pi)$, i.e. for each $i \in \{1, \dots, k-1\}$, X_i counts the number of type i results before the first type k result (failure) in a sequence of i.i.d. trials with k possible results with respective probabilities $\Pi_1, \dots, \Pi_k$. For any non-negative integers $n_1, \dots, n_{k-1}$ and with $N = n_1 + \dots + n_{k-1}$, we have

$$\mathbb{P}\big[X = (n_1, \dots, n_{k-1})\big] = \mathbb{E}\left[\frac{N!}{n_1! \dots n_{k-1}!} \prod_{i=1}^{k-1} \Pi_i^{n_i}\, \Pi_k\right] = \frac{1}{k}\,\frac{1}{1+N} \prod_{i=2}^{k} \frac{\Gamma(n_i + 1/k)}{\Gamma(1/k)\, n_i!}.$$

The random variable X is said to follow a $(k-1)$-dimensional *negative Dirichlet multinomial* distribution with parameters $(1; 1/k, \dots, 1/k)$ which is a multidimensional generalisation of the beta geometric distribution. Further observe that the sum $\|X\| = X_1 + \dots + X_{k-1}$ has beta geometric distribution with parameters $(1/k, 1 - 1/k)$ and that conditionally on $\|X\| = n$, X follows a $(k-1)$-dimensional Dirichlet multinomial distribution with parameters $(n\,;\, 1/k, \dots, 1/k)$.

Corresponding infinite tree Let $(X_n, \tau_{n,1}, \dots \tau_{n,k-1})_{n\geq 0}$ be a sequence of i.i.d. variables such that for all $n \geq 0$, X_n is distributed according to a $(k-1)$-dimensional $(1; 1/k, \dots, 1/k)$ negative Dirichlet multinomial distribution and conditionally on $X_n = (m_1, \dots, m_{k-1})$, $\tau_{n,1}, \dots, \tau_{n,k-1}$ are independent and have respective distributions $\mathrm{GT}_k^{m_1}$, $\dots$, $\mathrm{GT}_k^{m_{k-1}}$.

Conditionally on $(X_n, \tau_{n,1}, \dots \tau_{n,k-1})_{n\geq 0}$, let T_∞ be the tree obtained after grafting at each height $n \geq 0$ of an infinite branch the concatenation of $\tau_{n,i}$, $1 \leq i \leq k-1$, i.e. set

$$T_\infty := \mathtt{b}_\infty \bigotimes\nolimits_{n\geq 0} \big(\mathtt{v}_n, [\![\tau_{n,1}, \dots \tau_{n,k-1}]\!]\big),$$

and let GT_k^∞ be the distribution of T_∞.

Section 5.5.1 will prove the following proposition.

Proposition 5.26. *In the sense of the local limit topology, GT_k^n converges weakly to GT_k^∞ when n goes to ∞.*

Let Π be a $(k-1)$-dimensional Dirichlet variable with parameters $(1/k, \dots, 1/k)$. Following [32, Section 3.1], we define the dislocation measure ν_k^{GT} such that for all measurable $f : \mathcal{S}_{\leq 1}^{\downarrow} \to \mathbb{R}_+$

$$\int_{\mathcal{S}_{\leq 1}^{\downarrow}} f \,\mathrm{d}\nu_k^{\mathrm{GT}} = \frac{\Gamma(1/k)}{k}\, \mathbb{E}\left[\frac{f\big[(\Pi, 0, 0, \dots)^{\downarrow}\big]}{1 - \Pi_1}\right].$$

Let Δ be a $(k-2)$-dimensional Dirichlet variable with parameters $(1/k,\ldots,1/k)$. We also define the immigration measure I_k^{GT} for all measurable functions $F:\mathcal{S}^{\downarrow}\to\mathbb{R}_+$ by

$$\int_{\mathcal{S}^{\downarrow}} F(\mathbf{s})\, I_k^{\mathrm{GT}}(\mathrm{d}\mathbf{s}) := \frac{1/k}{\Gamma(1-1/k)}\int_0^{\infty} t^{-1-1/k}\,\mathbb{E}\Big[F\big(t\,(\Delta,0,0,\ldots)^{\downarrow}\big)\Big]\,\mathrm{d}t.$$

The aim of Section 5.5.2 will be to prove the next proposition.

Proposition 5.27. *Let T be a GT_k^{∞}-distributed tree and endow it with μ_T°, the counting measure on the set of its internal vertices. With respect to the topology induced by $\mathrm{D}_{\mathrm{GHP}}$, when $R\to\infty$, $(T/R,\mu_T^{\circ}/R^k)$ converges in distribution to a $(1/k,\nu_k^{\mathrm{GT}},I_k^{\mathrm{GT}})$-fragmentation tree with immigration.*

5.5.1 Markov branching property and local limits

For any $\mathtt{t}$ in $\mathtt{T}$, we define $\Lambda^{\circ}(\mathtt{t})$ as the decreasing rearrangement of the number of internal vertices of the sub-trees of $\mathtt{t}$ attached to its root, i.e. we let $\Lambda^{\circ}(\mathtt{t}) := \Lambda(\mathtt{t})-\Lambda^{\mathcal{L}}(\mathtt{t})$. In the setting of k-ary growing trees, $\Lambda^{\circ}(T_0)=\varnothing$ a.s. and if $n\geq 1$, $\Lambda^{\circ}(T_n)$ takes its values in the set of decreasing families of $(\mathbb{Z}_+)^k$ with sum $n-1$. Because of the deterministic relationship between n, $\#T_n$ and $\#_{\mathcal{L}}T_n$, we have $\Lambda(T_0)=\Lambda^{\mathcal{L}}(T_0)=\varnothing$ and for $n\geq 1$, $\Lambda(T_n)=k\Lambda^{\circ}(T_n)+(1,\ldots,1)$ in $\mathcal{P}_{kn}$ and $\Lambda^{\mathcal{L}}(T_n)=(k-1)\Lambda^{\circ}(T_n)+(1,\ldots,1)$ in $\mathcal{P}_{(k-1)n+1}$. For all $n\geq 1$, call q_{n-1}° the distribution of $\Lambda^{\circ}(T_n)$, that is the first-split distribution of T_n with respect to internal vertices.

Proposition 3.3 from [32] states that $(T_n)_{n\geq 0}$ satisfies the Markov branching property and the distribution of T_n may be expressed as either MB_{kn+1}^{q} or $\mathrm{MB}_{(k-1)n+1}^{\mathcal{L},q^{\mathcal{L}}}$ where q and $q^{\mathcal{L}}$ are both easily obtained from $(q_n^{\circ})_{n\geq 0}$. Rewriting the formula from this last proposition for our purposes (where partition blocs are arranged in decreasing order), for all $n\geq 1$ and $\lambda=(\lambda_1,\ldots,\lambda_k)$ decreasing with sum n, we get that

$$\begin{aligned} q_{n-1}^{\circ}(\lambda) = \frac{(k-1)!}{\prod_{j\geq 1} m_j(\lambda)!}\,\frac{1}{k}\,\frac{\Gamma(1/k)}{\Gamma(n+1+1/k)}\prod_{i=1}^{k}\frac{\Gamma(\lambda_i+1/k)}{\Gamma(1/k)\,\lambda_i!} \\ \times\sum_{i=1}^{k}\left(m_{\lambda_i}(\lambda)\lambda_i!\sum_{j=0}^{\lambda_i}\frac{(j+n-\lambda_i)!}{j!}\right). \end{aligned}$$

We can rewrite GT_k^{∞} as the distribution $\mathrm{MB}_{\infty}^{q,q_{\infty}}$ or $\mathrm{MB}_{\infty}^{\mathcal{L},q^{\mathcal{L}},q_{\infty}^{\mathcal{L}}}$ of an infinite Markov branching tree. The corresponding measures q_{∞} and $q_{\infty}^{\mathcal{L}}$ on $\mathcal{P}_{\infty}$ can also be easily deduced from the measure q_{∞}° on the set of decreasing k-tuples of $\mathbb{Z}_+\cup\{\infty\}$ with infinite sum such that $q_{\infty}^{\circ}(\lambda)=0$ if λ_2 is infinite and

$$q_{\infty}^{\circ}(\infty,\lambda_2,\ldots,\lambda_k) = \frac{(k-1)!}{\prod_{j\geq 1} m_j(\lambda)!}\,\frac{1}{k}\,\frac{1}{\|\lambda\|+1}\prod_{i=2}^{k}\frac{\Gamma(\lambda_i+1/k)}{\Gamma(1/k)\,\lambda_i!}$$

for any integers $\infty>\lambda_2\geq\cdots\geq\lambda_p\geq 0$. Observe that $q_{\infty}^{\circ}(\infty,\lambda_2,\ldots,\lambda_k)=\mathbb{P}\big[X^{\downarrow}=(\lambda_2,\ldots,\lambda_k)\big]$ where X is a $(k-1)$-dimensional negative Dirichlet multinomial variable with parameters $(1;1/k,\ldots,1/k)$. As a result, q_{∞}° is a probability measure.

Proof of Proposition 5.26. Let $\lambda=(\lambda_2,\ldots,\lambda_k)$ be a decreasing sequence of $(\mathbb{Z}_+)^{k-1}$ and set $L=\lambda_2+\cdots+\lambda_k$. For n large enough, we have

$$
\begin{aligned}
&q^\circ_n(n-L,\lambda_2,\dots,\lambda_k)\\
&= \frac{(k-1)!}{\prod\limits_{j\ge 1} m_j(\lambda)!}\,\frac{1}{k}\prod_{i=2}^{k}\frac{\Gamma(\lambda_i+1/k)}{\Gamma(1/k)\,\lambda_i!}\overbrace{\frac{\Gamma(n-L+1/k)}{\Gamma(n+1+1/k)}}^{\underset{n\to\infty}{\sim} n^{-L-1}}\left[\overbrace{\sum_{j=0}^{n-L}\frac{(j+L)!}{j!}}^{\underset{n\to\infty}{\sim} n^{L+1}\int_0^1 x^L\,\mathrm{d}x}+\overbrace{\sum_{i=2}^{k}\sum_{j=0}^{\lambda_i}\frac{\lambda_i!\,(j+n-\lambda_i)!}{(n-L)!\,j!}}^{\underset{n\to\infty}{=} O(n^L)}\right]\\
&\xrightarrow[n\to\infty]{}\frac{(k-1)!}{\prod_{j\ge1} m_j(\lambda)!}\,\frac{1}{k}\,\frac{1}{L+1}\prod_{i=2}^{k}\frac{\Gamma(\lambda_i+1/k)}{\Gamma(1/k)\,\lambda_i!}=q^\circ_\infty(\infty,\lambda).
\end{aligned}
$$

Corollary 2.10 concludes this proof. □

5.5.2 Scaling limits

Proposition 3.1 in [32] states that $n^{1/k}\,(1-s_1)\,\bar{q}^\circ_n(\mathrm{d}s) \Rightarrow (1-s_1)\,\nu^{\mathrm{GT}}_k(\mathrm{d}s)$ as $n\to\infty$ in the sense of finite measures on $\mathcal{S}^\downarrow_{\le 1}$. Assumption ($\mathtt{S}$) of Theorem 4.2 is thus met for the sequence q°. To prove Proposition 5.27, we will need the following lemma. Let $X=(X_1,\dots,X_{k-1})$ denote a negative Dirichlet multinomial variable with parameters $(1;1/k,\dots,1/k)$.

Lemma 5.28. *Let Δ be a $(k-2)$-dimensional Dirichlet $(1/k,\dots,1/k)$ variable. For all Lipschitz-continuous functions $G:[0,\infty)^{k-1}\longrightarrow\mathbb{R}_+$ such that $G(\boldsymbol{x})\le 1\wedge\|\boldsymbol{x}\|$ for all $\boldsymbol{x}$ in $[0,\infty)^{k-1}$,*

$$
R\,\mathbb{E}\left[G\left(\frac{X}{R^k}\right)\right]\xrightarrow[R\to\infty]{}\frac{1/k}{\Gamma(1-1/k)}\int_0^\infty t^{-1-1/k}\mathbb{E}[G(t\,\Delta)]\,\mathrm{d}t.
$$

Proof. Let $(Y_n)_{n\ge1}$ be i.i.d. and such that conditionally on Δ, Y_n is multinomial with parameters $(1;\Delta)$. Moreover, set $Z_n:=Y_1+\dots+Y_n$. The law of large numbers ensures that Z_n/n converges almost surely to Δ. Let N be independent of Δ and $(Z_n)_n$ and have beta geometric distribution with parameters $(1/k,1-1/k)$. Observe that X has the same distribution as Z_N.

Define $g:\mathbb{R}_+\to\mathbb{R}_+$ by $g(t):=\mathbb{E}[G(t\,\Delta)]$. The dominated convergence theorem implies that it is continuous and it clearly satisfies $g(t)\le 1\wedge t$. Lemma 4.18 then ensures that $R\,\mathbb{E}[g(N/R^k)]\to\left[k\,\Gamma(1-1/k)\right]^{-1}\int_0^\infty t^{-1-1/k}g(t)\,\mathrm{d}t$.

Since Z_n/n a.s. converges to Δ and because $\|(Z_n/n)-\Delta\|\le 2$, we can use the dominated convergence theorem to state that for all positive ε, there exists n_ε such that $\mathbb{E}[\|(Z_n/n)-\Delta\|]<\varepsilon$ as soon as $n\ge n_\varepsilon$. Therefore, if K is the Lipschitz constant of G,

$$
\begin{aligned}
\left|R\,\mathbb{E}\left[G\left(\frac{X}{R^k}\right)\right]-R\,\mathbb{E}\left[g\left(\frac{N}{R^k}\right)\right]\right| &\le R\,\mathbb{E}\left[\left|G\left[\frac{N}{R^k}\frac{Z_N}{N}\right]-G\left[\frac{N}{R^k}\,\Delta\right]\right|\right]\\
&\le R\,\mathbb{E}\left[1\wedge\left(K\varepsilon\frac{N}{R^k}\right)\right]+\frac{2Kn_\varepsilon}{R^{k-1}}\xrightarrow[R\to\infty]{}\frac{1/k}{\Gamma(1-1/k)}\int_0^\infty\frac{1\wedge(K\varepsilon\,t)}{t^{1+1/k}}\mathrm{d}t
\end{aligned}
$$

where we have used Lemma 4.18. This last quantity in turn converges to 0 when $\varepsilon\to0$ which proves the desired result. □

Proof of Proposition 5.27. Recall that if Λ is such that (∞,Λ) follows q°_∞, then Λ is distributed like $X^\downarrow$. We may then deduce from Lemma 5.28 and Lemma 3.21 that Assumption ($\mathtt{I}$) holds for q°_∞, $I=I^{\mathrm{GT}}_k$ and $\gamma=1/k$. As a result, Theorem 4.2 concludes this proof. □

References

[1] Romain Abraham and Jean-François Delmas, *Local limits of conditioned Galton-Watson trees: the condensation case*, Electron. J. Probab **19** (2014), no. 56, 1–29. MR-3227065

[2] Romain Abraham and Jean-François Delmas, *Local limits of conditioned Galton-Watson trees: the infinite spine case*, Electron. J. Probab. **19** (2014), no. 2, 1–19. MR-3164755

[3] Romain Abraham, Jean-François Delmas, and Patrick Hoscheit, *A note on the Gromov-Hausdorff-Prokhorov distance between (locally) compact metric measure spaces*, Electron. J. Probab **18** (2013), no. 14, 1–21. MR-3035742

[4] Louigi Addario-Berry, Nicolas Broutin, Christina Goldschmidt, and Grégory Miermont, *The scaling limit of the minimum spanning tree of the complete graph*, Ann. Probab. (2017), to appear.

[5] David Aldous, *The continuum random tree. I*, Ann. Probab. **19** (1991), no. 1, 1–28. MR-1085326

[6] David Aldous, *Probability distributions on cladograms*, Random discrete structures (Minneapolis, MN, 1993), IMA Vol. Math. Appl., vol. 76, Springer, New York, 1996, pp. 1–18. MR-1395604

[7] David Aldous and Jim Pitman, *Tree-valued Markov chains derived from Galton-Watson processes*, Ann. Inst. Henri Poincaré Probab. Stat. **34** (1998), no. 5, 637–686. MR-1641670

[8] Julien Berestycki, *Ranked fragmentations*, ESAIM Probab. Statist. **6** (2002), 157–175. MR-1943145

[9] Jean Bertoin, *Self-similar fragmentations*, Ann. Inst. Henri Poincaré Probab. Stat. **38** (2002), no. 3, 319–340. MR-1899456

[10] Jean Bertoin, *The asymptotic behavior of fragmentation processes*, J. Eur. Math. Soc. (JEMS) **5** (2003), no. 4, 395–416. MR-2017852

[11] Jean Bertoin, *Random fragmentation and coagulation processes*, Cambridge Studies in Advanced Mathematics, vol. 102, Cambridge University Press, Cambridge, 2006. MR-2253162

[12] Jean Bertoin, *Fires on trees*, Ann. Inst. Henri Poincaré Probab. Stat. **48** (2012), no. 4, 909–921. MR-3052398

[13] Jean Bertoin, *Sizes of the largest clusters for supercritical percolation on random recursive trees*, Random Structures & Algorithms **44** (2014), no. 1, 29–44. MR-3143589

[14] Jean Bertoin, *The cut-tree of large recursive trees*, Ann. Inst. Henri Poincaré Probab. Stat. **51** (2015), no. 2, 478–488. MR-3335011

[15] Jean Bertoin and Grégory Miermont, *The cut-tree of large Galton-Watson trees and the Brownian CRT*, Ann. Appl. Probab. **23** (2013), no. 4, 1469–1493. MR-3098439

[16] Patrick Billingsley, *Convergence of probability measures*, second ed., Wiley Series in Probability and Statistics: Probability and Statistics, John Wiley & Sons, Inc., New York, 1999. MR-1700749

[17] Nicolas Broutin, Luc Devroye, Erin McLeish, and Mickael de la Salle, *The height of increasing trees*, Random Structures & Algorithms **32** (2008), no. 4, 494–518. MR-2422392

[18] Nicolas Broutin and Cécile Mailler, *And/or trees: A local limit point of view*, Random Structures & Algorithms (2017), to appear.

[19] Bo Chen, Daniel J Ford, and Matthias Winkel, *A new family of Markov branching trees: the alpha-gamma model*, Electron. J. Probab. **14** (2009), no. 15, 400–430. MR-2480547

[20] Nicolas Curien and Igor Kortchemski, *Random non-crossing plane configurations: A conditioned Galton-Watson tree approach*, Random Structures & Algorithms **45** (2014), no. 2, 236–260. MR-3245291

[21] Daphné Dieuleveut, *The vertex-cut-tree of Galton-Watson trees converging to a stable tree*, Ann. Appl. Probab. **25** (2015), no. 4, 2215–2262. MR-3349006

[22] Thomas Duquesne, *Continuum random trees and branching processes with immigration*, Stochastic Process. Appl. **119** (2009), no. 1, 99–129. MR-3349006

[23] Thomas Duquesne and Jean-François Le Gall, *Random trees, Lévy processes and spatial branching processes*, Astérisque **281** (2002). MR-1954248

[24] Daniel J Ford, *Probabilities on cladograms: introduction to the alpha model*, arXiv preprint arXiv:math/0511246 (2005).

[25] Bénédicte Haas, *Loss of mass in deterministic and random fragmentations*, Stochastic Process. Appl. **106** (2003), no. 2, 245–277. MR-1989629

[26] Bénédicte Haas, *Equilibrium for fragmentation with immigration*, Ann. Appl. Probab. **15** (2005), no. 3, 1958–1996. MR-2152250

[27] Bénédicte Haas, *Fragmentation processes with an initial mass converging to infinity*, J. Theoret. Probab. **20** (2007), no. 4, 721–758. MR-2359053

[28] Bénédicte Haas, *Scaling limits of Markov-Branching trees and applications*, arXiv preprint arXiv:1605.07873 (2016).

[29] Bénédicte Haas and Grégory Miermont, *The genealogy of self-similar fragmentations with negative index as a continuum random tree*, Electron. J. Probab. **9** (2004), no. 4, 57–97. MR-2041829

[30] Bénédicte Haas and Grégory Miermont, *Scaling limits of Markov branching trees with applications to Galton-Watson and random unordered trees*, Ann. Probab. **40** (2012), no. 6, 2589–2666. MR-3050512

[31] Bénédicte Haas, Grégory Miermont, Jim Pitman, and Matthias Winkel, *Continuum tree asymptotics of discrete fragmentations and applications to phylogenetic models*, Ann. Probab. **36** (2008), no. 5, 1790–1837. MR-2440924

[32] Bénédicte Haas and Robin Stephenson, *Scaling limits of k-ary growing trees*, Ann. Inst. Henri Poincaré Probab. Stat. **51** (2015), no. 4, 1314–1341. MR-3414449

[33] Svante Janson, *Simply generated trees, conditioned Galton-Watson trees, random allocations and condensation*, Probab. Surv. **9** (2012), 103–252. MR-2908619

[34] Thordur Jonsson and Sigurður Örn Stefánsson, *Condensation in nongeneric trees*, J. Stat. Phys. **142** (2011), no. 2, 277–313. MR-2764126

[35] Olav Kallenberg, *Random measures*, fourth ed., Akademie-Verlag, Berlin; Academic Press, Inc., London, 1986. MR-854102

[36] Douglas P Kennedy, *The Galton-Watson process conditioned on the total progeny*, J. Appl. Probability **12** (1975), no. 4, 800–806. MR-0386042

[37] Harry Kesten, *Subdiffusive behavior of random walk on a random cluster*, Ann. Inst. Henri Poincaré Probab. Stat. **22** (1986), no. 4, 425–487. MR-871905

[38] J. F. C. Kingman, *Poisson processes*, Oxford Studies in Probability, vol. 3, The Clarendon Press, Oxford University Press, New York, 1993, Oxford Science Publications. MR-1207584

[39] Igor Kortchemski, *Invariance principles for Galton-Watson trees conditioned on the number of leaves*, Stochastic Process. Appl. **122** (2012), no. 9, 3126–3172. MR-2946438

[40] Amaury Lambert, *Probabilistic Models for the (sub)Tree(s) of Life*, Braz. J. Probab. Stat. **31** (2017), no. 3, 415–475.

[41] Jean-François Le Gall, *Random real trees*, Ann. Fac. Sci. Toulouse Math. (6) **15** (2006), no. 1, 35–62. MR-2225746

[42] Philippe Marchal, *A note on the fragmentation of a stable tree*, 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. MR-2508809

[43] Grégory Miermont, *Self-similar fragmentations derived from the stable tree. I. Splitting at heights*, Probab. Theory Related Fields **127** (2003), no. 3, 423–454. MR-2018924

[44] J. Pitman, *Combinatorial stochastic processes*, Lecture Notes in Mathematics, vol. 1875, Springer-Verlag, Berlin, 2006, Lectures from the 32nd Summer School on Probability Theory held in Saint-Flour, July 7–24, 2002. MR-2245368

[45] Jean-Luc Rémy, *Un procédé itératif de dénombrement d'arbres binaires et son application à leur génération aléatoire*, RAIRO Inform. Théor. **19** (1985), no. 2, 179–195. MR-803997

[46] Douglas 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. Henri Poincaré Probab. Stat. **51** (2015), no. 2, 512–532. MR-3335013

[47] Sigurður Örn Stefánsson, *The infinite volume limit of Ford's alpha model*, Acta Physica Polonica B Proc. Suppl. **2** (2009), no. 3, 555–560.

[48] Sigurður Örn Stefánsson, *Markov branching in the vertex splitting model*, J. Stat. Mech. Theory Exp. **4** (2012), no. 4, P04018. MR-2956262

[49] Robin Stephenson, *General fragmentation trees*, Electron. J. Probab. **18** (2013), no. 101, 1–45. MR-3141802

[50] Robin Stephenson, *Local convergence of large critical multi-type Galton-Watson trees and applications to random maps*, J. Theoret. Probab. (2017), to appear.

Acknowledgments. The author would like to thank his advisor, Bénédicte Haas, for many helpful discussions as well as an anonymous referee for their various suggestions which helped to improve this paper. This work was partially supported by the ANR GRAAL (ANR-14-CE25-0014).