

Decorated stable trees*

Delphin Sénizergues[†] Sigurdur Örn Stefánsson[‡] Benedikt Stufler[§]

Abstract

We define *decorated α -stable trees* which are informally obtained from an α -stable tree by blowing up its branchpoints into random metric spaces. This generalizes the α -stable looptrees of Curien and Kortchemski, where those metric spaces are just deterministic circles. We provide different constructions for these objects, which allows us to understand some of their geometric properties, including compactness, Hausdorff dimension and self-similarity in distribution. We prove an invariance principle which states that under some conditions, analogous discrete objects, random decorated discrete trees, converge in the scaling limit to decorated α -stable trees. We mention a few examples where those objects appear in the context of random trees and planar maps, and we expect them to naturally arise in many more cases.

Keywords: decorated trees; looptrees; stable trees; invariance principle; self-similarity.

MSC2020 subject classifications: 60F17; 60C05.

Submitted to EJP on May 19, 2022, final version accepted on October 30, 2023.

1 Introduction

In this paper we develop a general method for constructing random metric spaces by gluing together metric spaces referred to as *decorations* along a tree structure which is referred to as *the underlying tree*. The resulting object, which we define informally below and properly in Section 3 and Section 4, will be called a *decorated tree*. We will consider two frameworks, which we refer to as the *discrete case* and the *continuous case*.

In the discrete case the underlying tree is a finite, rooted plane tree T . To each vertex v in T with outdegree $d_T^+(v)$ there is associated a decoration $B(v)$ which is a metric space with one distinguished point called the inner root and $d_T^+(v)$ number of distinguished

*The second author acknowledges support from the Icelandic Research Fund, Grant Number: 185233-051, and is grateful for the hospitality at Université Paris-Sud Orsay. The first author was a post-doctoral fellow at the University of British Columbia when most of the research that led to this paper was conducted.

[†]MODALX, UPL, Univ. Paris Nanterre, CNRS, F92000 Nanterre, France. E-mail: dsenizer@parisnanterre.fr

[‡]University of Iceland. E-mail: sigurdur@hi.is

[§]Vienna University of Technology. E-mail: benedikt.stufler@tuwien.ac.at

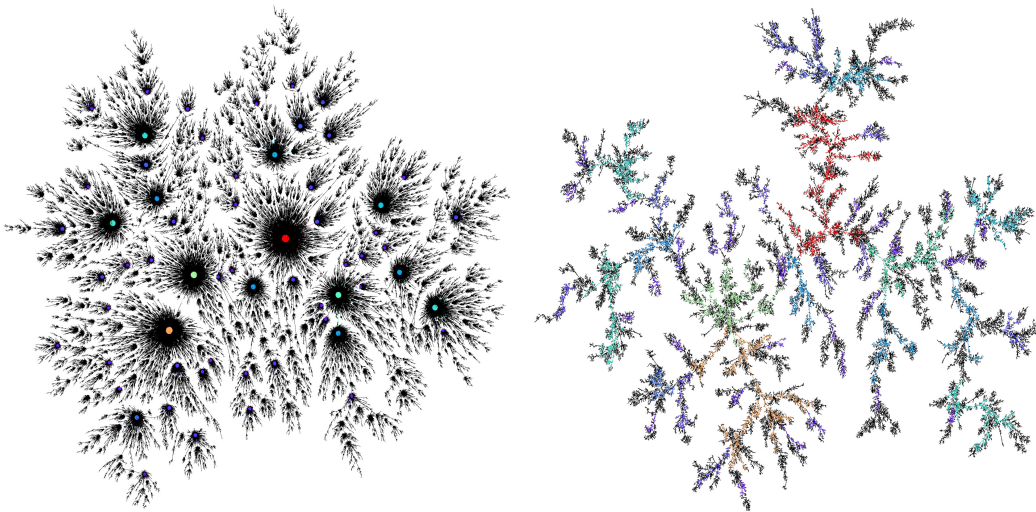


Figure 1: The right side illustrates a decorated α -stable tree for $\alpha = 5/4$. The decorations are Brownian trees (2-stable trees) rescaled according to the “width” of branchpoints. The left side shows the corresponding α -stable tree before blowing up its vertices. Colours mark vertices with high width and their corresponding decorations.

points called the outer roots. In applications, the decorations $B(v)$ may e.g. be finite graphs and they will only depend on the outdegree of v . We will thus sometimes write $B(v) = B_{d_T^+(v)}$. The decorated tree is constructed by identifying, for each vertex v , each of the outer roots of $B(v)$ with the inner root of $B(vi)$, $i = 1, \dots, d_T^+(v)$, where vi are the children of v . This identification will either be done at random or according to some prescribed labelling. Our aim is to describe the asymptotic behaviour of *random* spaces that are constructed in this way. To this end, let $(T_n)_{n \geq 1}$ be a sequence of random rooted plane trees. The sequence is chosen in such a way that the exploration process of the random trees, properly rescaled, converges in distribution towards an excursion of an α -stable Lévy process with $\alpha \in (1, 2)$. This particular choice is motivated by applications as will be explained below. We will further assume that there is a parameter $0 < \gamma \leq 1$, which we call the *distance exponent*, such that the (possibly random) decorations B_k , with the metric rescaled by $k^{-\gamma}$, converge in distribution towards a compact metric space \mathcal{B} as $k \rightarrow \infty$, in the Gromov–Hausdorff sense.

In the continuous case, we aim to construct a decorated tree which is the scaling limit of the sequence of discrete decorated trees described above. The underlying tree is chosen as the α -stable tree \mathcal{T}_α , with $\alpha \in (1, 2)$, which is a random compact real tree. One important feature of the α -stable tree is that it has countably many vertices of infinite degree. The size of such a vertex is measured by a number called its *width* which plays the same role as the degree in the discrete setting. The continuous decorated tree is constructed in a way similar as in the discrete case, where now each vertex v of infinite degree in \mathcal{T}_α is independently replaced by a decoration \mathcal{B}_v distributed as \mathcal{B} , in which distances are rescaled by width of v to the distance exponent γ . The gluing points are specified by singling out an inner root and sampling outer roots according to a measure μ and the decorations are then glued together along the tree structure. See Figure 1 for an example of a decorated tree. The precise construction is somewhat technical and is described in more detail in Section 2.2. In order for the construction to yield a compact metric space we require some control on the size of the decorations \mathcal{B}_v . More precisely,

we will assume that $\gamma > \alpha - 1$ and that there exists $p > \frac{\alpha}{\gamma}$ such that $\mathbb{E}[\text{diam}(\mathcal{B})^p] < \infty$. Under some further suitable conditions on the decorations and the underlying tree which are stated in Section 3.1 we arrive at the following main results:

- *Construction of the decorated α -stable tree:* The decorated α -stable tree defined above is a well defined random compact metric space, see Section 4 and Theorem 6.6.
- *Invariance principle:* The sequence of discrete random decorated trees, with a properly rescaled metric, converges in distribution towards a decorated α -stable tree in the Gromov–Hausdorff–Prokhorov sense. See Theorem 5.1 for the precise statement.
- *Hausdorff dimension:* Apart from degenerate cases, the Hausdorff dimension of the leaves of the decorated α -stable tree equals α/γ , see Theorem 6.6. It follows that if the decorations a.s. have a Hausdorff dimension $d_{\mathcal{B}}$ then the Hausdorff dimension of the decorated α -stable tree equals $\max\{d_{\mathcal{B}}, \alpha/\gamma\}$.

Remark 1.1. A heuristic argument for the condition $\alpha - 1 < \gamma$ is the following. If (T_n) is in the domain of attraction of \mathcal{T}_α then the maximum degree of a vertex in T_n is of the order $n^{1/\alpha}$ and the height of T_n is of the order $n^{\frac{\alpha-1}{\alpha}}$ (up to slowly varying functions). A decoration on a large vertex will then have a diameter of the order $n^{\gamma/\alpha}$. If $\alpha - 1 < \gamma$, the diameter of the decorations will thus dominate the height of T_n and the decorated tree will have height of the order $n^{\gamma/\alpha}$, coming only from the decorations. Furthermore, the decorations will “survive” in the limit since the distances of the decorated tree will be rescaled according to the inverse of its height which is of the order of the diameter of the decorations.

In the case $\alpha - 1 > \gamma$, the decorations will have a diameter which is of smaller order than the height of T_n . This means that the proper rescaling of the distances in the decorated tree should be the inverse of the height of T_n , i.e. $n^{-\frac{\alpha-1}{\alpha}}$ which would contract the decorations to points. We therefore expect that in this case, the scaling limit of the decorated tree is equal to a constant multiple of \mathcal{T}_α which has a Hausdorff dimension $\alpha/(\alpha - 1)$. We do not go through the details in the current work.

We do not know what happens exactly at the value $\alpha - 1 = \gamma$ at which there is a phase transition in the geometry and the Hausdorff dimension of the scaling limit. We leave this as an interesting open case.

1.1 Relations to other work

The only current explicit example of decorated α -stable trees are the α -stable looptrees which were introduced by Curien and Kortchemski [10]. The looptrees are constructed by decorating the α -stable trees by deterministic circles. The authors further showed, in a companion paper [11], that the looptrees appear as boundaries of percolation clusters on random planar maps and there are indications that they play a more general role in statistical mechanical models on random planar maps [31]. Looptrees have also been shown to arise as limits of various models of discrete decorated trees such as discrete looptrees, dissections of polygons [10] and outerplanar maps [35]. In these cases the discrete decorations B_k are graph cycles, chord restricted dissections of a polygon and dissections of a polygon respectively.

The idea of decorating α -stable trees by more general decorations was originally proposed by Kortchemski and Marzouk in [25]:

“Note that if one starts with a stable tree and explodes each branchpoint by simply gluing inside a “loop”, one gets the so-called stable looptrees.... More generally, one can imagine exploding branchpoints in stable trees and

glue inside any compact metric space equipped with a homeomorphism with $[0, 1]$.”

They mention explicitly an example where an α_1 -stable tree would be decorated by an α_2 -stable tree and so on and ask the question what the Hausdorff dimension of that object would be. We give a partial answer to their question in Section 7.1. See Figure 1 for a simulation where $\alpha_1 = 5/4$ and $\alpha_2 = 2$.

Archer studied a simple random walk on decorated Bienaymé–Galton–Watson trees in the domain of attraction of the α -stable tree [3]. Although her results do not explicitly involve constructing a Gromov–Hausdorff limit of the decorated trees, they do require an understanding of their global metric properties. We refer further to her work for examples of decorated trees, some of which we do not discuss in the current work.

Sénizergues has developed a general method for gluing together sequences of random metric spaces [32] and introduced a framework where metric spaces are glued together along a tree structure [34]. Section 6 relies heavily on the results of the latter.

Let us also mention that Stufler has studied local limits and Gromov–Hausdorff limits of decorated trees which are in the domain of attraction of the Brownian continuum random tree (equivalently the 2-stable tree) [37, 38]. The Brownian tree has no vertices of infinite degree and thus the rescaled decorations contract to points in the limit which only changes the metric of the decorated tree by a multiplicative factor. Its Gromov–Hausdorff limit then generically turns out to be a multiple of the Brownian tree itself (see [38, Theorem 6.60]).

1.2 Outline

The paper is organized as follows. In Section 2 we recall some important definitions. In Section 3 the construction of discrete decorated trees is provided and the required conditions on the underlying tree and decorations are stated. In Section 4 we give the definition of the continuous decorated trees and in Section 5 we state and prove the invariance principle, namely that the properly rescaled discrete decorated trees converge towards the continuous decorated trees in the Gromov–Hausdorff–Prokhorov sense. An alternative construction of the continuous decorated tree is given in Section 6 which allows us to prove that it is a compact metric space and gives a way to calculate its Hausdorff dimension. Section 7 provides several applications of our results. Finally, the Appendix is devoted to proving that the conditions on the discrete trees stated in Section 3 are satisfied for natural models of random trees, in particular Bienaymé–Galton–Watson trees which are in the domain of attraction of α -stable trees, with $\alpha \in (1, 2)$.

Notation

We let $\mathbb{N} = \{1, 2, \dots\}$ denote the set of positive integers. We usually assume that all considered random variables are defined on a common probability space whose measure we denote by \mathbb{P} . All unspecified limits are taken as n becomes large, possibly taking only values in a subset of the natural numbers. We write \xrightarrow{d} and \xrightarrow{p} for convergence in distribution and probability, and $\stackrel{d}{=}$ for equality in distribution. An event holds with high probability, if its probability tends to 1 as $n \rightarrow \infty$. We let $O_p(1)$ denote an unspecified random variable X_n of a stochastically bounded sequence $(X_n)_n$, and write $o_p(1)$ for a random variable X_n with $X_n \xrightarrow{p} 0$. The following list summarizes frequently used terminology.

$d_T^+(v)$	outdegree of vertex v in a tree T .
$ T $	the number of vertices of a tree T .
B	decoration in the discrete setting.
B_k	decoration of size k in the discrete setting.
d_k	metric on B_k .
$\mathbf{d}_k = k^{-\gamma} d_k$	rescaled metric on B_k .
\mathcal{B}	decoration in the continuous case.
\mathcal{B}_t	decoration in the continuous case indexed by $t \in [0, 1]$.
$\text{diam}(\mathcal{B})$	the diameter of the metric space \mathcal{B} .
T	discrete tree.
T_n	discrete tree of size n .
\mathcal{T}	real tree.
\mathcal{T}_α	the α -stable tree.
B	set of branchpoints of \mathcal{T}_α .
$X^{(\alpha)}$	a spectrally positive, α -stable Lévy process.
$X^{\text{exc},(\alpha)}$	an excursion of a spectrally positive, α -stable Lévy process.

2 Preliminaries

In this section we collect some definitions and refer to the key background results required in the current work. We start by introducing plane trees, then real trees and the α -stable tree and finally we recall the definition of the Gromov–Hausdorff–Prokhorov topology which is the topology we use to describe the convergence of compact measured metric spaces.

2.1 Plane trees

A rooted plane tree is a subgraph of the infinite Ulam–Harris tree, which is defined as follows. Its vertex set, \mathbb{U} , is the set of all finite sequences

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

where $\mathbb{N} = \{1, 2, \dots\}$ and $\mathbb{N}^0 = \{\emptyset\}$ contains the empty sequence denoted by \emptyset . If v and w belong to \mathbb{U} , their concatenation is denoted by vw . The edge set consists of all pairs (v, vi) , $v \in \mathbb{U}$, $i \in \mathbb{N}$. The vertex vi is said to be the i -th child of v and v is called its parent. In general, a vertex vw is said to be a descendant of v if $w \neq \emptyset$, and in that case v is called its ancestor. This ancestral relation is denoted by $v \prec w$. We write $v \preceq w$ if either $v = w$ or $v \prec w$. We denote the most recent common ancestor of v and w by $v \wedge w$, with the convention that $v \wedge w = v$ if $v \preceq w$ (and $v \wedge w = w$ if $w \preceq v$).

A rooted plane tree T is a finite subtree of the Ulam–Harris tree such that

1. \emptyset is in T , and is called its root.
2. If v is in T then all its ancestors are in T .
3. If v is in T there is a number $d_T^+(v) \geq 0$, called the outdegree of v , such that vi is in T if and only if $1 \leq i \leq d_T^+(v)$.

In the following we will only consider rooted plane trees. For simplicity, we will refer to them as trees. The number of vertices in a tree T will be denoted by $|T|$ and we denote them, listed in lexicographical order, by $v_1(T), v_2(T), \dots, v_{|T|}(T)$. When it is clear from the context, we may leave out the argument T for notational ease.

To the tree T we associate its Łukasiewicz path $(W_k(T))_{0 \leq k \leq |T|}$ by $W_0(T) = 0$ and

$$W_k(T) = \sum_{j=1}^k (d_T^+(v_j) - 1) \tag{2.1}$$

for $1 \leq k \leq |T|$.

Throughout the paper, we will consider sequences of *random trees* $(T_n)_{n \geq 1}$. The index n will usually refer to the size of the tree in some sense, e.g. the number of vertices or the number of leaves. In Section 3.1 we state the general conditions on these random trees. However, the prototypical example to keep in mind is when T_n is a Bienaymé–Galton–Watson tree, conditioned to have n vertices. In this case the Łukasiewicz path $(W_k(T_n))_{0 \leq k \leq n}$ is a random walk conditioned on staying non-negative until step n when it hits -1 .

2.2 Real trees and the α -stable trees

Real trees may be viewed as continuous counterparts to discrete trees. Informally, they are compact metric spaces obtained by gluing together line segments without creating loops. A formal definition may e.g. be found in [14] but for our purposes we give an equivalent definition in terms of excursions on $[0, 1]$, which are continuous functions $g : [0, 1] \rightarrow \mathbb{R}$ with the properties that $g \geq 0$ and $g(0) = g(1) = 1$. For $s, t \in [0, 1]$ define $m_g(s, t) = \min_{s \wedge t \leq r \leq s \vee t} g(r)$ and define the pseudo distance

$$d_g(s, t) = g(s) + g(t) - 2m_g(s, t).$$

The real tree \mathcal{T}^g encoded by the excursion g is defined as the quotient

$$\mathcal{T}^g = [0, 1] / \{d_g = 0\}$$

endowed with the metric inherited from d_g which will still be denoted d_g by abuse of notation.

We will consider a family of random real trees \mathcal{T}_α , $\alpha \in (1, 2)$ called α -stable trees, which serve as the underlying trees in the continuous random decorated trees. A brief definition of \mathcal{T}_α is given below, but we refer to definitions and more details in [12]. Let $(X_t^{(\alpha)}; t \in [0, 1])$ be a spectrally positive α -stable Lévy process, with $\alpha \in (1, 2)$, normalized such that

$$\mathbb{E}(\exp(-\lambda X_t)) = \exp(t\lambda^\alpha) \tag{2.2}$$

for every $\lambda > 0$. Let $(X_t^{\text{exc},(\alpha)}; t \in [0, 1])$ be an excursion of $X_t^{(\alpha)}$, i.e. the process conditioned to stay positive for $t \in (0, 1)$ and to be 0 for $t \in \{0, 1\}$. This involves conditioning on an event of probability zero which may be done via approximations, see e.g. [5, Chapter VIII.3]. Define the associated *continuous height process*

$$H_t^{\text{exc},(\alpha)} = \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} \int_0^t \mathbf{1}_{\{X_s^{\text{exc}} < I_s^t + \epsilon\}} ds$$

where

$$I_s^t = \inf_{[s,t]} X^{\text{exc}}.$$

Continuity of $H_t^{\text{exc},(\alpha)}$ is not guaranteed by the above definition but there exists a continuous modification which we consider from now on [16]. The α -stable tree is the random

real tree $\mathcal{T}_\alpha := \mathcal{T}^{H^{\text{exc},(\alpha)}}$ encoded by $H^{\text{exc},(\alpha)}$. We refer to [12, 13, 14] for more details on the construction and properties of \mathcal{T}_α .

Let T_n^{BGW} be a random BGW tree with a critical offspring distribution in the domain of attraction of an α -stable law with $\alpha \in (1, 2)$. Recall that this example is the prototype for the underlying random trees we consider in the current work. Since the Łukasiewicz path $W_n(T_n^{\text{BGW}})$ is in this case a conditioned random walk, it follows that by properly rescaling $(W_{[nt]}(T_n^{\text{BGW}}); t \in [0, 1])$, it converges in distribution towards the excursion $(X_t^{\text{exc},(\alpha)}; t \in [0, 1])$ in the Skorohod space $D([0, 1], \mathbb{R})$, [5, Chapter VIII]. This observation is the main idea in proving that T_n^{BGW} , with a properly rescaled graph metric, converges towards \mathcal{T}_α in the Gromov–Hausdorff sense [12, 14]. The pointed GHP-distance may and will be generalized in a straightforward manner by adding more points and more measures.

2.3 The Gromov-Hausdorff-Prokhorov topology

Let \mathbb{M} be the set of all compact measured metric spaces modulo isometries. An element $(X, d_X, \nu_X) \in \mathbb{M}$ consists of the space X , a metric d_X and a Borel-measure ν_X . In particular, \mathbb{M} contains finite graphs (e.g. with a metric which is a multiple of the graph metric) and real trees. The Gromov–Hausdorff–Prokhorov (GHP) distance between two elements $(X, d_X, \nu_X), (Y, d_Y, \nu_Y) \in \mathbb{M}$ is defined by

$$d_{\text{GHP}}(X, Y) = \inf \{ d_H^Z(\phi_1(X), \phi_2(Y)) + d_P^Z(\phi_1^* \nu_X, \phi_2^* \nu_Y) \} \tag{2.3}$$

where the infimum is taken over all measure preserving isometries $\phi_1 : X \rightarrow Z$ and $\phi_2 : Y \rightarrow Z$ and metric spaces Z . Here d_H^Z is the Hausdorff distance between compact metric spaces in Z , d_P^Z is the Prokhorov distance between measures on Z and $\phi^* \nu$ denotes the pushforward of the measure ν by ϕ . We refer to [39, Chapter 27] for more details. When we are only interested in the convergence of the metric spaces without the measures, we leave out the term in (2.3) with the Prokhorov metric and refer to the metric simply as the Gromov–Hausdorff (GH) metric, denoted d_{GH} .

The following formulation of GHP-convergence may be useful in calculations and we will use it repeatedly in the following sections. If (X, d_X) and (Y, d_Y) are metric spaces and $\epsilon > 0$, a function $\phi_\epsilon : X \rightarrow Y$ is called an ϵ -isometry if the following two conditions hold:

1. (Almost isometry)

$$\sup_{x_1, x_2 \in X} |d_Y(\phi_\epsilon(x_1), \phi_\epsilon(x_2)) - d_X(x_1, x_2)| < \epsilon.$$

2. (Almost surjective) For every $y \in Y$ there exists $x \in X$ such that

$$d_Y(\phi_\epsilon(x), y) < \epsilon.$$

If (X_n, d_n, ν_n) is a sequence of compact measured metric spaces which converge towards (X, d, ν) in the GHP topology then it is equivalent to the existence of ϵ_n -isometries $\phi_{\epsilon_n} : X_n \rightarrow X$, where $\epsilon_n \rightarrow 0$, such that

$$d_H^X(\phi_{\epsilon_n}(X_n), X) \rightarrow 0 \quad \text{and} \quad d_P^X(\phi_{\epsilon_n}^* \nu_n, \nu) \rightarrow 0 \tag{2.4}$$

as $n \rightarrow \infty$. When the GH topology is considered, only the first convergence of these two is assumed.

In most of the applications we will consider *pointed* metric spaces (X, d_X, ρ_X) where ρ_X is a distinguished element called a *point* or a *root*. The GHP-distance is adapted to

Decorated stable trees

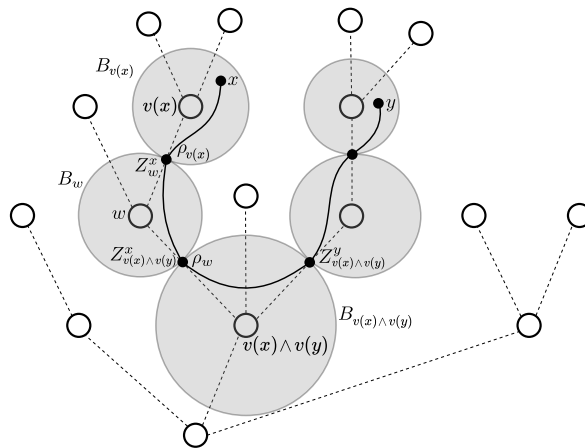


Figure 2: Definition of the metric d_n^{dec} .

this by requiring that the isometries in (2.3) and the ϵ_n -isometries in (2.4), send root to root. In this case we speak of the *pointed* GHP-distance. In the same vein, we will sometimes deal with metric spaces that carry some extra finite measure. In this case we just add a term in (2.3) that corresponds to the Prokhorov distance computed between the extra measures and modify (2.4) accordingly.

3 Construction in the discrete

In this section we define the discrete decorated trees and state the sufficient conditions on the decorations and the underlying trees for our main results to hold.

A *decoration* of size k is a compact metric space B equipped with a metric d . We distinguish a vertex ρ and call it the *internal root* of B and furthermore, we label k vertices (not necessarily different) and call them the *external roots* of B . This labeling is described by a function $\ell : \{1, \dots, k\} \rightarrow B$. Note that we explicitly allow external roots to coincide with the internal root. The space B is equipped with a finite Borel measure ν . We will use the notation (B, d, ρ, ℓ, ν) for a decoration. In practice, B will often be a finite graph, d the graph metric, and ν the counting measure on the vertex set. The labeled vertices will often be the boundary of B in some sense, e.g. the leaves of a tree or the vertices on the boundary of a planar map. Let $(T_n)_{n \geq 1}$ be a sequence of random trees and $(\tilde{B}_k, \tilde{\rho}_k, \tilde{d}_k, \tilde{\ell}_k, \tilde{\nu}_k)_{k \geq 0}$ a sequence of random decorations indexed by their size. Conditionally on T_n , for each $v \in T_n$, sample independently a decoration $(B_v, \rho_v, d_v, \ell_v, \nu_v)$ distributed as $(\tilde{B}_{d^+(v)}, \tilde{\rho}_{d^+(v)}, \tilde{d}_{d^+(v)}, \tilde{\ell}_{d^+(v)}, \tilde{\nu}_{d^+(v)})$. We aim to glue the decorations to each other along the structure of T_n which is informally explained as follows. Consider the disjoint union

$$T_n^{*, \text{dec}} = \bigsqcup_{v \in T_n} B_v$$

of decorations on which is defined an equivalence relation \sim , such that for each v in T_n , and each $1 \leq i \leq d^+(v)$ we have $\ell_v(i) \sim \rho_{v_i}$. In other words, the i -th external root of B_v is glued to the internal root of B_{v_i} . The precise construction is given below.

We first define a pseudo-metric $d_n^{*, \text{dec}}$ on $T_n^{*, \text{dec}}$. For each $x \in T_n^{*, \text{dec}}$ we let $v(x)$ denote the unique vertex in T_n such that $x \in B_{v(x)}$. If w is a vertex in T_n such that $w \preceq v(x)$, let

$Z_w^x \in B_w$ be defined by

$$Z_w^x = \begin{cases} \ell_w(i) & \text{if } wi \preceq v(x), \\ x & \text{if } w = v(x). \end{cases}$$

In order to clarify the first case on the right hand side, note that if $w \prec v(x)$ then there is a unique child wi , $1 \leq i \leq d_{T_n}^+(w)$, of w in T_n such that $wi \preceq v(x)$, and we set $Z_w^x = \ell_w(i)$.

Now, let $x, y \in T_n^{*,\text{dec}}$ and first assume that $v(x) \prec v(y)$ and set

$$d_n^{*,\text{dec}}(x, y) = d_n^{*,\text{dec}}(y, x) = d_{v(x)}(x, Z_{v(x)}^y) + \sum_{v(x) \prec w \preceq y} d_w(\rho_w, Z_w^y).$$

Otherwise, if $v(x)$ and $v(y)$ are not in an ancestral relation, set

$$\begin{aligned} d_n^{*,\text{dec}}(x, y) &= d_{v(x) \wedge v(y)}(Z_{v(x) \wedge v(y)}^x, Z_{v(x) \wedge v(y)}^y) + d_n^{*,\text{dec}}(Z_{v(x) \wedge v(y)}^x, x) \\ &\quad + d_n^{*,\text{dec}}(Z_{v(x) \wedge v(y)}^y, y). \end{aligned} \tag{3.1}$$

These definitions are clarified in Fig. 2. Finally, the decorated tree is the quotient

$$T_n^{\text{dec}} = T_n^{*,\text{dec}} / \sim \tag{3.2}$$

where $x \sim y$ if and only if $d_n^{*,\text{dec}}(x, y) = 0$, and the metric induced by $d_n^{*,\text{dec}}$ on the quotient is denoted by d_n^{dec} .

Define a measure ν_n^{dec} , on T_n^{dec} by

$$\nu_n^{\text{dec}} = \sum_{v \in T_n} \nu_v.$$

Here the mass of a gluing point is the sum of the masses of all vertices of its equivalence class. Finally, root T_n^{dec} at ρ_\emptyset . We view the the decorated tree as a random, compact, rooted and measured metric space

$$(T_n^{\text{dec}}, d_n^{\text{dec}}, \rho_\emptyset, \nu_n^{\text{dec}}).$$

3.1 Conditions on the trees and decorations

Assume in this subsection that $\alpha \in (1, 2)$ is fixed and that $(T_n)_n$ is a sequence of random trees with decorations $(B_v, d_v, \rho_v, \ell_v, \nu_v)_{v \in T_n}$ sampled independently according to $(\tilde{B}_k, \tilde{\rho}_k, \tilde{d}_k, \tilde{\ell}_k, \tilde{\nu}_k)_{k \geq 0}$. Let μ_k be a measure on \tilde{B}_k defined by

$$\mu_k = \frac{1}{k} \sum_{j=1}^k \delta_{\tilde{\ell}_k(j)}.$$

We let $v_1, \dots, v_{|T_n|}$ be the vertices of T_n listed in depth-first-search order and for all $1 \leq k \leq |T_n|$, we let

$$M_k = \sum_{i=1}^k \nu_{v_i}(B_{v_i})$$

be the sum of the total masses of the k first decorations in depth-first order. We impose the following conditions on the decorations (D), the trees (T) and on both trees and decorations (B).

Conditions on the decorations

D1) (*Pointed GHP limit*). There are $\beta, \gamma > 0$, a compact metric space (\mathcal{B}, d) with a point ρ , a Borel probability measure μ and a finite Borel measure ν such that

$$\left(\tilde{B}_k, \tilde{\rho}_k, k^{-\gamma} \tilde{d}_k, \mu_k, k^{-\beta} \tilde{\nu}_k \right) \rightarrow (\mathcal{B}, \rho, d, \mu, \nu)$$

as $k \rightarrow \infty$, weakly in the pointed Gromov–Hausdorff–Prokhorov sense.

D2) (*Moments of the diameter of the limit*). The limiting block $(\mathcal{B}, \rho, d, \mu, \nu)$ satisfies $\mathbb{E}[\text{diam}(\mathcal{B})^p] < \infty$, for some $p > \frac{\alpha}{\gamma}$.

D3) (*Limit of the expected mass*). It holds that

$$k^{-\beta} \cdot \mathbb{E} \left[\tilde{\nu}_k(\tilde{B}_k) \right] \xrightarrow[k \rightarrow \infty]{} \mathbb{E}[\nu(\mathcal{B})] < \infty.$$

Conditions on the trees

T1) (*Symmetries of subtrees*). The distribution of T_n is invariant under the permutation of any set of siblings.

T2) (*Invariance principle for the Łukasiewicz path*). There is a sequence $b_n \geq 0$ such that

$$(b_n^{-1} W_{\lfloor |T_n| t \rfloor}(T_n))_{0 \leq t \leq 1} \xrightarrow{d} X^{\text{exc},(\alpha)}$$

where $(W_k(T_n))_{0 \leq k \leq |T_n|}$ is the Łukasiewicz path of T_n and $X^{\text{exc},(\alpha)}$ is an excursion of the spectrally positive α -stable Lévy-process defined by (2.2).

Conditions on both

Let γ and b_n be as in conditions T and D in the previous sections.

B1) (*Small decorations do not contribute*). For any $\epsilon > 0$,

$$\mathbb{P} \left(\sup_{v \in T_n} \left\{ b_n^{-\gamma} \sum_{w \leq v} \text{diam}(B_w) \mathbf{1}_{\{d_{T_n}^+(w) \leq \delta b_n\}} \right\} > \epsilon \right) \xrightarrow[\delta \rightarrow 0]{} 0,$$

uniformly in $n \geq 1$ for which T_n is well-defined.

B2) (*Measure is spread out, case $\beta \leq \alpha$*). We assume that we have the following uniform convergence in probability, for some sequence (a_n) ,

$$\left(\frac{M_{\lfloor |T_n| t \rfloor}}{a_n} \right)_{t \in [0,1]} \xrightarrow[n \rightarrow \infty]{} (t)_{0 \leq t \leq 1}. \tag{3.3}$$

B3) (*Measure is on the blocks, case $\beta > \alpha$*). For any $\epsilon > 0$ we have

$$\limsup_{n \rightarrow \infty} \mathbb{P} \left(b_n^{-\beta} \sum_{i=1}^{|T_n|} \nu_{v_i}(B_{v_i}) \mathbf{1}_{\{d_{T_n}^+(v_i) \leq \delta b_n\}} > \epsilon \right) \xrightarrow[\delta \rightarrow 0]{} 0. \tag{3.4}$$

In this case we write $a_n = b_n^\beta$.

3.2 Reduction to the case of shuffled external roots

Given our Hypothesis T1 on the tree T_n , we can always shuffle the external roots of the blocks that we use without changing the law of the object T_n^{dec} that we construct.

More precisely, given the law of $(\tilde{B}_k, \tilde{d}_k, \tilde{\rho}_k, \tilde{\ell}_k, \tilde{\nu}_k)$ for any $k \geq 0$ we can define the block with shuffled external roots $(\hat{B}_k, \hat{d}_k, \hat{\rho}_k, \hat{\ell}_k, \hat{\nu}_k)$ as

$$(\hat{B}_k, \hat{d}_k, \hat{\rho}_k, \hat{\nu}_k) = (\tilde{B}_k, \tilde{d}_k, \tilde{\rho}_k, \tilde{\nu}_k) \quad \text{and} \quad \hat{\ell}_k(i) = \tilde{\ell}_k(\sigma(i)),$$

where σ is a uniform permutation of length k , independent of everything else. Then, assuming that Hypothesis T1 holds, the law of the object \hat{T}_n^{dec} constructed with underlying tree T_n and decorations $(\hat{B}_k, \hat{\rho}_k, \hat{d}_k, \hat{\ell}_k, \hat{\nu}_k)_{k \geq 1}$ is the same as the that of the object T_n^{dec} constructed with underlying tree T_n and decorations $(\tilde{B}_k, \tilde{d}_k, \tilde{\rho}_k, \tilde{\ell}_k, \tilde{\nu}_k)_{k \geq 0}$.

We will require the following lemma which informally states that the sequence of shuffled external roots in a decoration B_k converges weakly to an i.i.d. sequence of external roots in the GHP-limit \mathcal{B} of B_k

Lemma 3.1. *Let $(\tilde{B}_k, \tilde{d}_k, \tilde{\rho}_k, \tilde{\ell}_k, \tilde{\nu}_k)_{k \geq 1}$ be a (deterministic) sequence of decorations which converges to a decoration $(\mathcal{B}, \rho, d, \mu, \nu)$ in the GHP-topology. Let Y_1, Y_2, \dots be a sequence of independent points in \mathcal{B} sampled according to μ . Then there exists a sequence $\phi_k : \tilde{B}_k \rightarrow \mathcal{B}$ of ϵ_k -isometries, with $\epsilon_k \rightarrow 0$, such that*

$$(\phi_k(\hat{\ell}_k(1)), \phi_k(\hat{\ell}_k(2)) \dots, \phi_k(\hat{\ell}_k(k)), \rho, \rho, \dots) \xrightarrow{d} (Y_1, Y_2, \dots)$$

as $k \rightarrow \infty$.

Proof. First note that

$$\left\| \mu_k^{\otimes k}, \mathcal{L}(\hat{\ell}_k) \right\|_{TV} \rightarrow 0,$$

as $k \rightarrow \infty$, where $\mathcal{L}(\hat{\ell}_k)$ denotes the law of the uniformly shuffled labels $\hat{\ell}_k$. Thus, asymptotically, we may work with the measure $\mu_n^{\otimes k}$ instead of the shuffled labels. The GHP-convergence guarantees existence of ϵ_k -isometries $\phi_k : \tilde{B}_k \rightarrow \mathcal{B}$ with $\epsilon_k \rightarrow 0$, such that the pushforward $\phi_k^* \mu_k$ converges to μ in distribution. It follows that $\phi_k^* \mu_k^{\otimes N} \rightarrow \mu^{\otimes N}$ in distribution which concludes the proof. \square

In the rest of the paper, we are always going to assume that the external roots of the decorations that we use have been shuffled as above. Note that there can be situations where there is a natural ordering of the external roots e.g. when the decorations are oriented circles, maps with an oriented boundary, trees encoded by contour functions and more.

3.3 The case of Bienaymé–Galton–Watson trees

One important case of application is when the tree T_n is taken to be a Bienaymé–Galton–Watson (BGW) tree in the domain of attraction of an α -stable tree, conditioned on having n vertices. If we denote by μ a reproduction law on \mathbb{Z}_+ , this corresponds to having $\sum_{i=0}^{\infty} i\mu(i) = 1$ and the function

$$x \mapsto \frac{1}{\mu([x, \infty))}$$

to be α -regularly varying, see the Appendix for a discussion on regularly varying functions. Note that then T_n is possibly not defined for all $n \geq 1$; in that case we only consider values of n for which it is.

Condition T1 is satisfied for any BGW tree and T2 holds, thanks to results of Duquesne and Le Gall [12, 14], for a sequence $(b_n)_{n \geq 1}$ that depends on the tail behaviour of $\mu([n, \infty))$. In the rest of this section, we assume that $\alpha \in (1, 2)$ and $\gamma > \alpha - 1$ are fixed and that we are working with BGW trees with reproduction distribution μ as above and corresponding sequence $(b_n)_{n \geq 1}$.

The results below concern properties of the diameter and the measure on the decorations and they guarantee that the remaining conditions in Section 3.1 are satisfied. Their proofs are given in the Appendix. We first state a result that ensures that small decorations don't contribute to the macroscopic distances.

Proposition 3.2. *Under the following moment condition on the diameter of the decorations*

$$\sup_{k \geq 0} \mathbb{E} \left[\left(\frac{\text{diam}(\tilde{B}_k)}{k^\gamma} \right)^m \right] < \infty,$$

for some $m > \frac{4\alpha}{2\gamma+1-\alpha}$, condition B1 is satisfied.

Remark 3.3. In [3, Remark 1.5], it is stated that in this context $b_n^{-\gamma} \cdot \text{diam}(T_n^{\text{dec}})$ is tight under an assumption that corresponds to having $m > \frac{\alpha}{\alpha-1}$ in the above statement. Note that neither of those assumptions are optimal since $\frac{\alpha}{\alpha-1}$ and $\frac{4\alpha}{2\gamma-(\alpha-1)}$ are not always ordered in the same way for different values of α and γ . In particular, our bound does not blow up for a fixed γ and $\alpha \rightarrow 1$.

Now we state another result that ensures that either condition B2 or B3 is satisfied in this setting. We let D denote a random variable that has law given by the reproduction distribution μ of the BGW tree.

Lemma 3.4. *Assuming that D1 holds for some value $\beta > 0$ and that*

$$\sup_{k \geq 0} \mathbb{E} \left[\left(\frac{\nu_k(\tilde{B}_k)}{k^\beta} \right) \right] < \infty,$$

then condition B2 or B3 is satisfied, depending on the value of β :

- If $\beta < \alpha$ then B2 holds with $a_n := n \cdot \mathbb{E}[\nu_D(B_D)]$.
- If $\beta = \alpha$ then B2 holds with

$$a_n := \begin{cases} n \cdot \mathbb{E}[\nu_D(B_D)] & \text{if } \mathbb{E}[D^\alpha] < \infty, \\ n \cdot \mathbb{E}[\nu(\mathcal{B})] \cdot \mathbb{E}[D^\alpha \mathbf{1}_{\{D \leq b_n\}}] & \text{if } \mathbb{E}[D^\alpha] = \infty, \end{cases} \quad (3.5)$$

where in the second case, we further assume that $\sup_{k \geq 0} \mathbb{E} \left[\left(\frac{\nu_k(\tilde{B}_k)}{k^\beta} \right)^{1+\eta} \right] < \infty$ for some $\eta > 0$.

- If $\beta > \alpha$ then B3 holds with $a_n = b_n^\beta$.

Remark 3.5. Both those results are stated for BGW trees conditioned on having exactly n vertices. In fact, it is possible to use arguments from [23, Section 4, Section 5] to deduce that the same results hold for BGW trees conditioned on having $\lfloor \mu(\{0\})n \rfloor$ leaves. This is actually the setting that we need for most of our applications.

4 Construction in the continuous

Let $(X_t^{\text{exc},(\alpha)}; t \in [0, 1])$ be an excursion of an α -stable spectrally positive Lévy process, with $\alpha \in (1, 2)$ and denote the corresponding α -stable tree by \mathcal{T}_α . Denote the projection from $[0, 1]$ to \mathcal{T}_α by \mathbf{p} . Recall the definition

$$I_s^t = \inf_{[s,t]} X^{\text{exc},(\alpha)}.$$

We define a partial order \preceq on $[0, 1]$ as follows: for every $s, t \in [0, 1]$

$$s \preceq t \text{ if } s \leq t \text{ and } X_{s^-}^{\text{exc},(\alpha)} \leq I_s^t. \tag{4.1}$$

For $s, t \in [0, 1]$ with $s \preceq t$, define

$$x_s^t = I_s^t - X_{s^-}^{\text{exc}} \tag{4.2}$$

and let

$$u_s^t := \frac{x_s^t}{\Delta_s} \tag{4.3}$$

where $\Delta_s = X_s^{\text{exc},(\alpha)} - X_{s^-}^{\text{exc},(\alpha)}$ is the jump of $X_s^{\text{exc},(\alpha)}$ at s . If s is not a jump time of the process, we simply let $u_s^t := 0$.

Let us denote $\mathbb{B} = \{v \in [0, 1] \mid \Delta_v > 0\}$, which is in one-to-one correspondence with the set of branch points of \mathcal{T}_α . Almost surely, these points in the tree all have an infinite degree, meaning that for any $v \in \mathbb{B}$, the space $\mathcal{T} \setminus \{\mathbf{p}(v)\}$ has a countably infinite number of connected components. For any $v \in \mathbb{B}$, we let

$$\mathcal{A}_v := \{u_v^t \mid t \in [0, 1] \text{ and } \exists s \in [0, 1], v \prec s \prec t\}.$$

This set \mathcal{A}_v is in one-to-one correspondence with the connected components of $\mathcal{T} \setminus \{\mathbf{p}(v)\}$ above $\mathbf{p}(v)$, meaning not the one containing the root.

Conditionally on this, let $((\mathcal{B}_v, \rho_v, d_v, \mu_v, \nu_v))_{v \in \mathbb{B}}$ be random measured metric spaces, sampled independently with the distribution of $(\mathcal{B}, \rho, d, \mu, \nu)$ and for every $v \in \mathbb{B}$, let $(Y_{v,a})_{a \in \mathcal{A}_v}$ be i.i.d. with law μ_v . For $s \in [0, 1] \setminus \mathbb{B}$, let $\mathcal{B}_s = \{s\}$. Define for every $v \in \mathbb{B}$ the rescaled distance

$$\delta_v = \Delta_v^\gamma \cdot d_v,$$

and for any $t \succeq v$:

$$Z_v^t = \begin{cases} Y_{v,u_v^t} & \text{if } u_v^t \in \mathcal{A}_v, \\ \rho_v & \text{otherwise.} \end{cases} \tag{4.4}$$

Now we are ready to define a pseudo-distance d on the set $\mathcal{T}_\alpha^{*,\text{dec}} = \bigsqcup_{s \in [0,1]} \mathcal{B}_s$. Before doing that, let us define, for any $v, w \in \mathbb{B}$ such that $v \preceq w$, and any $x \in \mathcal{B}_w$,

$$Z_v^x = \begin{cases} Z_v^w & \text{if } v \prec w \\ x & \text{if } v = w. \end{cases}$$

We also extend the genealogical order on the whole space $\mathcal{T}_\alpha^{*,\text{dec}}$ by declaring that for any $v \in \mathbb{B}, x \in \mathcal{B}_v$, we have $v \preceq x \preceq v$, (thus making it a preorder, instead of an order relation). We also extend the definition of \wedge by saying that $a \wedge b$ is always chosen among $[0, 1]$.

Then we define $d_0(a, b)$ for $a \preceq b$:

$$d_0(a, b) = d_0(b, a) = \sum_{\substack{a \prec v \preceq b \\ v \in \mathbb{B}}} \delta_v(\rho_v, Z_v^b), \tag{4.5}$$

and for any $a, b \in \mathcal{T}_\alpha^{*,\text{dec}}$,

$$d(a, b) = d_0(a \wedge b, a) + d_0(a \wedge b, b) + \delta_{a \wedge b}(Z_{a \wedge b}^a, Z_{a \wedge b}^b),$$

We say that two elements a and b are equivalent $a \sim b$ if $d(a, b) = 0$. We quotient the space by the equivalence relation, $\mathcal{T}_\alpha^{\text{dec}} = \mathcal{T}_\alpha^{*,\text{dec}} / \sim$ and denote \mathbf{p}^{dec} the corresponding quotient map.

We can endow $\mathcal{T}_\alpha^{\text{dec}}$ with two types of measures:

- If $\beta \leq \alpha$, then we define the measure ν_β^{dec} on $\mathcal{T}_\alpha^{\text{dec}}$ as the push-forward of the Lebesgue measure on $[0, 1] \setminus \mathbb{B}$ through the projection map $\mathbf{p}^{\text{dec}} : \mathcal{T}_\alpha^{*,\text{dec}} \rightarrow \mathcal{T}_\alpha^{\text{dec}}$ onto the quotient.
- If $\beta > \alpha$, we define the measure ν_β^{dec} as (the push-forward through \mathbf{p}^{dec} of)

$$\sum_{t \in \mathbb{B}} \Delta_t^\beta \nu_t$$

which is almost surely a finite measure, provided that $\nu(\mathcal{B})$ has finite expectation, say.

Note that so far, it is not clear that the construction above yields a compact metric space (or even a metric space), as the metric d could in principle take arbitrarily large or even infinite values. This is handled by the work of Section 6. In particular, a result that is important in our approach is the following, which ensures that under some moment assumption on the diameter of the decorations the distances in the whole object are dominated by the contributions of the decorations corresponding to large jumps. This will in particular be useful in the next section.

Lemma 4.1. *Assume that $\gamma > \alpha - 1$ and that condition D2 is satisfied. Then,*

$$\sup_{s \in [0,1]} \left\{ \sum_{t < s} \Delta_t^\gamma \text{diam}(\mathcal{B}_t) \mathbf{1}_{\{\Delta_t \leq \delta\}} \right\} \rightarrow 0 \tag{4.6}$$

as $\delta \rightarrow 0$ in probability.

5 Invariance principle

In this section we state and prove one of the main results of the paper, namely that under the conditions stated in Section 3.1, the discrete decorated trees with a properly rescaled graph distance, converge towards α -stable decorated trees.

Theorem 5.1. *Let $(T_n)_{n \geq 1}$ be a sequence of random trees and $(B_v, \rho_v, d_v, \ell_v, \nu_v)_{v \in T_n}$ be random decorations on T_n sampled independently according to $(\tilde{B}_k, \tilde{\rho}_k, \tilde{d}_k, \tilde{\ell}_k, \tilde{\nu}_k)_{k \geq 0}$. Assume that conditions D, T and B in Section 3.1 are satisfied for some exponents α, β, γ and denote the weak limit of $(\tilde{B}_k, \tilde{\rho}_k, \tilde{d}_k, \tilde{\ell}_k, \tilde{\nu}_k)_{k \geq 0}$ in the sense of D1 by $(\mathcal{B}, \rho, d, \mu, \nu)$. Suppose that $\alpha - 1 < \gamma$ and let $\mathcal{T}_\alpha^{\text{dec}}$ be the α -stable tree \mathcal{T}_α , decorated according to $(\mathcal{B}, \rho, d, \mu, \nu)$ with distance exponent γ . Then*

$$(T_n^{\text{dec}}, b_n^{-\gamma} d_n^{\text{dec}}, a_n^{-1} \cdot \nu_n^{\text{dec}}) \rightarrow (\mathcal{T}_\alpha^{\text{dec}}, d, \nu_\beta^{\text{dec}})$$

in distribution in the Gromov-Hausdorff-Prokhorov topology.

Proof. At first we will describe in detail the GH-convergence of the sequence of decorated trees, leaving out the measures ν_n^{dec} , and then we give a less detailed outline of the GHP-convergence. We start by constructing, in several steps, a coupling that allows us to construct an ϵ -isometry between $(T_n^{\text{dec}}, b_n^{-\gamma} d_n^{\text{dec}})$ and $(\mathcal{T}_\alpha^{\text{dec}}, d)$ for arbitrary small ϵ .

Step 1, coupling of trees: By Condition T2, we may use Skorokhod’s representation theorem, and construct on the same probability space $X^{\text{exc},(\alpha)}$ and a sequence of random trees T_n such that

$$(b_n^{-1} W_{\lfloor T_n \rfloor} (T_n))_{0 \leq t \leq 1} \longrightarrow X^{\text{exc},(\alpha)}$$

almost surely. We will keep the same notation for all random elements on this new space. Order the vertices $(v_{n,1}, v_{n,2}, \dots)$ of T_n in non-increasing order of their degree (in case of ties use some deterministic rule) and denote their positions in the lexicographical

order on the vertices of T_n by $(t_{n,1}, t_{n,2}, \dots)$, in such a way that for any $k \geq 1$ we have $v_{t_{n,k}} = v_{n,k}$. Similarly, order the set B of jump times of $X^{\text{exc},\alpha}$ in decreasing order (t_1, t_2, \dots) of the sizes of the jumps (there is no tie almost surely). Then, by properties of the J_1 Skorokhod topology (see [28, Lem. 1]), it holds that for any $k \geq 1$,

$$\frac{t_{n,k}}{|T_n|} \rightarrow t_k \quad \text{and} \quad \frac{d^+(v_{n,k})}{b_n} \rightarrow \Delta_{t_k} \tag{5.1}$$

almost surely, as $n \rightarrow \infty$. Because of the way we can retrieve the genealogical order from the coding function, the genealogical order on $(v_{n,1}, v_{n,2}, \dots, v_{n,N})$ for any fixed N stabilizes for n large enough to that of (t_1, t_2, \dots, t_N) for the order \preceq defined in (4.1).

Step 2, coupling of decorations: Let us shuffle the external roots of the decorations $(\tilde{B}_k, \tilde{\rho}_k, \tilde{d}_k, \tilde{\ell}_k, \tilde{\nu}_k)_{k \geq 0}$ with independent uniform permutations as described in Section 3.2 and denote the shuffled decorations by $(\hat{B}_k, \hat{\rho}_k, \hat{d}_k, \hat{\ell}_k, \hat{\nu}_k)_{k \geq 0}$. From now on, sample $(B_v, \rho_v, d_v, \ell_v, \nu_v)_{v \in T_n}$ from the latter which, as explained before, does not affect the distribution of the decorated tree. Since for any $k \geq 1$, $d_{T_n}^+(v_{n,k}) \rightarrow \infty$ as $n \rightarrow \infty$, it holds, by Condition D1, that

$$(B_{v_{n,k}}, \rho_{v_{n,k}}, (d^+(v_{n,k}))^{-\gamma} d_{v_{n,k}}, \mu_{v_{n,k}}, (d^+(v_{n,k}))^{-\beta} \nu_{v_{n,k}}) \rightarrow (B, \rho, d, \mu, \nu)$$

weakly as $n \rightarrow \infty$. We again use Skorokhod's theorem and modify the probability space such that for each $k \geq 1$, this convergence holds almost surely and the a.s. limit will be denoted by $(\mathcal{B}_{t_k}, \rho_{t_k}, d_{t_k}, \mu_{t_k}, \nu_{t_k})$. We still keep the same notation for random elements.

For the rest of the proof, we will write \mathbf{d}_v in place of $b_n^{-1} \cdot d_v$ to lighten notation. From the convergence (5.1), and the definition of $\delta_v = \Delta_v^\gamma \cdot d_v$, we can re-express the above convergence as the a.s. convergence

$$(B_{v_{n,k}}, \rho_{v_{n,k}}, \mathbf{d}_{v_{n,k}}, \mu_{v_{n,k}}, (b_n)^{-\beta} \nu_{v_{n,k}}) \rightarrow (\mathcal{B}_{t_k}, \rho_{t_k}, \delta_{t_k}, \mu_{t_k}, \Delta_{t_k}^\beta \nu_{t_k}). \tag{5.2}$$

Step 3, coupling of gluing points: Fix a $\delta > 0$ and let $N(\delta)$ be the (finite) number of vertices $v \in B$ such that $\Delta_v > \delta$. Thanks to Step 1, we can almost surely consider n large enough so that the genealogical order on $(v_{n,1}, v_{n,2}, \dots, v_{n,N(\delta)})$ induced by T_n corresponds to that of $(t_1, t_2, \dots, t_{N(\delta)})$. Note that $X^{\text{exc},(\alpha)}$ almost surely does not have a jump of size exactly δ , and so for n large enough the vertices $(v_{n,1}, v_{n,2}, \dots, v_{n,N(\delta)})$ are exactly the vertices with degree larger than δb_n . Now recall the notation (4.3) and for any $k \in \{1, 2, \dots, N(\delta)\}$, consider the finite set

$$\{u_{t_k}^{t_i} \mid i \in \{1, 2, \dots, N(\delta)\}, t_i \succ t_k\} \subset \mathcal{A}_{t_k}$$

and enumerate its elements as $m_k(1), m_k(2), \dots, m_k(a_k)$ in increasing order. Similarly consider the set

$$\{\ell \in \mathbb{N} \mid v_{n,k} \ell \preceq v_{n,i} \text{ for some } i \in \{1, 2, \dots, N(\delta)\}\} \subset \{1, \dots, d^+(v_{n,k})\}.$$

and denote its elements j_1, j_2, \dots, j_{a_k} . Note that these two sets have the same cardinality a_k when n is large enough because of our assumptions.

For each $k \geq 1$, given $(B_{t_k}, \rho_{t_k}, d_{t_k}, \mu_{t_k}, \nu_{t_k})$ and $X^{\text{exc},(\alpha)}$ the points $(Y_{t_k,a})_{a \in \mathcal{A}_{t_k}}$ are sampled independently according to μ_{t_k} . By Lemma 3.1 and (5.2) we may find a sequence $\phi_{v_{n,k}} : B_{v_{n,k}} \rightarrow \mathcal{B}_{t_k}$ of $\epsilon_{v_{n,k}}$ -isometries, with $\epsilon_{v_{n,k}} \rightarrow 0$ as $n \rightarrow \infty$, such that

$$(\phi_{v_{n,k}}(\ell_{v_{n,k}}(j_1)), \phi_{v_{n,k}}(\ell_{v_{n,k}}(j_2)) \dots, \phi_{v_{n,k}}(\ell_{v_{n,k}}(j_{a_k}))) \xrightarrow{w} (Y_{t_k,m_k(1)}, Y_{t_k,m_k(2)}, \dots, Y_{t_k,m_k(a_k)}). \tag{5.3}$$

We modify the probability space once again, using Skorokhod's representation theorem, so that this convergence holds almost surely. When n is large enough, for k, k' such that

$v_{n,k} \prec v_{n,k'}$ we know that $v_{n,k} j_p \preceq v_{n,k'}$ and $u_{t_k}^{t_k'} = m_k(p)$ for some $p \in \{1, 2, \dots, a_k\}$ so we have $Z_{v_{n,k}}^{v_{n,k'}} = \ell_{v_{n,k}}(j_p)$ as well as $Z_{t_k}^{t_k'} = Y_{t_k, m_k(a_p)}$. Plugging this back in (5.3) that we have the convergence

$$\phi_{v_{n,k}}(Z_{v_{n,k}}^{v_{n,k'}}) \rightarrow Z_{t_k}^{t_k'}.$$

Step 4: convergence of the decorated trees

Now, construct the sequence of decorated trees T_n^{dec} from T_n and the decorations $(B_{v_{n,k}}, \rho_{v_{n,k}}, d_{v_{n,k}}, \ell_{v_{n,k}}, \nu_{v_{n,k}})_{k \geq 1}$. Similarly, construct the tree $\mathcal{T}_\alpha^{\text{dec}}$ from \mathcal{T}_α and the family of decorations $(\mathcal{B}_{t_k}, \rho_{t_k}, d_{t_k}, \mu_{t_k}, \nu_{t_k})_{k \geq 1}$. Denote the projections to the quotients by $\mathbf{p}_n^{\text{dec}}$ and \mathbf{p}^{dec} respectively. We claim that this coupling of trees, decorations and gluing points, guarantees that the convergence

$$(T_n^{\text{dec}}, b_n^{-\gamma} d_n^{\text{dec}}) \rightarrow (\mathcal{T}_\alpha^{\text{dec}}, d)$$

holds in probability for the Gromov-Hausdorff topology.

Introduce

$$r_n^\delta := b_n^{-\gamma} \cdot \sup_{v \in T_n} \left\{ \sum_{w \preceq v} \text{diam}(B_w) \mathbf{1}_{\{d_{T_n}^+(w) \leq \delta b_n\}} \right\}, \tag{5.4}$$

and also

$$r^\delta := \sup_{s \in [0,1]} \left\{ \sum_{t \prec s} \Delta_t^\gamma \text{diam}(B_t) \mathbf{1}_{\{\Delta_t \leq \delta\}} \right\}. \tag{5.5}$$

It holds, by condition B1 and Lemma 4.1 that

$$\lim_{\delta \rightarrow 0} \limsup_{n \rightarrow \infty} r_n^\delta = 0 \quad \text{and} \quad \lim_{\delta \rightarrow 0} r^\delta = 0 \tag{5.6}$$

in probability.

Let us consider the subset $\mathcal{T}_\alpha^{\text{dec}, \delta} \subset \mathcal{T}_\alpha^{\text{dec}}$ which consists only of the blocks corresponding to jumps larger than δ , i.e.

$$\mathcal{T}_\alpha^{\text{dec}, \delta} = \bigsqcup_{1 \leq k \leq N(\delta)} \mathbf{p}^{\text{dec}}(\mathcal{B}_{t_k}), \tag{5.7}$$

endowed with the induced metric d . From the construction of d it is clear that we can bound their Hausdorff distance

$$d_H(\mathcal{T}_\alpha^{\text{dec}, \delta}, \mathcal{T}_\alpha^{\text{dec}}) \leq 2r^\delta. \tag{5.8}$$

Let us do a similar thing with the discrete decorated trees. We introduce $T_n^{\text{dec}, \delta} \subset T_n^{\text{dec}}$ endowed with the induced metric from $b_n^{-\gamma} \cdot d_n^{\text{dec}}$, where

$$T_n^{\text{dec}, \delta} = \bigsqcup_{1 \leq k \leq N(\delta)} \mathbf{p}_n^{\text{dec}}(B_{v_{n,k}}). \tag{5.9}$$

We have, for the distance $b_n^{-\gamma} \cdot d_n^{\text{dec}}$,

$$d_H(T_n^{\text{dec}, \delta}, T_n^{\text{dec}}) \leq 2r_n^\delta. \tag{5.10}$$

We introduce the following auxiliary object

$$\widehat{T}_n^{\text{dec}, \delta} = \bigsqcup_{1 \leq k \leq N(\delta)} B_{v_{n,k}},$$

which we endow with the distance \hat{d} defined as

$$\hat{d}(x, y) = b_n^{-\gamma} \cdot d_n^{\text{dec}}(\mathbf{p}_n^{\text{dec}}(x), \mathbf{p}_n^{\text{dec}}(y)) + \delta \cdot \mathbf{1}_{\{x, y \text{ don't belong to the same block}\}}.$$

It is easy to check that \hat{d} is a distance on $\hat{T}_n^{\text{dec}, \delta}$ and that the projection $\mathbf{p}_n^{\text{dec}}$ is a δ -isometry, so that

$$d_{\text{GH}}\left(\left(\hat{T}_n^{\text{dec}, \delta}, \hat{d}\right), \left(T_n^{\text{dec}, \delta}, b_n^{-\gamma} \cdot d_n^{\text{dec}}\right)\right) \leq \delta.$$

The role of this auxiliary object is just technical: it allows to have a metric space very close to T_n^{dec} but where no two vertices of two different blocks are identified together. This allows to define function from $\hat{T}_n^{\text{dec}, \delta}$ to $\mathcal{T}_\alpha^{\text{dec}, \delta}$ by just patching together the almost isometries that we have on the blocks.

Step 5, constructing an almost isometry: Now, we construct a function from $\hat{T}_n^{\text{dec}, \delta}$ that we will show is an ϵ -isometry for some small ϵ when n is large enough. Recall the $\epsilon_{v_{n,k}}$ -isometries $\phi_{v_{n,k}}$ from (5.3). We define the function ϕ from $\hat{T}_n^{\text{dec}, \delta}$ to $\mathcal{T}_\alpha^{\text{dec}}$ as the function such that for any $1 \leq k \leq N(\delta)$ and any $x \in B_{v_{n,k}}$ we have

$$\phi(x) := \mathbf{p}^{\text{dec}}(\phi_{v_{n,k}}(x))$$

Now let us show that for n large enough ϕ is an ϵ isometry for

$$\epsilon = 3r^\delta + 3r_n^\delta + 3\delta.$$

First, from (5.8) and the definition of ϕ from the $\phi_{v_{n,k}}$, it is clear that the ϕ satisfies the almost surjectivity condition for that value of ϵ . Then we have to check the almost isometry condition. For that, let us take n large enough so that on the realization that we consider, the genealogical order on $(v_{n,1}, v_{n,2}, \dots, v_{n,N(\delta)})$ in T_n coincides with that on $(t_1, t_2, \dots, t_{N(\delta)})$ defined by (4.1).

Now for $k, k' \leq N(\delta)$ for which $t_k \prec t_{k'}$ (and equivalently $v_{n,k} \prec v_{n,k'}$) we have

$$\begin{aligned} \sum_{\substack{v_{n,k} \prec v_{n,i} \prec v_{n,k'} \\ 1 \leq i \leq N(\delta)}} \mathbf{d}_{v_{n,i}}\left(\rho_{v_{n,i}}, Z_{v_{n,i}}^{v_{n,k'}}\right) &\leq \sum_{w \in T_n: v_{n,k} \prec w \prec v_{n,k'}} \mathbf{d}_w\left(\rho_w, Z_w^{v_{n,k'}}\right) \\ &\leq \sum_{\substack{v_{n,k} \prec v_{n,i} \prec v_{n,k'} \\ 1 \leq i \leq N(\delta)}} \mathbf{d}_{v_{n,i}}\left(\rho_{v_{n,i}}, Z_{v_{n,i}}^{v_{n,k'}}\right) + r_n^\delta \end{aligned}$$

and similarly

$$\sum_{\substack{t_k \prec t_i \prec t_{k'} \\ 1 \leq i \leq N(\delta)}} \delta_{t_i}\left(\rho_{t_i}, Z_{t_i}^{t_{k'}}\right) \leq \sum_{w \in \mathbb{B}: t_k \prec w \prec t_{k'}} \delta_w\left(\rho_w, Z_w^{t_{k'}}\right) \leq \sum_{\substack{t_k \prec t_i \prec t_{k'} \\ 1 \leq i \leq N(\delta)}} \delta_{t_i}\left(\rho_{t_i}, Z_{t_i}^{t_{k'}}\right) + r^\delta$$

and so we have

$$\begin{aligned} &\left| \sum_{w \in T_n: v_{n,k} \prec w \prec v_{n,k'}} \mathbf{d}_w\left(\rho_w, Z_w^{v_{n,k'}}\right) - \sum_{w \in \mathbb{B}: t_k \prec w \prec t_{k'}} \delta_w\left(\rho_w, Z_w^{t_{k'}}\right) \right| \\ &\leq \sum_{i: v_{n,k} \prec v_{n,i} \prec v_{n,k'}} \left| \mathbf{d}_{v_{n,i}}\left(\rho_{v_{n,i}}, Z_{v_{n,i}}^{v_{n,k'}}\right) - \delta_{t_i}\left(\rho_{t_i}, Z_{t_i}^{t_{k'}}\right) \right| + r_n^\delta + r_n^\delta \\ &\leq \underbrace{\sum_{k, k', i=1}^{N(\delta)} \left| \mathbf{d}_{v_{n,i}}\left(\rho_{v_{n,i}}, Z_{v_{n,i}}^{v_{n,k'}}\right) - \delta_{t_i}\left(\rho_{t_i}, Z_{t_i}^{t_{k'}}\right) \right|}_{\epsilon_n} + r_n^\delta + r_n^\delta \end{aligned}$$

and by (5.2), the term ϵ_n in the above display tends to 0 for a fixed δ as n tends to infinity. In particular, it is eventually smaller than δ .

Now let us take two $x, x' \in \widehat{T}_n^{\text{dec}, \delta}$, where $x \in B_{v_{n,k}}$ and $x' \in B_{v_{n,k'}}$. Let us study the different cases

- Case 1: assume that $k = k'$, then

$$d(\phi(x), \phi(x')) = \delta_{t_k}(\phi_{v_{n,k}}(x), \phi_{v_{n,k}}(x')) \quad \text{and} \quad \hat{d}(x, x') = \mathbf{d}_{v_{n,k}}(x, x'). \quad (5.11)$$

and then we have $|d(\phi(x), \phi(x')) - \hat{d}(x, x')| \leq \epsilon_{v_{n,k}}$ by definition of $\phi_{v_{n,k}}$ being an $\epsilon_{v_{n,k}}$ -isometry. For a fixed δ , this quantity becomes smaller than δ for n large enough. Since $k \leq N(\delta) < \infty$ a.s., this holds uniformly in k .

- Case 2: assume that $t_k \prec t_{k'}$, then

$$\begin{aligned} d(\phi(x), \phi(x')) &= \delta_{t_k}(\phi(x), Z_{t_k}^{t_{k'}}) + \sum_{w \in \mathbb{B}: t_k \prec w \prec t_{k'}} \delta_w(\rho_w, Z_w^{t_{k'}}) + \delta_{t_{k'}}(\rho_{t_{k'}}, \phi(x')) \\ &= \delta_{t_k}(\phi_{v_{n,k}}(x), Z_{t_k}^{t_{k'}}) + \sum_{w \in \mathbb{B}: t_k \prec w \prec t_{k'}} \delta_w(\rho_w, Z_w^{t_{k'}}) + \delta_{t_{k'}}(\rho_{t_{k'}}, \phi_{t_{k'}}(x')) \end{aligned}$$

and

$$\hat{d}(x, x') = \mathbf{d}_{v_{n,k'}}(x, Z_{v_{n,k}}^{v_{n,k'}}) + \sum_{w \in T_n: v_{n,k} \prec w \prec v_{n,k'}} \mathbf{d}_w(\rho_w, Z_w^{v_{n,k'}}) + \mathbf{d}_{v_{n,k'}}(\rho_{v_{n,k'}}, x') + \delta.$$

Hence we have

$$\begin{aligned} |d(\phi(x), \phi(x')) - \hat{d}(x, x')| &\leq |\delta_{t_k}(\phi(x), Z_{t_k}^{t_{k'}}) - \mathbf{d}_{v_{n,k'}}(x, Z_{v_{n,k}}^{v_{n,k'}})| \\ &+ \left| \sum_{w \in \mathbb{B}: t_k \prec w \prec t_{k'}} \delta_w(\rho_w, Z_w^{t_{k'}}) - \sum_{w \in T_n: v_{n,k} \prec w \prec v_{n,k'}} \mathbf{d}_w(\rho_w, Z_w^{v_{n,k'}}) \right| \\ &+ |\delta_{t_{k'}}(\rho_{t_{k'}}, \phi_{t_{k'}}(x')) - \mathbf{d}_{v_{n,k'}}(\rho_{v_{n,k'}}, x')| \end{aligned}$$

and so, upper-bounding each term, the last display is eventually smaller than ϵ as $n \rightarrow \infty$.

- Case 3: assume that $t_k \not\prec t_{k'}$ and $t_{k'} \not\prec t_k$. Then, we can write

$$d(\phi(x), \phi(x')) = \delta_{t_k \wedge t_{k'}}(Z_{t_k \wedge t_{k'}}^{t_k}, Z_{t_k \wedge t_{k'}}^{t_{k'}}) + \sum_{t_k \wedge t_{k'} \prec w \preceq t_{k'}} \delta_w(\rho_w, Z_w^{t_{k'}}) + \delta_{t_{k'}}(\rho_{t_{k'}}, \phi(x)) \quad (5.12)$$

$$+ \sum_{t_k \wedge t_{k'} \prec w \preceq t_k} \delta_w(\rho_w, Z_w^{t_k}) + \delta_{t_k}(\rho_{t_k}, \phi(x')), \quad (5.13)$$

in the same fashion as before, and a similar expression can be written down for the discrete object.

Now, we can consider two cases: either $t_k \wedge t_{k'} = t_j$ for some $j \in \{1, \dots, N(\delta)\}$, in which case we can write

$$|\delta_{t_k \wedge t_{k'}}(Z_{t_k \wedge t_{k'}}^{t_k}, Z_{t_k \wedge t_{k'}}^{t_{k'}}) - \mathbf{d}_{v_{n,k} \wedge v_{n,k'}}(Z_{v_{n,k} \wedge v_{n,k'}}^{v_{n,k}}, Z_{v_{n,k} \wedge v_{n,k'}}^{v_{n,k'}})| \leq 2\epsilon_n,$$

which tends to 0 as $n \rightarrow \infty$. Otherwise, it means that $\Delta_{t_k \wedge t_{k'}} \leq \delta$ and so we can upper-bound the term in the last display by $r^\delta + r_n^\delta$.

With the other terms, we can reason similarly as in the previous case and get

$$|\hat{d}(x, x') - d(\phi(x), \phi(x'))| \leq 3r^\delta + 4\epsilon_n + 3r_n^\delta + 2\delta + \epsilon_{v_{n,k}} + \epsilon_{v_{n,k'}}, \quad (5.14)$$

which is smaller than ϵ for n large enough.

Adding the measures We sketch here a modification of the proof above to improve the convergence from Gromov–Hausdorff to Gromov–Hausdorff–Prokhorov topology.

Step 1, the total mass converges: First, remark that under the assumptions of the theorem with $\beta \leq \alpha$ the total mass $\nu_n^{\text{dec}}(T_n^{\text{dec}})$ of the measure carried on T_n^{dec} satisfies

$$a_n^{-1} \cdot \nu_n^{\text{dec}}(T_n^{\text{dec}}) \xrightarrow{n \rightarrow \infty} 1.$$

On the other hand, if $\beta > \alpha$, then using the coupling defined in the proof above, we claim that we have

$$b_n^{-\beta} \cdot \nu_n^{\text{dec}}(T_n^{\text{dec}}) = b_n^{-\beta} \cdot \sum_{i=1}^{|T_n|} \nu_{v_i}(B_{v_i}) \xrightarrow{n \rightarrow \infty} \sum_{0 \leq s \leq 1} \Delta_s^\beta \cdot \nu_s(\mathcal{B}_s). \tag{5.15}$$

Let us prove the above display. First, using (5.1) and (5.2), we can write for all $k \geq 1$,

$$b_n^{-\beta} \cdot \nu_{v_{n,k}}(B_{v_{n,k}}) \xrightarrow{n \rightarrow \infty} \Delta_{t_k}^\beta \cdot \nu_{t_k}(\mathcal{B}_{t_k}).$$

This ensures that for any given $\delta > 0$ we have

$$b_n^{-\beta} \cdot \sum_{k=0}^{\infty} \nu_{v_{n,k}}(B_{v_{n,k}}) \mathbf{1}_{\{d^+(v_{n,k}) > \delta b_n\}} \rightarrow \sum_{0 \leq s \leq 1} \Delta_s^\beta \cdot \nu_s(\mathcal{B}_s) \mathbf{1}_{\{\Delta_s > \delta\}},$$

and the quantity on the right-hand-side converges to $\sum_{0 \leq s \leq 1} \Delta_s^\beta \cdot \nu_s(\mathcal{B}_s)$ as $\delta \rightarrow 0$. From there, we can construct a sequence $\delta_n \rightarrow 0$ such that the following convergence holds on an event of probability 1

$$b_n^{-\beta} \cdot \sum_{k=0}^{\infty} \nu_{v_{n,k}}(B_{v_{n,k}}) \mathbf{1}_{\{d^+(v_{n,k}) \geq \delta_n b_n\}} \rightarrow \sum_{0 \leq s \leq 1} \Delta_s^\beta \cdot \nu_s(\mathcal{B}_s).$$

Now, using Condition B3 we get that

$$R(n, \delta) = b_n^{-\beta} \cdot \sum_{i=1}^{|T_n|} \nu_{v_i}(B_{v_i}) \mathbf{1}_{\{d^+(v_i) \leq \delta b_n\}}$$

tends to 0 in probability as $\delta \rightarrow 0$, uniformly in n large enough. This ensures that for our sequence δ_n tending to 0 we have the convergence $R(n, \delta_n) \xrightarrow{n \rightarrow \infty} 0$ in probability. Putting things together we get that (5.15) holds in probability.

Step 2, Convergence of the sampled points: Now that we know that the total mass of ν_n^{dec} appropriately normalized converges to that of ν_β^{dec} , we consider a point Y_n sampled under normalized version of ν_n^{dec} and Υ_β sampled under a normalized version of ν_β^{dec} . In order to prove that $(T_n^{\text{dec}}, d, \nu_n^{\text{dec}})$ converges to $(\mathcal{T}_\alpha^{\text{dec}}, d, \nu_\alpha^{\text{dec}})$ in distribution, we will prove that the pointed spaces $(T_n^{\text{dec}}, d, Y_n)$ converge to $(\mathcal{T}_\alpha^{\text{dec}}, d, \Upsilon_\beta)$ in distribution for the pointed Gromov–Hausdorff topology.

Step 3, Sampling a uniform point: To construct Y_n on T_n^{dec} , we first sample a point X_n on $[0, 1]$. Then we introduce

$$K_n := \inf \{k \geq 1 \mid M_k \geq X_n \cdot M_{|T_n|}\} \tag{5.16}$$

and then sample a point on $B_{v_{K_n}}$ using a normalized version of the measure $\nu_{v_{K_n}}$.

For $\beta \leq \alpha$, in the continuous setting, we construct a random point Υ_β on $\mathcal{T}_\alpha^{\text{dec}}$ distributed as ν_β^{dec} by defining it as $\mathbf{p}^{\text{dec}}(X)$ where $X \sim \text{Unif}[0, 1]$.

For $\beta > \alpha$ we can define the following random point Υ_β on $\mathcal{T}_\alpha^{\text{dec}}$ in three steps: sample $X \sim \text{Unif}[0, 1]$ and then set

$$Z := \inf \left\{ t \in [0, 1] \left| \sum_{0 \leq s \leq t} \Delta_s^\beta \nu_s(\mathcal{B}_s) \geq X \cdot \sum_{0 \leq s \leq 1} \Delta_s^\beta \nu_s(\mathcal{B}_s) \right. \right\}$$

and then sample a point on \mathcal{B}_Z using a normalized version of the measure ν_Z . (Note that in this case, Z is a jump time of $X^{\text{exc},(\alpha)}$ almost surely.) This construction ensures that conditionally on $X^{\text{exc},(\alpha)}$, and $(\mathcal{B}_s)_{s \in \mathbb{B}}$, for all $t \in [0, 1]$ we have

$$\mathbb{P}(Z = t) = \frac{\Delta_t^\beta \cdot \nu_t(\mathcal{B}_t)}{\sum_{0 \leq s \leq 1} \Delta_s^\beta \cdot \nu_s(\mathcal{B}_s)}.$$

Step 4, Coupling Y_n and Υ_β : In both cases $\alpha \leq \beta$ and $\alpha > \beta$ respectively, we couple the construction of Y_n and Υ by using the same uniform random variable $X_n = X$. Now let us consider the two cases separately.

Case $\beta \leq \alpha$. Consider the convergence (3.3). Since this convergence is in probability to a constant, it happens jointly with other convergence results used above using Slutsky’s lemma. Hence, using the Skorokhod embedding again, we can suppose that this convergence is almost sure, jointly with the other ones. Then, we can essentially go through the same proof as above and consider an extra decoration corresponding to $B_{v_{K_n}}$, which contains the random point Y_n .

Case $\beta > \alpha$. In that case, thanks to the reasoning of Step 1, we have the following convergence for the Skorokhod topology

$$b_n^{-\beta} \cdot \left(\sum_{i=1}^{\lfloor tT_n \rfloor} \nu_{v_i}(B_{v_i}) \right)_{0 \leq t \leq 1} \xrightarrow{n \rightarrow \infty} \left(\sum_{0 \leq s \leq t} \Delta_s^\beta \cdot \nu_s(\mathcal{B}_s) \right)_{0 \leq t \leq 1}.$$

With the definition of respectively Z and K_n using the coupling $X_n = X$, the above convergence ensures that for n large enough we have $K_n = v_{n,k}$ and $Z = t_k$ for some (random) $k \geq 1$. Then again, we can essentially go through the same proof as above and consider an extra random point Y_n sampled on $B_{v_{n,k}}$ and a point Υ_β sampled on \mathcal{B}_{t_k} in such a way that those points are close together. \square

6 Self-similarity properties and alternative constructions

The goal of this section is to decompose the decorated stable tree along a discrete tree, using a framework similar to that developed in Section 2.1, with the difference that the gluing will occur along the whole Ulam tree instead of a finite tree. In particular this decomposition, which follows directly from a similar description of the α -stable tree itself, will allow us to describe the decorated stable tree as obtained by a sort of line-breaking construction, and also identify its law as the fixed point of some transformation. We rely on a similar construction that holds jointly for the α -stable tree and its associated looptree described in [34].

In this section, we will always assume that $\beta > \alpha$ and that $\mathbb{E}[\nu(\mathcal{B})]$ and $\mathbb{E}[\text{diam}(\mathcal{B})]$ are finite.

6.1 Introducing the spine and decorated spine distributions

We construct two related random metric spaces, \mathcal{S} and \mathcal{S}^{dec} , which will be the building blocks used to construct \mathcal{T}_α and $\mathcal{T}_\alpha^{\text{dec}}$ respectively. This construction will depend on parameters $\alpha \in (1, 2)$ and $\gamma \in (\alpha - 1, 1]$ and $\beta > \alpha$. First, let

Decorated stable trees

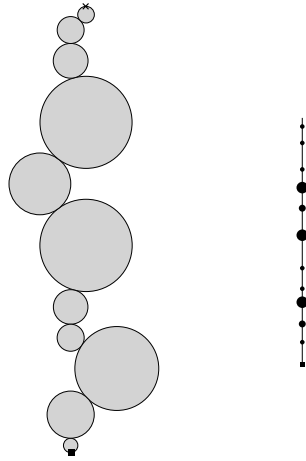


Figure 3: On the left, the space \mathcal{S}^{dec} . On the right, the corresponding space \mathcal{S} which is just a segment with atoms along it.

- $(Y_k)_{k \geq 1}$ be a sequence of i.i.d. uniform random variables in $[0, 1]$,
- $(P_k) \sim \text{GEM}(\alpha - 1, \alpha - 1)$, and L its $(\alpha - 1)$ -diversity.
- $(\mathcal{B}_k, d_k, \rho_k, \mu_k, \nu_k)$ i.i.d. random metric spaces with the same law as $(\mathcal{B}, d, \rho, \mu, \nu)$,
- for every $k \geq 1$, a random point Z_k taken on \mathcal{B}_k under μ_k .

We refer for example to [34, Appendix A.2] for the definitions of the distributions in the second bulletpoint. We first define the random spine

$$\mathcal{S} = (S, d_S, \rho_S, \mu_S)$$

as the segment $S = [0, L]$, rooted at $\rho_S = 0$, endowed with the probability measure $\mu_S = \sum_{k \geq 1} P_k \delta_{L \cdot Y_k}$. In order to construct \mathcal{S}^{dec} we are going to informally replace every atom of μ_S with a metric space scaled by an amount that corresponds to the mass of that atom. To do so, we consider the disjoint union

$$\bigsqcup_{i=1}^{\infty} \mathcal{B}_i, \tag{6.1}$$

which we endow with the distance $d_{\mathcal{S}^{\text{dec}}}$ defined as

$$d_{\mathcal{S}^{\text{dec}}}(x, y) = \begin{cases} P_i^\gamma d_i(x, y) & \text{if } x, y \in \mathcal{B}_i, \\ P_i^\gamma d_i(x, Z_i) + \sum_{k: Y_i < Y_k < Y_j} P_i^\gamma d_k(\rho_k, Z_k) + P_j^\gamma d_j(\rho_j, y) & \text{if } x \in \mathcal{B}_i, y \in \mathcal{B}_j, Y_i < Y_j. \end{cases}$$

Then \mathcal{S}^{dec} is defined as the completion of $\bigsqcup_{i=1}^{\infty} \mathcal{B}_i$ equipped with this distance. Its root ρ can be obtained as a limit $\rho = \lim_{i \rightarrow \infty} \rho_{\sigma_i}$ for any sequence $(\sigma_i)_{i \geq 1}$ for which $Y_{\sigma_i} \rightarrow 0$. Essentially, \mathcal{S}^{dec} looks like a skewer of decorations that are arranged along a line, in uniform random order. We can furthermore endow \mathcal{S}^{dec} with a probability measure $\mu_{\mathcal{S}^{\text{dec}}} = \sum_{k=1}^{\infty} P_k \mu_k$. If $\beta > \alpha$ and $\mathbb{E}[\nu(\mathcal{B})]$ is finite, we also define the measure $\nu_{\mathcal{S}^{\text{dec}}} = \sum_{k=1}^{\infty} P_k^\beta \nu_k$. In the end, we have defined

$$\mathcal{S}^{\text{dec}} = (\mathcal{S}^{\text{dec}}, d_{\mathcal{S}^{\text{dec}}}, \rho_{\mathcal{S}^{\text{dec}}}, \mu_{\mathcal{S}^{\text{dec}}}, \nu_{\mathcal{S}^{\text{dec}}}).$$

It is important to note that conditionally on \mathcal{S} , constructing \mathcal{S}^{dec} consists in replacing every atom of μ_S by an independent copy of a random metric space appropriately rescaled. This exactly corresponds to what we are trying to do with the entire tree \mathcal{T}_α .

Up to now, it is not clear whether the last display is well defined though, because it is not clear whether the function $d_{S^{dec}}$ only takes finite values. The next lemma ensures that it is the case as long as $\text{diam}(\mathcal{B})$ has a finite first moment. Introduce the quantity $R = \sum_{i \geq 1} P_i^\gamma \text{diam}(\mathcal{B}_i)$.

Lemma 6.1. *If $\mathbb{E}[\text{diam}(\mathcal{B})^p] < \infty$ for some $p \geq 1$ then $\mathbb{E}[R^p]$ is finite so S^{dec} is almost surely a compact metric space with $\mathbb{E}[\text{diam}(S^{dec})^p] < \infty$.*

In order to get Lemma 6.1 we will apply the following lemma, with $\xi = \theta = \alpha - 1$ and $Z_i = \text{diam}(\mathcal{B}_i)$.

Lemma 6.2. *Let $(Z_i)_{i \geq 1}$ be i.i.d. positive random variables with moment of order $p \geq 1$, and $(P_i)_{i \geq 1}$ an independent random variable with GEM(ξ, θ) distribution, and $\gamma > \xi$. Then the random variable*

$$\sum_{i=1}^{\infty} P_i^\gamma Z_i,$$

admits moments of order p .

Note that this lemma also guarantees that the measure $\nu_{S^{dec}} = \sum_{k=1}^{\infty} P_k^\beta \nu_k$ is almost surely a finite measure, as soon as $\mathbb{E}[\nu(B)]$ is finite.

Proof. Suppose that $\xi < \gamma \leq 1$ and write

$$\begin{aligned} \left(\sum_{i=1}^{\infty} P_i^\gamma Z_i \right)^p &= \left(\sum_{i=1}^{\infty} P_i^\gamma \right)^p \left(\sum_{i=1}^{\infty} \left(\frac{P_i^\gamma}{\sum_{i=1}^{\infty} P_i^\gamma} \right) Z_i \right)^p \\ &\stackrel{\text{convexity}}{\leq} \left(\sum_{i=1}^{\infty} P_i^\gamma \right)^p \left(\sum_{i=1}^{\infty} \left(\frac{P_i^\gamma}{\sum_{i=1}^{\infty} P_i^\gamma} \right) Z_i^p \right) \\ &\leq \left(\sum_{i=1}^{\infty} P_i^\gamma \right)^{p-1} \left(\sum_{i=1}^{\infty} P_i^\gamma Z_i^p \right). \end{aligned}$$

In the case $\gamma > 1$, since all the P_i are smaller than 1, we can write

$$\left(\sum_{i=1}^{\infty} P_i^\gamma Z_i \right)^p \leq \left(\sum_{i=1}^{\infty} P_i Z_i \right)^p \leq \left(\sum_{i=1}^{\infty} P_i Z_i^p \right),$$

where the second inequality comes from convexity. Now, for any $\xi < \gamma \leq 1$, for $n = \lceil p-1 \rceil$ we have

$$\begin{aligned} \mathbb{E} \left[\left(\sum_{i=1}^{\infty} P_i^\gamma \right)^{p-1} \left(\sum_{i=1}^{\infty} P_i^\gamma Z_i^p \right) \right] &\leq \mathbb{E} \left[\left(\sum_{i=1}^{\infty} P_i^\gamma Z_i^p \right) \cdot \left(\sum_{i=1}^{\infty} P_i^\gamma \right)^n \right] \\ &= \mathbb{E} \left[\sum_{(i_0, i_1, i_2, \dots, i_n) \in (\mathbb{N}^*)^{n+1}} P_{i_0}^\gamma \cdot Z_{i_0}^p \cdot P_{i_1}^\gamma \dots P_{i_n}^\gamma \right] \\ &\leq \sum_{(i_0, i_1, i_2, \dots, i_n) \in (\mathbb{N}^*)^{n+1}} \mathbb{E} [P_{i_1}^\gamma \dots P_{i_n}^\gamma] \cdot \mathbb{E} [Z_{i_0}^p] \\ &< \infty. \end{aligned}$$

The last term is finite provided that $\gamma > \xi$, using Lemma 5.4 in [18]. □

6.2 Description of an object as a decorated Ulam tree

Now that we have defined the random object S^{dec} , we define some decorations of the Ulam tree that are constructed in such a way that up to some scaling factor, all the decorations are i.i.d. with the same law as S^{dec} . Gluing all those decorations along the structure of the tree, as defined in the introduction, will give us a random object $\tilde{\mathcal{T}}_\alpha^{\text{dec}}$ that has the same law as $\mathcal{T}_\alpha^{\text{dec}}$ (the proof of that fact will come later in the section).

In fact, in what follows, we provide two equivalent descriptions of the same object $\tilde{\mathcal{T}}_\alpha^{\text{dec}}$:

- One of them is a description as a random self-similar metric space: this construction will ensure that the object that we construct is compact under the weakest possible moment assumption on the diameter of the decoration and also give us the Hausdorff dimension of the object.
- The other one is through an iterative gluing construction. It is that one that we use to identify the law of $\tilde{\mathcal{T}}_\alpha^{\text{dec}}$ with that of $\mathcal{T}_\alpha^{\text{dec}}$.

The fact that the two constructions yield the same object is a result from [34].

6.2.1 Decorations on the Ulam tree

We extend the framework introduced in the first section for finite trees to the entire Ulam tree, following [34]. We work with families of decorations on the Ulam tree $\mathcal{D} = (\mathcal{D}(v), v \in \mathbb{U})$ of measured metric spaces indexed by the vertices of \mathbb{U} . Each of the decorations $\mathcal{D}(v)$ indexed by some vertex $v \in \mathbb{U}$ is rooted at some ρ_v and carries a sequence of gluing points $(\ell_v(i))_{i \geq 1}$. The way to construct the associated decorated Ulam tree is to consider as before

$$\bigsqcup_{v \in \mathbb{U}} \mathcal{D}(v)$$

and consider the metric gluing of those blocks obtained by the relations $\ell_v(i) \sim \rho_{vi}$ for all $v \in \mathbb{U}$ and $i \geq 1$. For topological considerations, we actually consider the metric completion of the obtained object. The final result is denoted $\mathcal{G}(\mathcal{D})$.

The completed Ulam tree Recall the definition of the Ulam tree $\mathbb{U} = \bigcup_{n \geq 0} \mathbb{N}^n$ with $\mathbb{N} = \{1, 2, \dots\}$. Here $\mathbb{N}^0 = \{\emptyset\}$. Introduce the set $\partial\mathbb{U} = \mathbb{N}^\mathbb{N}$, to which we refer as the *leaves* of the Ulam tree, which we see as the infinite rays joining the root to infinity and set $\bar{\mathbb{U}} := \mathbb{U} \cup \partial\mathbb{U}$. On this set, we have a natural genealogical order \preceq defined in such a way that $v \preceq w$ if and only if v is a prefix of w . From this order relation we can define for any $v \in \mathbb{U}$ the *subtree descending from v* as the set $T(v) := \{w \in \bar{\mathbb{U}} \mid v \preceq w\}$. The collection of sets $\{T(v), v \in \mathbb{U}\} \sqcup \{v\}, v \in \mathbb{U}$ generates a topology over $\bar{\mathbb{U}}$, which can also be generated using an appropriate ultrametric distance. Endowed with this distance, the set $\bar{\mathbb{U}}$ is then a separable and complete metric space.

Identification of the leaves Suppose that \mathcal{D} is such that $\mathcal{G}(\mathcal{D})$ is compact. We can define a map

$$\iota_{\mathcal{D}} : \partial\mathbb{U} \rightarrow \mathcal{G}(\mathcal{D}), \tag{6.2}$$

that maps every leaf of the Ulam-Harris tree to a point of $\mathcal{G}(\mathcal{D})$. For any $\mathbf{i} = i_1 i_2 \dots \in \partial\mathbb{U}$, the map is defined as

$$\iota_{\mathcal{D}}(\mathbf{i}) := \lim_{n \rightarrow \infty} \rho_{i_1 i_2 \dots i_n} \in \mathcal{G}(\mathcal{D}),$$

The fact that this map is well-defined and continuous is proved in [34].

Scaling of decorations In the rest of this section, for $\mathcal{X} = (X, d, \rho, \mu, \nu)$ a pointed metric space endowed with finite measures (as well as some extra structure like a sequence of points for example) we will use the notation

$$\text{Scale}(a, b, c; \mathcal{X}) = (X, a \cdot d, \rho, b \cdot \mu, c \cdot \nu),$$

where the resulting object still carries any extra structure that \mathcal{X} originally carried.

6.2.2 A self-similar decoration

We introduce the following random variables:

- For every $v \in \mathbb{U}$, sample independently $(Q_{vi})_{i \geq 1} \sim \text{GEM}(\frac{1}{\alpha}, 1 - \frac{1}{\alpha})$ and denote D_v the $\frac{1}{\alpha}$ -diversity of $(Q_{vi})_{i \geq 1}$.
- This defines for every $v \in \mathbb{U}$,

$$\overline{Q}_v = \prod_{w \preceq v} Q_w \quad \text{and} \quad w_v = (\overline{Q}_v)^{\frac{1}{\alpha}} \cdot D_v$$

This in fact also defines a probability measure η on $\partial\mathbb{U}$ the frontier of the tree which is characterized by $\eta(T(v)) = \overline{Q}_v$, for all $v \in \mathbb{U}$.

- For every $v \in \mathbb{U}$ we sample independently of the rest $\mathcal{S}_v^{\text{dec}}$ that has the same law as $\mathcal{S}^{\text{dec}} = (\mathcal{S}^{\text{dec}}, d_{\mathcal{S}^{\text{dec}}}, \rho_{\mathcal{S}^{\text{dec}}}, \mu_{\mathcal{S}^{\text{dec}}}, \nu_{\mathcal{S}^{\text{dec}}})$ and sample a sequence of points from its measure $\mu_{\mathcal{S}^{\text{dec}}}$. We call those points $(X_{vi})_{i \geq 1}$ and define $\ell_v : i \mapsto X_{vi} \in \mathcal{S}_v^{\text{dec}}$. Then we can consider the following decorations on the Ulam tree as, for $v \in \mathbb{U}$,

$$\mathcal{D}(v) := \text{Scale}(w_v^\gamma, w_v, w_v^\beta; \mathcal{S}_v^{\text{dec}}). \tag{6.3}$$

From these decorations on the Ulam tree, we can define the corresponding decorated tree (and consider its metric completion) that we write as

$$\tilde{\mathcal{T}}_\alpha^{\text{dec}} := \mathcal{G}(\mathcal{D}).$$

Note that the object defined above is not necessarily compact. Whenever the underlying block (and hence also the decorated spine thanks to Lemma 6.1) has a moment of order $p > \frac{\alpha}{\gamma}$, a result of [34, Section 4.2] ensures that the obtained metric space $\tilde{\mathcal{T}}_\alpha^{\text{dec}}$ is almost surely compact.

The uniform measure Assume that the underlying block has a moment of order $p > \frac{\alpha}{\gamma}$ so that $\tilde{\mathcal{T}}_\alpha^{\text{dec}}$ is almost surely compact. Then the maps $\iota_{\mathcal{D}} : \partial\mathbb{U} \rightarrow \mathcal{G}(\mathcal{D})$ is almost surely well-defined and continuous so we can consider the measure $(\iota_{\mathcal{D}})_* \eta$ on $\mathcal{G}(\mathcal{D})$, the push-forward of the measure η .

The block measure If $\beta > \alpha$ and $\mathbb{E}[\nu(\mathcal{B})] < \infty$ then we can check that the total mass of the ν measures has finite expectation so it's almost surely finite. Hence we can endow $\tilde{\mathcal{T}}_\alpha^{\text{dec}}$ with the measure $\sum_u \nu_u$.

6.2.3 The iterative gluing construction for $\tilde{\mathcal{T}}_\alpha^{\text{dec}}$

Before diving into the construction of $\tilde{\mathcal{T}}_\alpha^{\text{dec}}$, we define a family of time-inhomogeneous Markov chains called Mittag-Leffler Markov chains, first introduced by Goldschmidt and Haas [17], see also [34].

Mittag-Leffler Markov chains Let $0 < \eta < 1$ and $\theta > -\eta$. The generalized Mittag-Leffler $\text{ML}(\eta, \theta)$ distribution has p th moment

$$\frac{\Gamma(\theta)\Gamma(\theta/\eta + p)}{\Gamma(\theta/\eta)\Gamma(\theta + p\eta)} = \frac{\Gamma(\theta + 1)\Gamma(\theta/\eta + p + 1)}{\Gamma(\theta/\eta + 1)\Gamma(\theta + p\eta + 1)} \tag{6.4}$$

and the collection of p -th moments for $p \in \mathbb{N}$ uniquely characterizes this distribution.

A Markov chain $(M_n)_{n \geq 1}$ has the distribution $\text{MLMC}(\eta, \theta)$ if for all $n \geq 1$,

$$M_n \sim \text{ML}(\eta, \theta + n - 1),$$

and its transition probabilities are characterized by the following equality in law

$$(M_n, M_{n+1}) = (\beta_n \cdot M_{n+1}, M_{n+1}),$$

where $\beta_n \sim \text{Beta}\left(\frac{\theta+k-1}{\eta} + 1, \frac{1}{\eta} - 1\right)$ is independent of M_{n+1} .

Construction of $\tilde{\mathcal{T}}_\alpha^{\text{dec}}$ We can now express our second construction of $\tilde{\mathcal{T}}_\alpha^{\text{dec}}$. We start with a sequence $(M_k)_{k \geq 1} \sim \text{MLMC}\left(\frac{1}{\alpha}, 1 - \frac{1}{\alpha}\right)$. Then we defined the sequence $(m_k)_{k \geq 1} = (M_k - M_{k-1})_{k \geq 1}$ of increments of that chain, where we assume by convention that $M_0 = 0$. From there, conditionally on $(M_k)_{k \geq 1}$ we define an independent sequence \mathcal{Y}_k of metric spaces such that

$$\mathcal{Y}_k \stackrel{(d)}{=} \text{Scale}\left(m_k^\gamma, m_k, m_k^\beta; \mathcal{S}^{\text{dec}}\right).$$

Then, we define a sequence of increasing subtrees of the Ulam tree as follows: we let T_1 be the tree containing only one vertex $v_1 = \emptyset$. Then if T_n is constructed, we sample a random number K_{n+1} in $\{1, \dots, n\}$ such that conditionally on all the rest $\mathbb{P}(K_n = k) \propto m_k$. Then, we add the next vertex v_{n+1} to the tree as the right-most child of $v_{K_{n+1}}$. The sequence $(T_n)_{n \geq 1}$ is said to have the distribution of a weighted recursive tree with sequence of weights $(m_k)_{k \geq 1}$. A property of this sequence of trees is the fact that the $\{v_k \mid k \geq 1\} = \mathbb{U}$. Hence for any $v \in \mathbb{U}$ we denote k_v the unique k such that $v_k = v$. Then we consider the decorations on the Ulam tree defined by

$$\widehat{\mathcal{D}}(v) = \mathcal{Y}_{k_v}. \tag{6.5}$$

In this setting, we can again define a measure $\tilde{\eta}$ on the leaves $\partial\mathbb{U}$ of the Ulam tree by taking the weak limit of the uniform measure on $\{v_1, v_2, \dots, v_n\}$ as $n \rightarrow \infty$, the almost sure existence of the limit being guaranteed by [33, Theorem 1.7, Proposition 2.4].

Then, from [34, Proposition 3.2] we have

$$((m_{k_v})_{v \in \mathbb{U}}, \tilde{\eta}) \stackrel{(d)}{=} ((w_v)_{v \in \mathbb{U}}, \eta).$$

From that equality in distribution and the respective definitions (6.3) of \mathcal{D} and (6.5) of $\widehat{\mathcal{D}}$, it is clear that those two families of decorations have the same distribution.

6.2.4 Strategy

The rest of this section is about proving that the random decorated tree $\tilde{\mathcal{T}}_\alpha^{\text{dec}}$ that we constructed above has the same distribution as $\mathcal{T}_\alpha^{\text{dec}}$. For that we are going to characterize their “finite-dimensional marginals” and show that they are the same for the two processes: Using the second description of $\tilde{\mathcal{T}}_\alpha^{\text{dec}}$, we can consider for any $k \geq 1$ the subset $\tilde{\mathcal{T}}_k^{\text{dec}}$ that corresponds to keeping only the blocks corresponding to v_1, \dots, v_k . We compare this to $\mathcal{T}_k^{\text{dec}}$ which is the subset of $\mathcal{T}_\alpha^{\text{dec}}$ spanned by k uniform leaves.

6.3 Finite-dimensional marginals of \mathcal{T}_α and $\mathcal{T}_\alpha^{\text{dec}}$

6.3.1 Approximating the decorated tree by finite dimensional marginals

For any tuple of points x_1, x_2, \dots, x_k in $[0, 1]$, we can consider

$$\text{Span}(X^{\text{exc},(\alpha)}; x_1, x_2, \dots, x_k) = \{x \in [0, 1] \mid x \preceq x_i, \text{ for some } i \in \{1, 2, \dots, k\}\},$$

using the definition of \preceq derived from $X^{\text{exc},(\alpha)}$. We further define

$$\text{Span}(X^{\text{exc},(\alpha)}, (\mathcal{B}_t)_{t \in [0,1]}; x_1, x_2, \dots, x_k) = \bigsqcup_{t \in \text{Span}(X^{\text{exc},(\alpha)}; x_1, x_2, \dots, x_k)} \mathcal{B}_t.$$

Using independent uniform random points $(U_i)_{i \geq 1}$ on $[0, 1]$, we define

$$\mathcal{T}_k = \mathbf{p} \left(\text{Span}(X^{\text{exc},(\alpha)}; U_1, U_2, \dots, U_k) \right)$$

and

$$\mathcal{T}_k^{\text{dec}} = \mathbf{p}^{\text{dec}} \left(\text{Span}(X^{\text{exc},(\alpha)}, (\mathcal{B}_t)_{t \in [0,1]}; U_1, U_2, \dots, U_k) \right),$$

where $\mathbf{p} : [0, 1] \rightarrow \mathcal{T}_\alpha$ and $\mathbf{p}^{\text{dec}} : \sqcup_{t \in [0,1]} \mathcal{B}_t \rightarrow \mathcal{T}_\alpha^{\text{dec}}$ are the respective quotient maps. We say that \mathcal{T}_k (resp. $\mathcal{T}_k^{\text{dec}}$) is the discrete random finite-dimensional marginal of \mathcal{T}_α (resp. $\mathcal{T}_\alpha^{\text{dec}}$), following the standard definition from trees, introduced by Aldous [2].

Remark 6.3. Note that for $\mathcal{T}_k^{\text{dec}}$ to be almost surely well-defined, we don't need the whole decorated tree $\mathcal{T}_\alpha^{\text{dec}}$ to be well-defined as a compact measured metric space. In fact, it is easy to check that if the tail of $\text{diam}(\mathcal{B})$ is such that $\mathbb{P}(\text{diam}(\mathcal{B}) \geq x) \sim x^{-p}$ with $1 < p < \frac{\alpha}{\gamma}$, then $\sup_{v \in \mathbb{B}} \Delta_v^\gamma \cdot \text{diam}(\mathcal{B}_v) = \infty$, even though the distance of a random point to the root in $\mathcal{T}_\alpha^{\text{dec}}$ is almost surely finite.

Lemma 6.4. Identifying \mathcal{B}_v for any $v \in \mathbb{B}$ as a subset of $\mathcal{T}_\alpha^{\text{dec}}$ we almost surely have

$$\bigcup_{v \in \mathbb{B}} \mathcal{B}_v \subset \overline{\bigcup_{n \geq 0} \mathcal{T}_n^{\text{dec}}}.$$

Proof. By properties of the stable excursion, for any $t \in \mathbb{B}$ the set $\{s \succeq t \mid s \in [0, 1]\}$ has a positive Lebesgue measure. Hence almost surely there is some k such that $t \preceq U_k$. \square

Corollary 6.5. The space $\overline{\bigcup_{n \geq 0} \mathcal{T}_n^{\text{dec}}}$ is compact if and only if $\mathcal{T}_\alpha^{\text{dec}}$ is well-defined as a compact metric space.

6.3.2 Description of \mathcal{T}_k and $\mathcal{T}_k^{\text{dec}}$

First, for any $k \geq 1$, we let L_k be the distance $d(U_k, \mathcal{T}_{k-1})$ computed in the tree \mathcal{T}_α . Then we consider the set of branch-points $\mathbb{B} \cap (\mathcal{T}_k \setminus \mathcal{T}_{k-1})$ and define the decreasing sequence $(N_k(\ell))_{\ell \geq 1}$ which is the decreasing reordering of the sequence $(\Delta_t)_{t \in \mathbb{B} \cap (\mathcal{T}_k \setminus \mathcal{T}_{k-1})}$. Denote $N_k = \sum_{\ell \geq 1} N_k(\ell)$. Note that the jumps of the excursion process are all distinct almost surely. We denote $(t_k(\ell))_{\ell \geq 1}$ the corresponding sequence of jump times. We also consider the sequence $(L_k(\ell))_{\ell \geq 1}$ defined as $d(t_k(\ell), \mathcal{T}_{k-1})$. Let us also write $N_k^r(\ell)$ for the quantity $x_{t_k(\ell)}^{U_k}$, defined in (4.2).

Let us check that these random variables are enough to reconstruct $\mathcal{T}_k^{\text{dec}} \setminus \mathcal{T}_{k-1}^{\text{dec}}$. We have

$$\overline{\mathcal{T}_k^{\text{dec}} \setminus \mathcal{T}_{k-1}^{\text{dec}}} = \overline{\bigcup_{t \in \mathcal{T}_k \setminus \mathcal{T}_{k-1}} \mathcal{B}_t}$$

seen as a subset of $\mathcal{T}_\alpha^{\text{dec}}$. Now, from the form of the distance on $\mathcal{T}_\alpha^{\text{dec}}$, the induced distance between two point in $\mathcal{T}_k^{\text{dec}} \setminus \mathcal{T}_{k-1}^{\text{dec}}$ only depends on:

- the ordering of the jump times $(t_k(\ell))_{\ell \geq 1}$ by the relation \preceq (they are all comparable because by definition we have $t_k(\ell) \preceq U_k$ for all $\ell \geq 1$),
- the sizes of the jumps $(\Delta_{t_k(\ell)})$ and the corresponding blocks $\mathcal{B}_{t_k(\ell)}$,
- the position of the gluing points $Z_{t_k(\ell)}^{U_k}$ on the corresponding block $\mathcal{B}_{t_k(\ell)}$.

We can also check that the topological closure $\overline{\partial(\mathcal{T}_k^{\text{dec}} \setminus \mathcal{T}_{k-1}^{\text{dec}})}$ contains a single point, we call this point Z_k . When considering the compact object $\overline{\mathcal{T}_k^{\text{dec}} \setminus \mathcal{T}_{k-1}^{\text{dec}}}$, we see it as rooted at Z_k .

Now, if we want to reconstruct the entire $\mathcal{T}_k^{\text{dec}}$ from $(\mathcal{T}_k^{\text{dec}} \setminus \mathcal{T}_{k-1}^{\text{dec}})$ and $\mathcal{T}_{k-1}^{\text{dec}}$ we also need some additional information. For that we consider the finite sequence $U_1 \wedge U_k, U_2 \wedge U_k, \dots, U_{k-1} \wedge U_k$ which are all $\preceq U_k$ by definition. Because they are all comparable, we can consider V_k the maximal element of this sequence. We let I_k be the unique $i \leq k - 1$ such that $V_k \in (\mathcal{T}_i^{\text{dec}} \setminus \mathcal{T}_{i-1}^{\text{dec}})$. Additionally, we let $W_k = u_{V_k}^{U_k}$. We can check that the corresponding point $Z_k = Y_{V_k, W_k}$ is such that, conditionally on $\mathcal{T}_{k-1}^{\text{dec}}$ and V_k , distributed as a random point under the measure ν_{V_k} carried by the block \mathcal{B}_{V_k} , by definition.

Now, we use the results of [18, Proposition 2.2] that identify the joint law of these quantities as scaling limit of analogous quantities defined in a discrete setting on trees constructed using the so-called Marchal algorithm. Those quantities have been studied with a slightly different approach in [34], which provides an explicit description of those random variables, jointly in k and ℓ . The following can be read from [34, Lemma 5.5, Proposition 5.7]:

- the sequence $(N_k)_{k \geq 1}$ has the distribution of the sequence $(m_k)_{k \geq 1} = (M_k - M_{k-1})_{k \geq 1}$ where $(M_k)_{k \geq 1} \sim \text{MLMC}(\frac{1}{\alpha}, 1 - \frac{1}{\alpha})$.
- the sequences $(\frac{N_k(\ell)}{N_k})_{\ell \geq 1}$ are i.i.d. with law $\text{PD}(\alpha - 1, \alpha - 1)$ and $L_k = \alpha \cdot N_k^{\alpha-1} \cdot S_k$ where S_k is the $(\alpha - 1)$ diversity of $(\frac{N_k(\ell)}{N_k})_{\ell \geq 1}$.
- the random variables $\frac{L_k(\ell)}{L_k}$ are i.i.d. uniform on $[0, 1]$,
- the random variables $\frac{N_k^k(\ell)}{N_k(\ell)}$ are i.i.d. uniform on $[0, 1]$.

In particular, from the above description, we can check that conditionally on the sequence $(N_k)_{k \geq 1}$, the random variables $(\overline{\mathcal{T}_k^{\text{dec}} \setminus \mathcal{T}_{k-1}^{\text{dec}}})_{k \geq 1}$ are independent with distribution given by

$$\overline{\mathcal{T}_k^{\text{dec}} \setminus \mathcal{T}_{k-1}^{\text{dec}}} = \text{Scale}(N_k^\gamma, N_k, N_k^\beta; \mathcal{S}^{\text{dec}}).$$

We have identified the laws of the two sequences $(\mathcal{T}_k^{\text{dec}} \setminus \mathcal{T}_{k-1}^{\text{dec}})_{k \geq 1}$ and $(\tilde{\mathcal{T}}_k^{\text{dec}} \setminus \tilde{\mathcal{T}}_{k-1}^{\text{dec}})_{k \geq 1}$. In order to identify the law of the sequences $(\mathcal{T}_k^{\text{dec}})_{k \geq 1}$ and $(\tilde{\mathcal{T}}_k^{\text{dec}})_{k \geq 1}$, we just have to check that the attachment procedure is the same.

Recall the definition of the random variables $(I_k)_{k \geq 1}$, $(V_k)_{k \geq 1}$ and $(W_k)_{k \geq 1}$. Still from [34, Lemma 5.5, Proposition 5.7], conditionally on all the quantities whose distributions were identified above, the I_k 's are independent and $\mathbb{P}(I_k = i) = \frac{N_i}{N_1 + \dots + N_{k-1}}$ for all $i \leq k - 1$. From those random variables we can construct an increasing sequence of trees $(S_n)_{n \geq 1}$ in such a way that the parent of the vertex with label k is the vertex with label I_k . From the observation above, the law of $(S_n)_{n \geq 1}$ conditionally on everything else mentioned before only depends on $(N_k)_{k \geq 1}$. This law is the same as that of $(T_n)_{n \geq 1}$ conditionally on $(m_k)_{k \geq 1}$ (and everything else).

Now we just need to check the law of the gluing points: recall how the point Z_k is given as $Z_k = Z_{V_k}^{U_k} = Y_{V_k, W_k}$. We just need to argue that this point, conditionally on $I_k = i$ the point Z_k is taken under a normalized version of the measure $\sum_{\ell=1}^{\infty} \Delta_{t_i(\ell)} \mu_{t_i(\ell)}$. In fact, still from [34], conditionally on $I_k = i$ and everything else mentioned before we

have $\mathbb{P}(V_k = t_i(\ell)) = \frac{N_i(\ell)}{N_i} = \frac{\Delta_{t_i(\ell)}}{\sum_{\ell=1}^{\infty} \Delta_{t_i(\ell)}}$ for $\ell \geq 1$ and W_k is independent uniform on $[0, 1]$. Since by definition, $W_k \in \mathcal{A}_{V_k}$, then Y_{V_k, W_k} is distributed as μ_{V_k} by construction. Now, since the uniform distribution has no atom, it is almost surely the case that $W_k \notin \left\{u_{V_k}^{U_r} \mid r \leq k-1, V_k \leq U_r\right\}$ so that the point Y_{V_k, W_k} has not been used in the construction up to time $k-1$, so the sampling of Z_k is indeed independent of the rest.

6.4 Properties of the construction

We introduce the set of leaves $\mathcal{L} = \mathcal{T}_\alpha^{\text{dec}} \setminus \cup_{n \geq 1} \mathcal{T}_n^{\text{dec}}$. Then still from [34] we have the following

Theorem 6.6. *If $\mathbb{E}[\text{diam}(\mathcal{B})^p] < \infty$ for some $p > \frac{\alpha}{\gamma}$, then $\mathcal{T}_\alpha^{\text{dec}}$ is almost surely a compact metric space and $\mathbb{E}[\text{diam}(\mathcal{T}_\alpha^{\text{dec}})^p] < \infty$. Under the assumption that the measure on \mathcal{B} is not concentrated on its root almost surely, the Hausdorff dimension of the set of leaves \mathcal{L} is given by*

$$\dim_H(\mathcal{L}) = \frac{\alpha}{\gamma}$$

almost surely.

We can now provide a proof of Lemma 4.1 which can just be seen as a corollary of the above theorem.

Proof of Lemma 4.1. Introduce the random block $\widehat{\mathcal{B}}$ defined on the same probability space as \mathcal{B} as the interval $[0, \text{diam } \mathcal{B}]$ rooted at 0 and whose sampling measure is a Dirac point mass at $\text{diam } \mathcal{B}$. We also consider the corresponding object $\widehat{\mathcal{T}}^{\text{dec}}$. Because $\widehat{\mathcal{B}}$ also satisfies the assumptions of the theorem above, $\widehat{\mathcal{T}}^{\text{dec}}$ is almost surely compact and by Corollary 6.5 we have $d_H(\widehat{\mathcal{T}}^{\text{dec}}, \widehat{\mathcal{T}}_n^{\text{dec}}) \rightarrow 0$ almost surely as $n \rightarrow \infty$. Then we can check that, for uniform random variables U_1, U_2, \dots, U_n that were used to define $\widehat{\mathcal{T}}_n^{\text{dec}}$, we have

$$\sup_{s \in [0, 1]} \sum_{t \prec s} \text{diam}(\mathcal{B}_t) \mathbf{1}_{\{\Delta_t < \delta\}} \leq \sup_{s \in \{U_1, U_2, \dots, U_n\}} \sum_{t \prec s} \text{diam}(\mathcal{B}_t) \mathbf{1}_{\{\Delta_t < \delta\}} + d_H(\widehat{\mathcal{T}}^{\text{dec}}, \widehat{\mathcal{T}}_n^{\text{dec}}).$$

Thanks to the above theorem, $\widehat{\mathcal{T}}^{\text{dec}}$ is compact and the second term tends to 0 as $n \rightarrow \infty$. Then, for a large fixed n , the first term tends to 0 as $\delta \rightarrow 0$. □

7 Applications

We present several applications of the invariance principle in Theorem 5.1 to block-weighted models of random discrete structures. The limits of these objects are stable trees decorated by other stable trees, stable looptrees, or even Brownian discs.

7.1 Marked trees and iterated stable trees

Define a class of combinatorial objects \mathcal{M} consisting of all marked rooted plane trees where the root and each leaf receives a mark and internal vertices may or may not receive a mark. The size of an object M in \mathcal{M} is its number of leaves and is denoted by $|M|$. For a given n there are infinitely many trees with size n so for simplicity we assume that there are no vertices of out-degree 1, rendering the number of trees of size n finite. A subclass of \mathcal{M} , which we denote by \mathcal{L} , consists of trees where no internal vertex, except the root, receives a mark, and so an n -sized object from \mathcal{L} is a rooted plane tree with n leaves. The ordinary generating series of the classes \mathcal{M} and \mathcal{L} satisfy the equation

$$\mathcal{M}(z) = z + \mathcal{L}(\mathcal{M}(z)) - \mathcal{M}(z). \tag{7.1}$$

Decorated stable trees

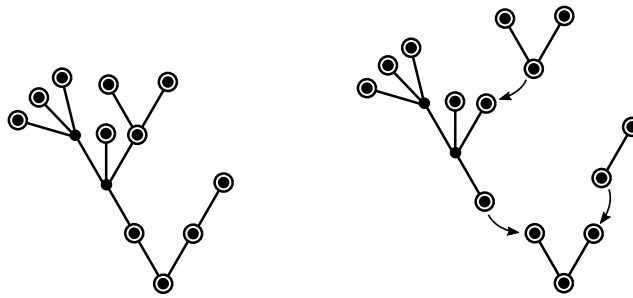


Figure 4: On the left is a tree from \mathcal{M} and on the right it is shown how it is decomposed into its tree blocks. The root is at the bottom and marked vertices are denoted by black and white circles.

The interpretation is that either an object from \mathcal{M} is a single vertex (the root of the tree) or a tree in \mathcal{L} , different from a single root, in which each marked vertex is identified with the root of an object from \mathcal{M} .

We will call a subtree of an element M from \mathcal{M} a *tree block* if it has more than one vertex, all of its leaves are marked, its root is marked and no other vertices are marked. We further require that it is a maximal subtree with this property. The tree blocks may also be understood as the subtrees obtained by cutting the marked tree at each marked internal vertex, see Fig. 4.

Remark 7.1. One may introduce marks with k different colors and define ‘color blocks’ and assign weights to them. This model is a candidate for a discrete version of further iterated trees $\mathcal{T}_{\alpha_1, \alpha_2, \dots, \alpha_{k+1}}$. Denote the set of such structures with marks of k colours by \mathcal{M}_k , with $\mathcal{L} = \mathcal{M}_0$ and $\mathcal{M} = \mathcal{M}_1$. One then has an equation of generating series

$$\mathcal{M}_k(z) = z + \mathcal{M}_{k-1}(\mathcal{M}_k(z)) - \mathcal{M}_k(z). \quad (7.2)$$

Remark 7.2. It is a standard result (and easy to check) that

$$\mathcal{L}(z) = \frac{1 + z - \sqrt{z^2 - 6z + 1}}{4} = z + z^2 + 3z^3 + 11z^4 + 45z^5 + 197z^6 + \dots$$

and from this and Equation (7.1) one may deduce that

$$\mathcal{M}(z) = \frac{1 + 7z - \sqrt{z^2 - 10z + 1}}{12} = z + z^2 + 5z^3 + 31z^4 + 215z^5 + 1597z^6 + \dots$$

This sequence of coefficients is not in the OEIS. Going further one finds that

$$\mathcal{M}_2(z) = \frac{1 + 17z - \sqrt{z^2 - 14z + 1}}{24} = z + z^2 + 7z^3 + 61z^4 + 595z^5 + 6217z^6 + \dots$$

In general, by induction

$$\begin{aligned} \mathcal{M}_k(z) &= \frac{1 + (2k^2 + 4k + 1)z - \sqrt{z^2 - (4k + 6)z + 1}}{2(k + 1)(k + 2)} \\ &= z + z^2 + (2(k + 1) + 1)z^3 + (5(k + 1)^2 + 5(k + 1) + 1)z^4 \\ &\quad + (14(k + 1)^3 + 21(k + 1)^2 + 9(k + 1) + 1)z^5 \dots \end{aligned}$$

Note that $[z^n(k + 1)^m]\mathcal{M}_k(z)$ is the number of rooted plane trees with no vertex of outdegree 1 which have n leaves and m internal vertices (not counting the root). In particular it holds that

$$\sum_{m=0}^{n-2} [z^n(k + 1)^m]\mathcal{M}_k(z) = [z^n]\mathcal{L}(z).$$

To each element L of \mathcal{L} we assign a weight $\gamma(L)$ and denote the class of such weighted structures by \mathcal{L}^γ . We assume the weight of the marked tree consisting of a single vertex is equal to 1. Then, we assign a weighting ω to elements M from \mathcal{M} by

$$\omega(M) = \prod_L \gamma(L)$$

where the index L ranges over the tree blocks in M . Denote the corresponding class of weighted structures by \mathcal{M}^ω . The weighed ordinary generating series satisfy a similar equation as before

$$\mathcal{M}^\omega(z) = z + \mathcal{L}^\gamma(\mathcal{M}^\omega(z)) - \mathcal{M}^\omega(z).$$

Define a random element M_n^ω from the set of n -sized elements of \mathcal{M}^ω which is selected with a probability proportional to its weight. Let $(\iota_n)_{n \geq 0}$ and $(\zeta_n)_{n \geq 0}$ be sequences of non-negative weights, with $\iota_0 = \zeta_1 = 1$ and $\iota_1 = \zeta_1 = 0$. We will be interested in weights γ of the form

$$\gamma(L) = \zeta_{|L|} W(L)$$

with

$$W(L) = \prod_{v \text{ internal vertex in } L} \iota_{d^+(v)}.$$

Let $(w_n)_{n \geq 0}$ be a sequence of non-negative numbers such that $w_0 = 1$ and $w_1 = 0$. For $n \geq 2$ we will write the weights ζ_n as follows

$$\zeta_n = w_n Z_n^{-1}$$

where

$$Z_n = \sum_{L \in \mathcal{L}, |L|=n} W(L)$$

is the *partition function* of elements from \mathcal{L}^γ of size n . In particular, when we choose $\zeta_n = 1$ (i.e. $w_n = Z_n$) for all $n \geq 2$, we say that M_n^ω is an n -leaf *simply generated marked tree* with branching weights $(\iota_k)_k$.

Proposition 7.3. *The random element M_n^ω may be sampled as follows:*

1. *Sample an n -leaf simply generated tree $\tau_n = \tau_n^\omega$ with branching weights $[z^k] \mathcal{L}^\gamma(z) = w_k$ assigned to each internal vertex of out-degree $k > 1$ and branching weight $w_0 = 1$ assigned to each leaf.*
2. *For each vertex v of τ_n sample independently a $d^+(v)$ -leaf simply generated tree $\delta(v)$ with branching weights ι_k assigned to each vertex of outdegree $k \geq 0$. Glue together according to the tree structure of τ_n (see Fig. 5).*

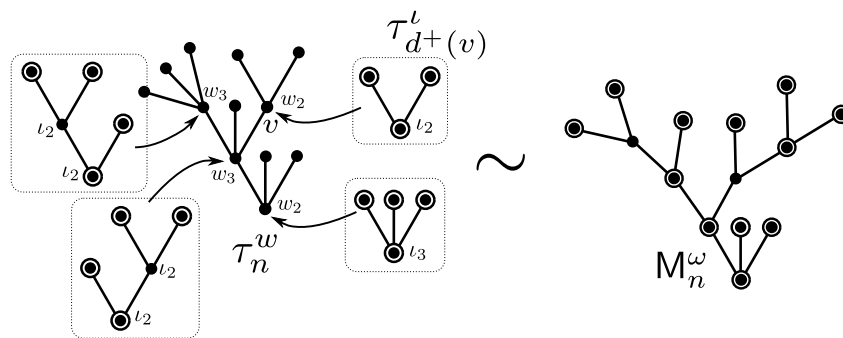


Figure 5: The coupling of M_n^ω with simply generated trees with leaves as atoms.

We refer to the surveys [21, 38] for details on simply generated models of trees with fixed numbers of vertices or leaves. The proof of Proposition 7.3 is by a straightforward calculation, or alternatively by applying a general result on sampling trees with decorations [38, Lem. 6.7]. We will only be interested in the case where $\mathcal{M}^\omega(z)$ has positive radius of convergence. In this case one may rescale the weights w_n and ι_n such that they are probabilities without affecting the distribution of M_n^ω , and we will assume in the following that this has been done. We denote by ξ a random variable with distribution w_n and χ a random variable with distribution ι_n . The tree τ_n may then be viewed as a Bienaymé–Galton–Watson tree with offspring distribution ξ conditioned on having n leaves and $\delta(v)$ may be viewed as a Bienaymé–Galton–Watson tree with offspring distribution χ conditioned on having $d^+(v)$ leaves.

We let $d_{M_n^\omega}$ denote the graph-distance on M_n^ω and ν_{M_n} the counting measure on its set of non-root vertices. Proposition 7.3 ensures that $(M_n^\omega, d_{M_n^\omega}, \nu_{M_n})$ falls into the framework of discrete decorated trees described in Section 3, with the random tree given by τ_n and decorations according to a sequence $(\tilde{B}_k, \tilde{\rho}_k, \tilde{d}_k, \tilde{\ell}_k, \tilde{\nu}_k)_{k \geq 0}$ given as follows. The space \tilde{B}_k is a χ -BGW-tree conditioned on having k leaves, $\tilde{\rho}_k$ is its root vertex, and \tilde{d}_k is the graph distance on that space. The labeling function $\tilde{\ell}_k$ is chosen to be some bijection between $\{1, \dots, k\}$ and the leaves of \tilde{B}_k . Thus μ_k is the uniform measure on the k leaves of \tilde{B}_k . The measure $\tilde{\nu}_k$ is the counting measure on the non-root vertices of \tilde{B}_k .

Choose $(w_n)_{n \geq 0}$ and $(\iota_n)_{n \geq 0}$ such that $\mathbb{E}[\xi] = \mathbb{E}[\chi] = 1$, and such that ξ, χ follow asymptotic power laws so that ξ is in the domain of attraction of a stable law with index $\alpha_2 \in]1, 2[$ and χ lies in the domain of attraction of a stable law with index $\alpha_1 \in]1, 2[$. By [23, Thm. 6.1, Rem. 5.4] there is a sequence

$$b_n = \left(\frac{|\Gamma(1 - \alpha_2)|}{\mathbb{P}(\xi = 0)} \right)^{1/\alpha_2} \inf\{x \geq 0 \mid \mathbb{P}(\xi > x) \leq 1/n\} \sim c_1 n^{1/\alpha_2}$$

for some constant $c_1 > 0$ such that

$$(b_n^{-1} W_{\lfloor \tau_n |t \rfloor}(\tau_n))_{0 \leq t \leq 1} \xrightarrow{d} X^{\text{exc}, (\alpha_2)}.$$

The number of vertices of \tilde{B}_k concentrates at $k/\mathbb{P}(\chi = 0)$, and by [23, Thm. 5.8] it follows that for some constant $c_2 > 0$

$$\left(\tilde{B}_k, \tilde{\rho}_k, c_2 k^{-1-1/\alpha_1} \tilde{d}_k, \mu_k, (k/\mathbb{P}(\chi = 0))^{-1} \tilde{\nu}_k \right) \rightarrow (\mathcal{T}_{\alpha_1}, \rho_{\mathcal{T}_{\alpha_1}}, d_{\mathcal{T}_{\alpha_1}}, \nu_{\mathcal{T}_{\alpha_1}}, \nu_{\mathcal{T}_{\alpha_1}})$$

with $(\mathcal{T}_{\alpha_1}, \rho_{\mathcal{T}_{\alpha_1}}, d_{\mathcal{T}_{\alpha_1}}, \nu_{\mathcal{T}_{\alpha_1}})$ denoting the α_1 -stable tree with root vertex $\rho_{\mathcal{T}_{\alpha_1}}$. For $\alpha_1 > \frac{1}{2-\alpha_2}$ we may use the construction from Section 4 to form a decorated stable tree $(\mathcal{T}_{\alpha_1, \alpha_2}, d_{\mathcal{T}_{\alpha_1, \alpha_2}}, \nu_{\mathcal{T}_{\alpha_1, \alpha_2}})$ with distance exponent $\gamma_1 := 1 - 1/\alpha_1$, obtained by blowing up the branchpoints of the α_2 -stable tree with rescaled independent copies of the α_1 -stable tree. The measure $\nu_{\mathcal{T}_{\alpha_1, \alpha_2}}$ is taken to be the push-forward of the Lebesgue measure, corresponding to the case $\beta = 1$. Theorem 5.1 yields the following scaling limit:

Theorem 7.4. *Choose $(w_n)_{n \geq 0}$ and $(\iota_n)_{n \geq 0}$ such that $\mathbb{E}[\xi] = \mathbb{E}[\chi] = 1$, and such that ξ, χ follow asymptotic power laws so that ξ is in the domain of attraction of a stable law with index $\alpha_2 \in]1, 2[$ and χ lies in the domain of attraction of a stable law with index $\alpha_1 \in]1, 2[$. Suppose that $\alpha_1 > \frac{1}{2-\alpha_2}$. Then*

$$\left(M_n^\omega, c_1^{-1+1/\alpha_1} c_2 n^{-\frac{\alpha_1-1}{\alpha_2 \alpha_1}} d_{M_n^\omega}, \frac{1}{|M_n^\omega|} \nu_{M_n^\omega} \right) \xrightarrow{d} (\mathcal{T}_{\alpha_1, \alpha_2}, d_{\mathcal{T}_{\alpha_1, \alpha_2}}, \nu_{\mathcal{T}_{\alpha_1, \alpha_2}})$$

in the Gromov–Hausdorff–Prokhorov topology.

By Theorem 6.6 and standard arguments we also obtain the Hausdorff-dimension of the iterated stable tree:

Theorem 7.5. *For every $\alpha_2 \in (1, 2)$ and $\alpha_1 \in (\frac{1}{2-\alpha_2}, 2)$, almost surely the random tree $\mathcal{T}_{\alpha_1, \alpha_2}$ has Hausdorff dimension $\frac{\alpha_2 \alpha_1}{\alpha_1 - 1}$.*

The construction may be iterated: With $\gamma_2 := \gamma_1/\alpha_2$ we may choose any $\alpha_3 \in]1, 1 + \gamma_2[$ and build the iterated stable tree $\mathcal{T}_{\alpha_1, \alpha_2, \alpha_3}$ as in Section 4 with distance exponent γ_2 , by blowing up the branchpoints of an α_3 -stable tree by rescaled independent copies of the iterated stable tree $\mathcal{T}_{\alpha_1, \alpha_2}$. By Remark 7.1 and analogous arguments as for Theorem 7.4, the tree $\mathcal{T}_{\alpha_1, \alpha_2, \alpha_3}$ arises as scaling limits of random finite marked trees with diameter having order n^{γ_3} for $\gamma_3 := \gamma_2/\alpha_3$. This construction may be iterated indefinitely often, by choosing $\alpha_4 \in]1, 1 + \gamma_3[$ and setting $\gamma_4 = \gamma_3/\alpha_4$, and so on. This yields a sequence $(\alpha_i)_{i \geq 1}$ so that $\mathcal{T}_{\alpha_1, \dots, \alpha_k}$ is well-defined for all $k \geq 2$. We pose the following question:

Question 7.6. Is there a non-trivial scaling limit for $\mathcal{T}_{\alpha_1, \dots, \alpha_k}$ as $k \rightarrow \infty$?

Note that the associated sequence $(\gamma_i)_{i \geq 1}$ is strictly decreasing and satisfies $\gamma_{i+1} < \gamma_i/(\gamma_i + 1)$, which yields $\lim_{i \rightarrow \infty} \gamma_i = 0$ and hence $\lim_{i \rightarrow \infty} \alpha_i = 1$. The intuition for the question is that if α is close to 1, then the vertex with maximal width in \mathcal{T}_α should dominate and hence blowing up all branchpoints by some decoration should yield something close to a single rescaled version of that decoration. Hence $\mathcal{T}_{\alpha_1, \dots, \alpha_{k+1}}$ should be close to a constant multiple of $\mathcal{T}_{\alpha_1, \dots, \alpha_k}$ for k large enough. Since $\mathcal{T}_{\alpha_1, \dots, \alpha_k}$ has Hausdorff dimension $\alpha_1 \cdots \alpha_k / (\alpha_1 - 1)$, we expect that $(\alpha_i)_{i \geq 1}$ needs to be chosen so that $\prod_{i=1}^\infty \alpha_i$ converges in order to get a scaling limit.

7.2 Weighted outerplanar maps

Planar maps may roughly be described as drawings of connected graphs on the 2-sphere, such that edges are represented by arcs that may only intersect at their endpoints. The connected components that are created when removing the map from the plane are called the faces of the map. The number of edges on the boundary of a face is its degree. In order avoid symmetries, usually a root edge is distinguished and oriented. The origin of the root-edge is called the root vertex. The face to the right of the root edge is called the outer face, the face to the left the root face.

An outerplanar map is a planar map where all vertices lie on the boundary of the outer face. The geometric shape of outerplanar maps has received some attention in the literature [8, 36, 35, 38]. Throughout this section we fix two sequences $\iota = (\iota_k)_{k \geq 3}$ and $\kappa = (\kappa_k)_{k \geq 2}$ of non-negative weights. We are interested in random outerplanar maps that are generated according to κ -weights on their blocks and ι -weights on their faces. Our goal in this section is to describe phases in which we obtain decorated stable trees as scaling limits. Specifically, we will obtain stable trees decorated with looptrees and Brownian trees. This is motivated by the mentioned work [35], which described a phase transition of random face-weighted outerplanar maps from a deterministic circle to the Brownian tree via loop trees. By utilizing the second weight sequence κ we hence obtain a completely different phase diagram.

We will only consider outerplanar maps without multi edges or loop edges. Recall that a block of a graph is a connected subgraph D that is maximal with the property that removing any of its vertices does not disconnect D . Blocks of outerplanar maps are precisely dissections of polygons. We define the weight of a dissection D by

$$\gamma(D) = \kappa_{|D|} \prod_F \iota_{\deg(F)}, \tag{7.3}$$

with $|D|$ denoting the number of vertices of D , the index F ranging over the faces of D , and $\deg(F)$ denoting the degree of the face F . This includes the case where D consists of two vertices joined by a single edge. The weight of an outerplanar map O is then

defined by

$$\omega(O) = \prod_D \gamma(D), \tag{7.4}$$

with the index D ranging over the blocks of O . Given $n \geq 1$ such that at least one n -sized outerplanar map has positive ω -weight, this allows us to define a random outerplanar map O_n that is selected with probability proportional to its ω -weight among the finitely many outerplanar graphs with n vertices. We will only consider the case where infinitely many integers n with this property exist. Likewise, we define the random dissection D_n that is sampled with probability proportional to its weight among all dissections with n vertices.

The random outerplanar map O_n fits into the framework considered in the present work. We will make this formal. For each $k \geq 2$ let us set

$$d_k = \sum_{D:|D|=k} \gamma(D), \tag{7.5}$$

with the sum index D ranging over all dissections of k -gons. We will only consider the case where the power series $D(z) := \sum_{k \geq 2} d_k z^k$ has positive radius of convergence $\rho_D > 0$. For any $t > 0$ with $D(t) < t$ we define the probability weight sequence

$$(p_i(t))_{i \geq 0} = (1 - D(t)/t, 0, d_2 t, d_3 t^2, d_4 t^3, \dots). \tag{7.6}$$

Its first moment is given by $\sum_{k \geq 2} k d_k t^{k-1}$, and we set

$$m_O = \lim_{t \nearrow \rho_D} \sum_{k \geq 2} k d_k t^{k-1} \in]0, \infty]. \tag{7.7}$$

If $m_O \geq 1$, then there is a unique $0 < t_O \leq \rho_D$ for which $\sum_{k \geq 2} k d_k t_O^{k-1} = 1$. If $m_O < 1$ we set $t_O = \rho_D$. Furthermore, we let ξ denote a random non-negative integer with probabilities

$$\mathbb{P}(\xi = i) = p_i(t_O), \quad i \geq 0. \tag{7.8}$$

Let T_n denote a ξ -BGW tree conditioned on having n leaves. Note that T_n has no vertex with outdegree 1. Let $g_n > 0$ be a positive real number. For each $k \geq 2$ let $(\tilde{B}_k, \tilde{d}_k, \tilde{\rho}_k, \tilde{\ell}_k, \tilde{\nu}_{n,k})$ denote the decoration with $\tilde{B}_k = D_k$, \tilde{d}_k the graph distance on D_k multiplied by g_n , $\tilde{\rho}_k$ the root vertex of D_k , $\tilde{\ell}_k : \{1, \dots, k\} \rightarrow D_k$ any fixed enumeration of the k vertices of D_k , and $\tilde{\nu}_k$ the counting measure on the non-root vertices (that is, all vertices except for the origin of the root edge) of D_k . We let \tilde{B}_0 denote a trivial one point space with no mass.

Lemma 7.7 ([38, Lem. 6.7, Sec. 6.1.4]). *The decorated tree $(T_n^{dec}, d_n^{dec}, \nu_n^{dec})$ is distributed like the random weighted outerplanar map O_n equipped with the graph distance multiplied by g_n and the counting measure on the non-root vertices.*

See Figure 6 for an illustration. We refer to the cited sources for detailed justifications.

Similarly as for parameter m_O , we let ρ_ι denote the radius of convergence of the series $\sum_{k \geq 2} \iota_{k+1} z^k$ and set

$$m_D := \lim_{t \nearrow \rho_\iota} \sum_{k \geq 2} k \iota_{k+1} t^{k-1} \in]0, \infty].$$

Decorated stable trees

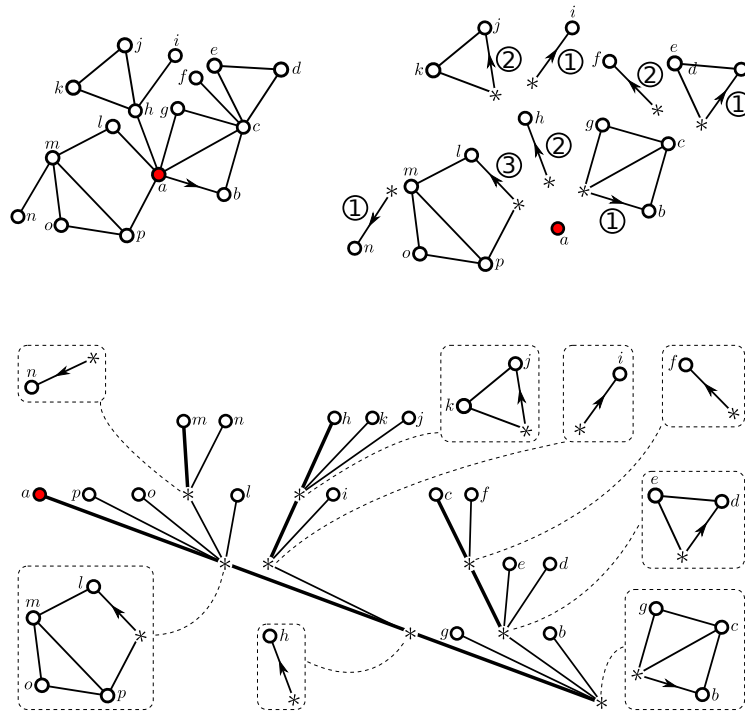


Figure 6: Correspondence of an outerplanar map (top left corner) to a tree decorated by dissections of polygons (bottom half). (Source of image: [35].)

If $m_D \geq 1$, there is a unique constant $0 < t_D \leq \rho_\iota$ such that $\sum_{k \geq 2} k \iota_{k+1} t_D^{k-1} = 1$. For $m_D < 1$ we set $t_D = \rho_\iota$. Furthermore, we let ζ denote a random non-negative integer with distribution given by the probability weight sequence

$$(q_i)_{i \geq 0} := \left(1 - \sum_{k \geq 2} \iota_{k+1} t_D^{k-1}, 0, \iota_3 t_D, \iota_4 t_D^2, \dots \right).$$

We list three known non-trivial scaling limits for the random face-weighted dissection D_n . If $m_D > 1$, then by [9] D_n lies in the universality class of the Brownian tree \mathcal{T}_2 , with a diameter of order \sqrt{n} . (That is, \mathcal{T}_2 is given by $\sqrt{2}$ times the random real tree obtained by identifying points of a Brownian excursion of duration 1.) Specifically,

$$\frac{\sqrt{2}}{\sqrt{\mathbb{V}[\zeta] q_0}} \frac{1}{4} \left(\mathbb{V}[\zeta] + \frac{q_0 \mathbb{P}(\zeta \in 2\mathbb{N}_0)}{2\mathbb{P}(\zeta \in 2\mathbb{N}_0) - q_0} \right) \frac{1}{\sqrt{n}} D_n \xrightarrow{d} \mathcal{T}_2. \quad (7.9)$$

If $m_D = 1$ and $\mathbb{P}(\zeta \geq n) \sim c_D n^{-\eta}$ for $1 < \eta < 2$ and $c_D > 0$, then by [10] D_n lies in the universality class of the η -stable looptree \mathcal{L}_η :

$$(c_D q_0 |\eta(1-\eta)|)^{1/\eta} \frac{1}{n^{1/\eta}} D_n \xrightarrow{d} \mathcal{L}_\eta \quad (7.10)$$

in the Gromov–Hausdorff–Prokhorov sense. Here and in the following we use a shortened notation, where the product of a real number and a random finite graph refers to the vertex set space with the corresponding rescaled graph metric and the uniform measure on the set of vertices.

If $0 < m_D < 1$ and $\mathbb{P}(\zeta = n) = f(n)n^{-\theta}$ for some constant $\theta > 2$ and a slowly varying function f , then D_n lies in the universality class of a deterministic loop. That is,

$$\frac{q_0}{n(1 - \mathbb{E}[\zeta])} D_n \xrightarrow{d} S^1 \quad (7.11)$$

with S^1 denoting the 1-sphere. See for example [35] for a detailed justification.

Knowing the asymptotic behaviour of D_n allows us to describe the asymptotic shape of the random weighted outerplanar map O_n :

- If $m_O > 1$, then by [38, Thm. 6.60] there is a constant $c_\omega > 0$ (defined in a complicated manner using expected distances in bi-pointed Boltzmann dissections) such that

$$c_\omega \frac{1}{\sqrt{n}} O_n \xrightarrow{d} \mathcal{T}_2. \tag{7.12}$$

In this case the asymptotics of weighted dissections only influence the constant c_ω , but not the universality class.

- If $m_O = 1$ and $\mathbb{P}(\xi \geq n) \sim c_O n^{-\alpha}$ for $1 < \alpha < 2$, then T_n lies in the universality class of an α -stable tree. That is, by [23, Thm. 6.1, Rem. 5.4] there is a sequence

$$b_n = \left(\frac{|\Gamma(1 - \alpha)|}{\mathbb{P}(\xi = 0)} \right)^{1/\alpha} \inf\{x \geq 0 \mid \mathbb{P}(\xi > x) \leq 1/n\}$$

of order $n^{1/\alpha}$ for which

$$(b_n^{-1} W_{\lfloor T_n | t \rfloor}(T_n))_{0 \leq t \leq 1} \xrightarrow{d} X^{\text{exc}, (\alpha)}.$$

Now, the outcome of applying Theorem 5.1 depends on the decoration. Suppose that $\gamma > \alpha - 1$. Then in each of the three discussed cases (7.9), (7.10), and (7.11) we obtain a scaling limit of the form

$$\frac{1}{b_n^\gamma} O_n \xrightarrow{d} \mathcal{T}_\alpha^{\text{dec}}. \tag{7.13}$$

Specifically:

a) If $m_D > 1$ then $\gamma = 1/2$ and $\mathcal{T}_\alpha^{\text{dec}}$ is the α -stable tree decorated according to a constant multiple of the Brownian tree \mathcal{T}_2 (with the constant given by the inverse of the scaling factor in (7.9)).

b) If $m_D = 1$ and $\mathbb{P}(\zeta \geq n) \sim c_D n^{-\eta}$ for $1 < \eta < 2$ and $c_D > 0$, then $\gamma = 1/\eta$ and $\mathcal{T}_\alpha^{\text{dec}}$ is the α -stable tree decorated according to the stretched η -stable looptree $(c_D q_0 |\eta(1 - \eta)|)^{-1/\eta} \mathcal{L}_\eta$.

c) If $0 < m_D < 1$ and $\mathbb{P}(\zeta = n) = f(n)n^{-\theta}$ for some constant $\theta > 2$ and a slowly varying function f , then $\gamma = 1$ and $\mathcal{T}_\alpha^{\text{dec}}$ is the α -stable tree decorated according to the stretched circles $\frac{n(1 - \mathbb{E}[\zeta])}{q_0} S^1$. In other words, $\mathcal{T}_\alpha^{\text{dec}}$ is distributed like the stretched α -stable loop tree $\frac{n(1 - \mathbb{E}[\zeta])}{q_0} \mathcal{L}_\alpha$.

Here condition B1 (which is necessary for applying Theorem 5.1) may be verified using Proposition 3.2 and Remark 3.5. The necessary sufficiently high uniform integrability of the diameter of the decorations is trivial in case c), and follows from [38, Lem. 6.61, Sec. 6.1.3] in case a), and analogously from tail-bounds of conditioned BGW trees [24] in case b).

- Finally, if $0 < m_O < 1$ and $\mathbb{P}(\xi = n) = f_O(n)n^{-\theta_O}$ for some constant $\theta_O > 2$ and a slowly varying function f , then O_n has giant 2-connected component of size about $n(1 - \mathbb{E}[\xi])/\mathbb{P}(\xi = 0)$. Hence, if D_n admits a scaling limit like in the three discussed cases, then the scaling limit for M_n is, up to a constant multiplicative factor, the same as for D_n . See for example [35] for details on such approximations.

7.3 Weighted planar maps with a boundary

A planar map with a boundary refers to a planar map where the outer face typically plays a special role, and the perimeter (number of half-edges on the boundary of the outer face) serves as a size parameter. The reason for counting half-edges instead of edges is that both sides of an edge may be adjacent to the outer face, and in this case such an edge is counted twice. We say the boundary of a planar map is simple, if it is a cycle. Here we explicitly allow the case of degenerate 2-cycles and 1-cycles. That is, a map consisting of two vertices joined by single edge has a simple boundary. A map consisting of a loop with additional structures on the inside has a simple boundary, and so has a map consisting of two vertices joined by two edges and additional structures on the inside.

Planar maps with a boundary fit into the framework of decorated trees by identical arguments as for outerplanar maps. That is, we may decompose a planar map into its components with a simple boundary in the same way as an outerplanar map may be decomposed into dissections of polygons. They may be bijectively encoded as trees decorated by maps with a simple boundary in the same way as illustrated in Figure 6. Since we allow multi-edges and loops, the leaves of the encoding tree canonically and bijectively correspond to the half-edges on the boundary of the map.

There are infinitely many planar maps with a fixed positive perimeter, hence it makes sense to assign summable weights. Say, for each planar map D with a simple boundary we are given a weight $\gamma(D) \geq 0$. Like in Equation (7.4), we then define the weight $\omega(M)$ of a planar map M by

$$\omega(M) = \prod_D \gamma(D), \tag{7.14}$$

with the index D ranging over the components of M with a simple boundary. Thus, for any positive integer n for which the sum of all ω -weights of n -perimeter maps is finite and non-zero we may define the random n -perimeter map M_n that is drawn with probability proportional to its weight. Note that this formally encompasses the random outerplanar map O_n , for which $\gamma(D) = 0$ whenever D is not a dissection of a polygon of perimeter at least 3.

Using the same definitions (7.5)–(7.8) for m_O and ξ , it follows that the tree T_n corresponding to the random map M_n is distributed like the result of conditioning a ξ -BGW tree T_n on having n leaves. Furthermore, employing analogous definitions for the decorations $(\tilde{B}_k, \tilde{d}_k, \tilde{\rho}_k, \tilde{\ell}_k, \tilde{\nu}_k)_{k \geq 0}$, it follows like in Lemma 7.7 that M_n with distances rescaled by some scaling sequence $a_n > 0$ is a discrete decorated tree:

Corollary 7.8. *The decorated tree $(T_n^{dec}, d_n^{dec}, \nu_n^{dec})$ is distributed like the random weighted map M_n with perimeter n , equipped with the graph distance multiplied by a_n and the counting measure on the non-root vertices.*

The boundary of M_n also fits into this framework but its scaling limits have already been determined in pioneering works [31, 26], with the parameter m_O and the tail of ξ ultimately determining whether its a deterministic loop, a random loop tree, or the Brownian tree.

In order to apply Theorem 5.1 to M_n (as opposed to its boundary) we need to be able to look inside of the components, that is, we need a description of the asymptotic geometric shape of Boltzmann planar maps with a large simple boundary. However no such results appear to be known.

What is known by [31] is that so-called non-generic critical face weight sequences with parameter $\alpha' \in]1, 3/2[$ lie in a stable regime with $m_O = 1$ and T_n^{dec} in the universality class of an α -stable tree for $\alpha = 1/(\alpha' - 1/2)$. The boundary of M_n lies in the universality

class of an α -stable looptree by [31, Thm. 1.1]. It is natural to wonder whether additional knowledge of Boltzmann planar maps with a simple boundary in this regime enables the application of Theorem 5.1. This motivates the following question:

Question 7.9. Do decorated stable trees constructed in the present work arise as scaling limits of face-weighted Boltzmann planar maps with a boundary?

We note that there is a connection to the topic of stable planar maps arising as scaling limits (along subsequences) of Boltzmann planar maps without a boundary [27, 29]. Roughly speaking, Boltzmann planar maps with a large boundary are thought to describe the asymptotic geometric behaviour of macroscopic faces of Boltzmann planar maps without a boundary.

There are results for related models of triangulations and quadrangulations with a simple boundary with an additional weight on the vertices [19, 1]. Scaling limits for models with a non-simple boundary have been determined for the special case of uniform quadrangulations with a boundary [6, 4] that are condition on both the number of vertices and the boundary length.

For Boltzmann triangulations with the vertex weights as in [1] we would expect to be in the regime $m_O < 1$ where the shape of the map is dominated by a unique macroscopic component with a simple boundary. However, we could introduce block-weights as before in order to force the model into a stable regime. At least in principle, Theorem 5.1 then yields convergence towards a decorate stable tree obtained by blowing up the branchpoints of a stable tree by rescaled Brownian discs. Checking the requirements of Theorem 5.1 (such as verifying convergence of the moments of the rescaled diameter of Boltzmann triangulations with a simple boundary) does not appear involve any mayor obstacles, but we did not go through the details. To do so, we would need to recall extensive combinatorial background, and this appears to be beyond the scope of the present work.

A Appendix

This appendix contains the proofs of two technical statements that are used in the paper, Proposition 3.2 and Lemma 3.4, which ensure that the main assumptions under which the rest of the paper is stated are satisfied for reasonable models of decorated BGW trees. Section A.1 recalls some general results about regularly varying functions and domain of attraction of stable random variables. Section A.2 presents some estimates related to critical BGW trees whose reproduction law is in the domain of attraction of a stable law that are useful later on. In Section A.3, we introduce the notion of trees with marks on the vertices and prove a spine decomposition result for those marked trees. Finally in Section A.4 we prove Proposition 3.2, which is the main technical result of this Appendix, using all the results and estimates derived before. At the end, in Section A.5, which is independent of what comes before, we prove Lemma 3.4.

A.1 Regularly varying functions and domains of attraction

A.1.1 Compositional inverse of regularly varying functions

Following standard terminology, we say that a function f defined on a neighbourhood of infinity is regularly varying (at infinity) with exponent $\alpha \in \mathbb{R}$ if for every $\lambda \in \mathbb{R} \setminus \{0\}$ we have

$$\frac{f(\lambda x)}{f(x)} \rightarrow \lambda^\alpha.$$

We consider those functions up to the equivalence relation of having a ratio tending to 1 at infinity. When the index of regularity is positive $\alpha > 0$, a regularly varying function

f with exponent α tends to infinity and we can define (at least at a neighbourhood of infinity):

$$f^{[-1]}(x) := \inf \{y \in \mathbb{R}_+ \mid f(y) \geq x\},$$

and the equivalence class of $f^{[-1]}$ only depends on the equivalence class of f . Then $f^{[-1]}$ is a regularly varying function with index α^{-1} and satisfies:

$$f \circ f^{[-1]}(x) \underset{x \rightarrow \infty}{\sim} f^{[-1]} \circ f(x) \underset{x \rightarrow \infty}{\sim} x.$$

A.1.2 Asymmetric stable random variable

For $\alpha \in (0, 1) \cup (1, 2)$, we let Y_α be a random variable with so-called asymmetric stable law of index α , with distribution characterized by its Laplace transform, for all $\lambda > 0$,

$$\begin{aligned} \mathbb{E}[\exp(-\lambda Y_\alpha)] &= \exp(-\lambda^\alpha) \quad \text{if } 0 < \alpha < 1, \\ &= \exp(\lambda^\alpha) \quad \text{if } 1 < \alpha < 2. \end{aligned}$$

It is ensured by [41, Eq.(I20)] that those distributions have a density d_α and that this density is continuous and bounded.

Domain of attraction Let X be a random variable such that $\mathbb{P}(|X| > x)$ is regularly varying with index $-\alpha$, centred if $\alpha \in (1, 2)$. Consider a sequence X_1, X_2, \dots of i.i.d random variables with the law of X . Suppose that

$$\frac{\mathbb{P}(X > x)}{\mathbb{P}(|X| > x)} \rightarrow 1.$$

Then, let

$$B_X(x) := \left| \frac{1 - \alpha}{\Gamma(2 - \alpha)} \right|^{-\frac{1}{\alpha}} \left(\frac{1}{\mathbb{P}(|X| \geq x)} \right)^{[-1]}. \tag{A.1}$$

We also denote this function by B_ν if ν is the law of X . Since $x \mapsto \frac{1}{\mathbb{P}(|X| \geq x)}$ is α -regularly varying, B_X is α^{-1} -regularly varying. Then, by [40, Theorem 4.5.1], we have:

$$\frac{1}{B_X(n)} \cdot \sum_{i=1}^n X_i \xrightarrow[n \rightarrow \infty]{(d)} Y_\alpha,$$

where Y_α has the asymmetric α -stable law, which we recall has a density d_α . In fact, if the random variable X is integer-valued (and not supported on a non-trivial arithmetic progression), a more precise version of the above convergence known as *local limit theorem*, is true, see [20, Theorem 4.2.1]. Let $S_n = \sum_{i=1}^n X_i$, then we have

$$\sup_{k \in \mathbb{Z}} \left| B_X(n) \mathbb{P}(S_n = k) - d_\alpha \left(\frac{k}{B_X(n)} \right) \right| \xrightarrow[n \rightarrow \infty]{} 0. \tag{A.2}$$

Using the fact that the asymmetric α -stable density d_α is bounded, note that the above convergence ensures in particular that there exists a constant C so that for all $n \geq 1$ and all $k \in \mathbb{Z}$ we have

$$\mathbb{P}(S_n = k) \leq \frac{C}{B_X(n)}. \tag{A.3}$$

A.1.3 The Potter bounds

From [7, Theorem 1.5.6(iii)]: Let f be a regularly varying function of index ρ . Then for every $A > 1$ and ϵ , there exists B such that for all $x, y \geq B$ we have

$$\frac{f(y)}{f(x)} \leq A \cdot \max \left\{ \left(\frac{y}{x}\right)^{\rho-\epsilon}, \left(\frac{y}{x}\right)^{\rho+\epsilon} \right\}. \tag{A.4}$$

When f is defined on the whole interval $(0, \infty)$ and bounded below by a positive constant, we can increase the constant A so that the last display holds for any $x, y \geq 1$.

Let us apply this to B_X , the slowly varying function associated to some positive random variable X in the domain of attraction of an θ -stable law. For all $\epsilon > 0$ there exists a constant C such that for all $n \geq 1$ and all $\frac{1}{B_X(n)} \leq \lambda \leq 1$,

$$C^{-1} \cdot n \cdot \lambda^{\theta+\epsilon} \leq B_X^{[-1]}(\lambda B_X(n)) \leq C \cdot n \cdot \lambda^{\theta-\epsilon}, \tag{A.5}$$

and so that in particular, possibly for another constant C ,

$$C^{-1} \cdot \frac{1}{n} \cdot \lambda^{-\theta+\epsilon} \leq \mathbb{P}(X \geq \lambda B_X(n)) \leq C \cdot \frac{1}{n} \cdot \lambda^{-\theta-\epsilon}, \tag{A.6}$$

A.2 Estimates for Bienaymé–Galton–Watson trees with α -stable tails

Let μ be a critical reproduction law in the domain of attraction of an α -stable distribution, with $\alpha \in (1, 2)$, and $b_n = B_\mu(n)$ the associated $\frac{1}{\alpha}$ -regularly varying function. Recall that d_α is the density function of the random variable Y_α . The following lemma contains all the results that we need in the subsequent sections of the appendix.

Lemma A.1. *Let D be a random variable with distribution μ^* the size-biased version of μ and $(\tau_i)_{i \geq 1}$ be independent BGW trees with reproduction law μ . Then, for any $\eta > 0$, there exists a constant C such that for all $n \geq 1$ and for $\lambda \in [\frac{1}{b_n}, 1]$, we have*

(i) $\mathbb{P}(D \geq \lambda b_n) \leq C \cdot \lambda^{1-\alpha-\eta} \cdot \frac{b_n}{n},$

(ii) For all $k \in \mathbb{Z}$,

$$\mathbb{P}\left(\sum_{i=1}^{\lambda b_n} |\tau_i| = k\right) \leq C \cdot \frac{\lambda^{-\alpha-\eta}}{n}.$$

(iii) $\mathbb{E}\left[\frac{1}{\sum_{i=1}^{\lambda b_n} |\tau_i|}\right] \leq C \cdot \frac{\lambda^{-\alpha-\eta}}{n},$

Proof of Lemma A.1. We first prove (i). We have using [24, Eq.(45)]

$$\mathbb{P}(D \geq k) = \mu^*([k, \infty)) \underset{k \rightarrow \infty}{\sim} \frac{\alpha}{\alpha - 1} \cdot k \cdot \mu([k, \infty)),$$

so for C a constant chosen large enough we get that for all $k \geq 1$,

$$\mathbb{P}(D \geq k) \leq C \cdot k \cdot \mu([k, \infty)).$$

Applying this to $k = \lceil \lambda b_n \rceil$ and using the Potter bounds yields, for a value of C that may change from line to line,

$$\begin{aligned} \mathbb{P}(D \geq \lambda b_n) &\leq C \cdot \lceil \lambda b_n \rceil \cdot \mu([\lambda b_n, \infty)) \leq C \cdot \lceil \lambda b_n \rceil \cdot \frac{1}{n} \cdot \lambda^{-\alpha-\eta} \\ &\leq C \cdot \frac{b_n}{n} \cdot \lambda^{1-\alpha-\eta}. \end{aligned}$$

So (i) is proved. Now let us turn to (ii). Using, for example, [24, Eq.(26)], we have

$$\mathbb{P}(|\tau| = n) \underset{n \rightarrow \infty}{\sim} \frac{d_\alpha(0)}{nB_\mu(n)} \quad \text{and so} \quad \mathbb{P}(|\tau| \geq n) \underset{n \rightarrow \infty}{\sim} \frac{\alpha d_\alpha(0)}{B_\mu(n)}. \tag{A.7}$$

This ensures that the random variable $|\tau|$ is in the domain of attraction of the asymmetric $1/\alpha$ -stable random variable. We can also check from the last display and the definition (A.1) that

$$B_{|\tau|}(n) \sim C \cdot B_\mu^{[-1]}(n), \tag{A.8}$$

for some constant C . Using (A.3) we then get

$$\mathbb{P}\left(\sum_{i=1}^{\lambda b_n} |\tau_i| = k\right) \leq \frac{C}{B_\mu^{[-1]}(\lambda b_n)} \leq C \cdot \frac{\lambda^{-\alpha-\eta}}{n},$$

where the last inequality follows from (A.5). This finishes the proof of (ii). The last point (iii) follows from an application of Lemma A.2 stated below to the distribution of $|\tau|$ using (A.8) and an application of the Potter bounds to conclude. \square

Lemma A.2. *Let X be a distribution on $[1, \infty)$, in the domain of attraction of an asymmetric θ -stable law, with $\theta \in (0, 1)$, and X_1, X_2, \dots i.i.d random variables with the law of X . Then, there exists a constant C such that for all $N \geq 1$,*

$$\mathbb{E}\left[\frac{1}{\sum_{i=1}^N X_i}\right] \leq \frac{C}{B_X(N)}.$$

Proof. Write

$$\mathbb{E}\left[\frac{1}{\sum_{i=1}^N X_i}\right] = \frac{1}{B_X(N)} \cdot \mathbb{E}\left[\frac{B_X(N)}{\sum_{i=1}^N X_i}\right] \leq \frac{1}{B_X(N)} \cdot \mathbb{E}\left[\frac{B_X(N)}{\max_{1 \leq i \leq N} X_i}\right].$$

Then

$$\begin{aligned} \mathbb{E}\left[\frac{B_X(N)}{\max_{1 \leq i \leq N} X_i}\right] &= \int_0^\infty \mathbb{P}\left(\frac{B_X(N)}{\max_{1 \leq i \leq N} X_i} > x\right) dx \\ &= \int_0^\infty \mathbb{P}\left(\max_{1 \leq i \leq N} X_i < x^{-1} \cdot B_X(N)\right) dx. \end{aligned}$$

Since the values of X_i are at least 1, the integrand of the last display becomes 0 when $x > B_X(N)$. But when $x \leq B_X(N)$ we have

$$\begin{aligned} \mathbb{P}\left(\max_{1 \leq i \leq N} X_i < x^{-1} \cdot B_X(N)\right) &\leq \mathbb{P}(X < x^{-1} \cdot B_X(N))^N \\ &\leq (1 - \mathbb{P}(X \geq x^{-1} B_X(N)))^N \\ &\leq \exp(-N \cdot \mathbb{P}(X \geq x^{-1} B_X(N))) \\ &\leq \exp\left(-N \cdot C \cdot \frac{x^{\theta/2}}{N}\right) \leq \exp\left(-C \cdot x^{\theta/2}\right), \end{aligned}$$

which is integrable over \mathbb{R} . In the last line, we use that

$$\mathbb{P}(X \geq x^{-1} B_X(N)) \geq C \cdot \frac{1}{N} \cdot (x^{-1})^{-\frac{\theta}{2}}$$

for some constant C that does not depend on N , thanks to the Potter bounds with $\epsilon = \theta/2$, which apply here because $x^{-1} \geq \frac{1}{B_X(N)}$. \square

A.3 Spine decomposition for marked Bienaymé–Galton–Watson trees

Definitions We first recall a few definitions from Section 2.1. We have $\mathbb{U} = \bigcup_{n \geq 0} \mathbb{N}^n$ the Ulam tree, and we denote \mathbb{T} the set of plane trees. For any $\mathbf{t} \in \mathbb{T}$, and $v \in \mathbf{t}$, we define $\theta_v(\mathbf{t}) = \{u \in \mathbb{U} \mid vu \in \mathbf{t}\}$. For $w \in \mathbb{U}$, we let $w\mathbf{t} = \{wv \mid v \in \mathbf{t}\}$. If $\mathbf{t} \in \mathbb{T}$ and $u \in \mathbf{t}$, then we define $d_{\mathbf{t}}^+(u)$ to be the number of children of u in \mathbf{t} . We drop the index if it is clear from the context. If $v = v_1 \dots v_n$ we say that $|v| = n$ is the height of v , and for any $k \leq n$, we set $[v]_k = v_1 \dots v_k$. If $v \neq \emptyset$ we also let $\hat{v} = [v]_{n-1}$ be the father of v . We define $\mathbb{T}^* = \{(\mathbf{t}, v) \mid \mathbf{t} \in \mathbb{T}, v \in \mathbf{t}\}$ the set of plane trees with a marked vertex.

Let μ be a probability measure on the integers, such that $\sum_{k \geq 0} k\mu(k) = 1$. The associated probability measure BGW_μ on plane trees is the measure such that for T a random tree taken under this measure, we have for every $\mathbf{t} \in \mathbb{T}$, finite:

$$\mathbb{P}(T = \mathbf{t}) = \text{BGW}_\mu(\{\mathbf{t}\}) = \prod_{u \in \mathbf{t}} \mu(d_{\mathbf{t}}^+(u)). \tag{A.9}$$

Marks on a tree Let E be any measurable space. A *finite tree with marks* is an ordered pair $\tilde{\mathbf{t}} = (\mathbf{t}, (x_u)_{u \in \mathbf{t}})$ where $\mathbf{t} \in \mathbb{T}$ and $(x_u)_{u \in \mathbf{t}}$ are elements of E (corresponding to marks of the vertices). We denote by $\tilde{\mathbb{T}}$ the space of marked trees, and similarly, we let $\tilde{\mathbb{T}}^*$ the set of marked trees with a distinguished vertex.

Let $(\pi_k)_{k \geq 0}$ be a sequence of probability measures on E . A natural way to put random marks on a tree \mathbf{t} is to mark it with $(X_u)_{u \in \mathbf{t}}$ that is a family of independent random variables with respective distribution $X_u \sim \pi_{d_{\mathbf{t}}^+(u)}$.

Cutting a tree at a vertex If $\mathbf{t} \in \mathbb{T}$, and $v \in \mathbf{t}$, we define:

$$\text{Cut}(\mathbf{t}, v) := \mathbf{t} \setminus \{vu \mid u \in \mathbb{U}, u \neq \emptyset\} = \mathbf{t} \setminus (v\theta_v\mathbf{t}) \cup \{v\}. \tag{A.10}$$

The Kesten tree We let

- $(D_i)_{i \geq 0}$ be a sequence of i.i.d. random variables with distribution μ^* , where $\mu^*(k) = k\mu(k)$, for any $k \geq 0$,
- $(K_i)_{i \geq 1}$, which conditionally on the sequence (D_i) are uniform on $\llbracket 1, D_i \rrbracket$,
- $\mathbf{U}_n^* = K_1 K_2 \dots K_n$.
- $(\tau_u)_{u \in \mathbb{U}}$ are BGW trees with reproduction law μ .
- $\mathcal{H} = \{\mathbf{U}_n^* \mid n \geq 0, i \in \llbracket 1, D_i \rrbracket \setminus K_i\}$.

Then the Kesten tree is defined as

$$\mathbf{T}^* = \{\mathbf{U}_n^* \mid n \geq 0\} \cup \bigcup_{u \in \mathcal{H}} u\tau_u.$$

Cut Kesten tree Let τ be a BGW tree with offspring distribution μ , independent of everything else. We let:

$$\mathbf{T}_k^* := \text{Cut}(\mathbf{T}^*, \mathbf{U}_k^*) \cup (\mathbf{U}_k^* \tau),$$

which is almost surely a finite tree.

As before, we endow the finite random tree \mathbf{T}_k^* with some marks $(X_u)_{u \in \mathbf{T}_k^*}$ such that conditionally on \mathbf{T}_k^* , the marks are independent with respective conditional distribution $X_u \sim \pi_{d^+(u)}$. We denote by $\tilde{\mathbf{T}}_k^*$ the obtained tree with marks.

The ancestral line of a vertex We let

$$\mathbb{A} := \bigcup_{n \geq 1} (\mathbb{N}_0 \times E)^n.$$

If $\tilde{\mathbf{t}} = (\mathbf{t}, (x_u)_{u \in \mathbf{t}}) \in \mathbb{T}$, and $v = v_1 v_2 \dots v_k \in \mathbf{t}$ we define:

$$A(\tilde{\mathbf{t}}, v) := ((d_{\mathbf{t}}^+([v]_0), x_{[v]_0}), (d_{\mathbf{t}}^+([v]_1), x_{[v]_1}), \dots, (d_{\mathbf{t}}^+([v]_k), x_{[v]_k})) \in \mathbb{A},$$

the marked ancestral line of vertex v in the marked tree $\tilde{\mathbf{t}}$.

BGW tree with a uniform vertex satisfying some property Let $\mathcal{P} \subset \mathbb{A}$ be some property. Consider $\tilde{T} = (T, (X_u)_{u \in T})$ a marked BGW tree with reproduction distribution μ and mark distributions $(\pi_k)_{k \geq 0}$, and on the event $\{\exists v \in T : A(\tilde{T}, v) \in \mathcal{P}\}$ where at least one vertex of \tilde{T} has the property \mathcal{P} (which depends only on its ancestral line and the marks along it), we let U be a uniform vertex such that this is the case.

Proposition A.3. For any non-negative function $F : \tilde{\mathbb{T}}^* \rightarrow \mathbb{R}$, we have

$$\mathbb{E} \left[F(\tilde{T}, U) \mathbf{1}_{\{\exists v \in T : A(\tilde{T}, v) \in \mathcal{P}\}} \right] = \sum_{k \geq 0} \mathbb{E} \left[\frac{F(\tilde{\mathbf{T}}_k^*, \mathbf{U}_k^*) \mathbf{1}_{\{A(\tilde{\mathbf{T}}_k^*, \mathbf{U}_k^*) \in \mathcal{P}\}}}{\#\{v \in \mathbf{T}_k^* \mid A(\tilde{\mathbf{T}}_k^*, v) \in \mathcal{P}\}} \right]. \tag{A.11}$$

Proof. For any \mathbf{t} and $u \in \mathbf{t}$ of height k , we have

$$\begin{aligned} & \mathbb{P}((\mathbf{T}_k^*, \mathbf{U}_k^*) = (\mathbf{t}, u)) \\ &= \prod_{1 \leq i \leq k} \mathbb{P}(D_i = d_{\mathbf{t}}^+([u]_{i-1})) \cdot \mathbb{P}(K_i = u_i \mid D_i = d_{\mathbf{t}}^+([u]_{i-1})) \cdot \prod_{v \in \mathbf{t}, v \neq u} \mu(d_{\mathbf{t}}^+(v)) \\ &= \left(\prod_{1 \leq i \leq k} d_{\mathbf{t}}^+([u]_{i-1}) \mu(d_{\mathbf{t}}^+([u]_{i-1})) \frac{1}{d_{\mathbf{t}}^+([u]_{i-1})} \right) \cdot \prod_{v \in \mathbf{t}, v \neq u} \mu(d_{\mathbf{t}}^+(v)) \\ &= \prod_{v \in \mathbf{t}} \mu(d_{\mathbf{t}}^+(v)) \\ &= \mathbb{P}(T = \mathbf{t}). \end{aligned} \tag{A.12}$$

We now expand the expectation as a sum

$$\mathbb{E} \left[F(\tilde{T}, U) \mathbf{1}_{\{\exists v \in T : A(\tilde{T}, v) \in \mathcal{P}\}} \right] = \sum_{\mathbf{t} \in \mathbb{T}, u \in \mathbf{t}} \mathbb{E} \left[F(\tilde{T}, U) \mathbf{1}_{\{\exists v \in T : A(\tilde{T}, v) \in \mathcal{P}\}} \cdot \mathbf{1}_{\{T=\mathbf{t}, U=u\}} \right].$$

For every (\mathbf{t}, u) with $|u| = k$, we have

$$\begin{aligned} & \mathbb{E} \left[F(\tilde{T}, U) \mathbf{1}_{\{\exists v \in T : A(\tilde{T}, v) \in \mathcal{P}\}} \cdot \mathbf{1}_{\{T=\mathbf{t}, U=u\}} \right] \\ &= \mathbb{P}(T = \mathbf{t}) \cdot \mathbb{E} \left[F(\tilde{T}, u) \cdot \mathbf{1}_{\{A(\tilde{T}, u) \in \mathcal{P}\}} \cdot \mathbb{P}(U = u \mid \tilde{T}) \mid T = \mathbf{t} \right] \\ &= \mathbb{P}(T = \mathbf{t}) \cdot \mathbb{E} \left[\frac{F(\tilde{T}, u) \cdot \mathbf{1}_{\{A(\tilde{T}, u) \in \mathcal{P}\}}}{\#\{v \in \mathbf{t} \mid A(\tilde{T}, v) \in \mathcal{P}\}} \mid T = \mathbf{t} \right] \\ &= \mathbb{P}((\mathbf{T}_k^*, \mathbf{U}_k^*) = (\mathbf{t}, u)) \cdot \mathbb{E} \left[\frac{F(\tilde{\mathbf{T}}_k^*, u) \cdot \mathbf{1}_{\{A(\tilde{\mathbf{T}}_k^*, u) \in \mathcal{P}\}}}{\#\{v \in \mathbf{t} \mid A(\tilde{\mathbf{T}}_k^*, v) \in \mathcal{P}\}} \mid \mathbf{T}_k^* = \mathbf{t} \right] \\ &= \mathbb{P}((\mathbf{T}_k^*, \mathbf{U}_k^*) = (\mathbf{t}, u)) \cdot \mathbb{E} \left[\frac{F(\tilde{\mathbf{T}}_k^*, u) \cdot \mathbf{1}_{\{A(\tilde{\mathbf{T}}_k^*, u) \in \mathcal{P}\}}}{\#\{v \in \mathbf{t} \mid A(\tilde{\mathbf{T}}_k^*, v) \in \mathcal{P}\}} \mid \mathbf{T}_k^* = \mathbf{t}, \mathbf{U}_k^* = u \right] \end{aligned}$$

$$= \mathbb{E} \left[\frac{F(\tilde{\mathbf{T}}_k^*, \mathbf{U}_k^*) \cdot \mathbf{1}_{\{A(\tilde{\mathbf{T}}_k^*, \mathbf{U}_k^*) \in \mathcal{P}\}} \cdot \mathbf{1}_{\{(\mathbf{T}_k^*, \mathbf{U}_k^*) = (\mathbf{t}, u)\}}}{\#\{v \in \mathbf{t} \mid A(\tilde{\mathbf{T}}_k^*, v) \in \mathcal{P}\}} \right].$$

The third equality uses (A.12) and the fact that conditionally on the tree $T = \mathbf{t}$ or $\mathbf{T}_k^* = \mathbf{t}$, the distribution of the marks on the tree is the same. The fourth equality uses the fact that the quantity in the conditional expectation does not depend on \mathbf{U}_k^* . The result is then obtained by summing over all heights $k \geq 0$ and all finite trees \mathbf{t} with a distinguished vertex u at height k . \square

Remark A.4. For our purposes, we will just use $E = \mathbb{R}_+$ but we want to emphasize that this result still holds true for other types of marks: the key point for the proof is that the law of the marks conditionally on the tree only depends on the degree of the corresponding vertex.

A.4 The contribution of vertices of small degree is small

The goal of this section is to prove Proposition 3.2. For that, we prove Proposition A.5 which implies the former directly.

Let $\alpha \in (1, 2)$ and $\gamma > \alpha - 1$. Let μ be a critical reproduction distribution in the domain of attraction of an α -stable law. We let $b_n = B_\mu(n)$ be the $\frac{1}{\alpha}$ -regularly varying function associated to μ by (A.1). For all $n \geq 1$ for which the conditioning is non-degenerate, we let T_n be a BGW tree with reproduction law μ , conditioned to have exactly n vertices.

This random tree is endowed with marks $(X_u)_{u \in T_n}$ such that conditionally on T_n , the marks are independent with distribution that only depends on the degree of the corresponding vertex $X_u \sim \pi_{d_{T_n}^+(u)}$, where the sequence $(\pi_k)_{k \geq 0}$ is a sequence of probability measures on \mathbb{R}_+ . We further require that for some real number $m > \frac{4\alpha}{(2\gamma+1-\alpha)}$, we have

$$\sup_{k \geq 0} \mathbb{E} \left[\left(\frac{X^{(k)}}{k^\gamma} \right)^m \right] < \infty,$$

where for every $k \geq 0$, the random variable $X^{(k)}$ has distribution π_k . We prove the following.

Proposition A.5. *In the setting described above, for any $\epsilon > 0$ we have*

$$\mathbb{P} \left(\sup_{v \in T_n} \left\{ \sum_{u \preceq v} X_u \mathbf{1}_{\{d_{T_n}^+(u) \leq \delta b_n\}} \right\} > \epsilon b_n^\gamma \right) \xrightarrow{\delta \rightarrow 0} 0,$$

uniformly in $n \geq 1$ such that T_n is well-defined.

Remark that this proposition directly implies Proposition 3.2 when applying it to marks distribution $(\pi_k)_{k \geq 0}$ that are respectively the laws of $(\text{diam}(B_k))_{k \geq 0}$. In order to prove this result, we first prove an intermediate lemma.

Lemma A.6. *For any small enough $\eta > 0$, $\epsilon > 0$ and all $\delta \in (0, \epsilon^{\frac{1}{\eta}})$, we have simultaneously for all $n \geq 1$,*

$$\mathbb{P} \left(\sup_{v \in T_n} \left\{ \sum_{u \preceq v} X_u \mathbf{1}_{\{\frac{\delta}{2} b_n < d_{T_n}^+(u) \leq \delta b_n\}} \right\} > \epsilon b_n^\gamma \right) \leq \delta^\beta,$$

where $\beta = (\gamma + \frac{1-\alpha}{2} - 5\eta)m - 2\alpha - 4\eta$, which is positive if η is chosen small enough.

Let us show how the result we want follows from this lemma.

Proof of Proposition A.5. Let us take $\eta > 0$ small enough such that Lemma A.6 holds. For any $\epsilon > 0$ small enough, and $0 < \delta < \epsilon^{\frac{1}{\eta}}$ we have

$$\sup_{v \in T_n} \left(\sum_{u \preceq v} X_u \mathbf{1}_{\{d_{T_n}^+(u) \leq \delta b_n\}} \right) \leq \sum_{i=0}^{\infty} \sup_{v \in T_n} \left\{ \sum_{u \preceq v} X_u \mathbf{1}_{\{\delta 2^{-i-1} b_n < d_{T_n}^+(u) \leq \delta 2^{-i} b_n\}} \right\}.$$

Write $\frac{\epsilon}{1-2^{-\eta}} = \sum_{i=0}^{\infty} \epsilon \cdot 2^{-i\eta}$ and use a union bound and the result of Lemma A.6 for all pairs $(\epsilon', \delta') \in \{(2^{-i\eta} \epsilon, 2^{-i} \delta), i \geq 0\}$ which thanks to our assumption, still satisfy $\delta' < (\epsilon')^{\frac{1}{\eta}}$. This yields

$$\begin{aligned} & \mathbb{P} \left(\sup_{v \in T_n} \left(\sum_{u \preceq v} X_u \mathbf{1}_{\{d_{T_n}^+(u) \leq \delta b_n\}} \right) > \frac{\epsilon}{1-2^{-\eta}} b_n^\gamma \right) \\ & \leq \sum_{i=0}^{\infty} \mathbb{P} \left(\sup_{v \in T_n} \left\{ \sum_{u \preceq v} X_u \mathbf{1}_{\{\delta 2^{-i-1} b_n < d_{T_n}^+(u) \leq \delta 2^{-i} b_n\}} \right\} > 2^{-i\eta} \cdot \epsilon \cdot b_n^\gamma \right) \\ & \leq \sum_{i=0}^{\infty} (2^{-i} \delta)^\beta \xrightarrow{\delta \rightarrow 0} 0 \end{aligned}$$

which is what we wanted to prove. □

The rest of the section is then devoted to the proof of Lemma A.6.

A.4.1 Proof of Lemma A.6

Now, let us turn to the proof of Lemma A.6. For this one, we are going to use Proposition A.3 with a certain property $\tilde{P}(\epsilon, \delta, n)$, which is defined as the subset of \mathbb{A} such that for any $\tilde{\mathbf{t}} = (\mathbf{t}, (x_u)_{u \in \mathbf{t}})$ and $v \in \mathbf{t}$,

$$A(\tilde{\mathbf{t}}, v) \in \tilde{P}(\epsilon, \delta, n) \iff \sum_{u \preceq v} x_u \mathbf{1}_{\{\frac{\delta}{2} b_n < d_{\tilde{\mathbf{t}}}^+(u) \leq \delta b_n\}} > \epsilon b_n^\gamma.$$

Let $T_n \sim \text{BGW}_\mu^n$. For a small $\eta > 0$ we can write using a union bound

$$\begin{aligned} & \mathbb{P} \left(\sup_{v \in T_n} \left\{ \sum_{u \preceq v} X_u \mathbf{1}_{\{\frac{\delta}{2} b_n < d_{T_n}^+(u) \leq \delta b_n\}} \right\} > \epsilon b_n^\gamma \right) \\ & \leq \mathbb{P} \left(\sup_{v \in T_n} \left\{ \sum_{u \preceq v} X_u \mathbf{1}_{\{\frac{\delta}{2} b_n < d_{T_n}^+(u) \leq \delta b_n\}} \right\} > \epsilon b_n^\gamma \text{ and } \text{ht}(T_n) \leq \delta^{-\eta} \frac{n}{b_n} \right) \\ & \quad + \mathbb{P} \left(\text{ht}(T_n) > \delta^{-\eta} \frac{n}{b_n} \right). \end{aligned} \tag{A.13}$$

Thanks to [24, Theorem 2], the second term is already smaller than some $C_1 \exp(-C_2 \delta^{-\eta})$, uniformly in n . We then just have to take care of the first term, for which we write

$$\begin{aligned} & \mathbb{P} \left(\sup_{v \in T_n} \left\{ \sum_{u \preceq v} X_u \mathbf{1}_{\{\frac{\delta}{2} b_n < d_{T_n}^+(u) \leq \delta b_n\}} \right\} > \epsilon b_n^\gamma \text{ and } \text{ht}(T_n) \leq \delta^{-\eta} \frac{n}{b_n} \right) \\ & = \frac{1}{\mathbb{P}(|T| = n)}. \end{aligned}$$

$$\mathbb{P} \left(|T| = n \text{ and } \exists v \in T : \sum_{u \preceq v} X_u \mathbf{1}_{\{\frac{\delta}{2} b_n < d_{T_n}^+(u) \leq \delta b_n\}} > \epsilon b_n^\gamma \text{ and } \text{ht}(T) \leq \delta^{-\eta} \frac{n}{b_n} \right), \tag{A.14}$$

where here T is an unconditioned BGW tree. We already know from (A.7) that

$$\mathbb{P}(|T| = n) \underset{n \rightarrow \infty}{\sim} \frac{d_\alpha(0)}{n b_n}.$$

Using the spine decomposition Proposition A.3, we can write the second factor as

$$\begin{aligned} & \mathbb{P} \left(|T| = n, \text{ and } \exists v \in T : \sum_{u \preceq v} X_u \mathbf{1}_{\{\frac{\delta}{2} b_n < d_{T_n}^+(u) \leq \delta b_n\}} > \epsilon b_n^\gamma \text{ and } \text{ht}(v) \leq \delta^{-\eta} \frac{n}{b_n} \right) \\ &= \mathbb{E} \left[\sum_{k=0}^{\delta^{-\eta} \frac{n}{b_n}} \frac{\mathbf{1}_{\{|\mathbf{T}_k^*|=n\}} \cdot \mathbf{1}_{\{A(\tilde{\mathbf{T}}_k^*, \mathbf{U}_k^*) \in \tilde{P}(\epsilon, \delta, n)\}}}{\#\{v \in \mathbf{T}_k^* \mid A(\tilde{\mathbf{T}}_k^*, v) \in \tilde{P}(\epsilon, \delta, n)\}} \right]. \end{aligned} \tag{A.15}$$

Then we are going to upper-bound the last display by using the fact that from the definition of $\tilde{P}(\epsilon, \delta, n)$, for any $v \in \mathbf{T}_k^*$ such that $A(\tilde{\mathbf{T}}_k^*, v) \in \tilde{P}(\epsilon, \delta, n)$, any descendant u of v satisfies $A(\tilde{\mathbf{T}}_k^*, u) \in \tilde{P}(\epsilon, \delta, n)$. We first introduce

$$N(\epsilon, \delta, n) := \inf \left\{ i \geq 0 \mid \sum_{u \preceq \mathbf{U}_i^*} X_u \mathbf{1}_{\{\frac{\delta}{2} b_n < d_{\mathbf{T}_k^*}^+(u) \leq \delta b_n\}} > \epsilon b_n^\gamma \right\},$$

which is the height of the first vertex on the spine that satisfies $\tilde{P}(\epsilon, \delta, n)$. This, in particular, entails that $\frac{\delta}{2} b_n \leq d_{\mathbf{T}_k^*}^+(\mathbf{U}_{N(\epsilon, \delta, n)}^*) \leq \delta b_n$. For technical reasons that will be explained later, let us split the offspring of $u = \mathbf{U}_{N(\epsilon, \delta, n)}^*$ into two subsets whose sizes are of the same order: those written as ui for $i \leq \frac{\delta}{4} b_n$ and those written as ui for $i > \frac{\delta}{4} b_n$. We only consider the progeny of the former for now and use that to lower-bound the total number $\#\{v \in \mathbf{T}_k^* \mid A(\tilde{\mathbf{T}}_k^*, v) \in \tilde{P}(\epsilon, \delta, n)\}$. We then get

$$\begin{aligned} & \mathbb{E} \left[\sum_{k=0}^{\delta^{-\eta} \frac{n}{b_n}} \frac{\mathbf{1}_{\{|\mathbf{T}_k^*|=n\}} \cdot \mathbf{1}_{\{A(\tilde{\mathbf{T}}_k^*, \mathbf{U}_k^*) \in \tilde{P}(\epsilon, \delta, n)\}}}{\#\{v \in \mathbf{T}_k^* \mid A(\tilde{\mathbf{T}}_k^*, v) \in \tilde{P}(\epsilon, \delta, n)\}} \right] \\ & \leq \sum_{k=0}^{\delta^{-\eta} \frac{n}{b_n}} \mathbb{E} \left[\frac{\mathbf{1}_{\{|\mathbf{T}_k^*|=n\}} \cdot \mathbf{1}_{\{A(\mathbf{T}_k^*, \mathbf{U}_k^*) \in \tilde{P}(\epsilon, \delta, n)\}}}{\#\{v \in \mathbf{T}_k^* \mid v \succeq \mathbf{U}_{N(\epsilon, \delta, n)}^* i \text{ for some } i \leq \frac{\delta}{4} b_n\}} \right]. \end{aligned} \tag{A.16}$$

Now let us state a lemma and conclude from there. We prove the lemma at the end of the section.

Lemma A.7. *For all n and $k \leq \delta^{-\eta} \frac{n}{b_n}$, and $\delta < \epsilon^{\frac{1}{\eta}}$ small enough, there exists a constant C such that*

(i)

$$\mathbb{P} \left(|\mathbf{T}_k^*| = n \mid \frac{\mathbf{1}_{\{A(\mathbf{T}_k^*, \mathbf{U}_k^*) \in \tilde{P}(\epsilon, \delta, n)\}}}{\#\{v \in \mathbf{T}_k^* \mid v \succeq \mathbf{U}_{N(\epsilon, \delta, n)}^* i \text{ for some } i \leq \frac{\delta}{4} b_n\}} \right) \leq \frac{C \cdot \delta^{-\alpha-\eta}}{n},$$

(ii)

$$\mathbb{P} \left(A(\tilde{\mathbf{T}}_k^*, \mathbf{U}_k^*) \in \tilde{P}(\epsilon, \delta, n) \right) \leq \delta^{(\gamma + \frac{1-\alpha}{2} - 5\eta)m},$$

(iii)

$$\mathbb{E} \left[\frac{1}{\#\left\{v \in \mathbf{T}_k^* \mid v \succeq \mathbf{U}_{N(\epsilon, \delta, n)}^* i \text{ for } i \leq \frac{\delta}{4} b_n\right\}} \mid A(\mathbf{T}_k^*, \mathbf{U}_k^*) \in \tilde{P}(\epsilon, \delta, n) \right] \leq \frac{C \delta^{-\alpha-\eta}}{n}.$$

Using the above lemma to take care of every term in the sum appearing on the right-hand-side of (A.16) and sum over k to get the following

$$\begin{aligned} \sum_{k=0}^{\delta^{-\eta} \frac{n}{b_n}} \mathbb{E} \left[\frac{\mathbf{1}_{\{|\mathbf{T}_k^*|=n\}} \cdot \mathbf{1}_{\{A(\mathbf{T}_k^*, \mathbf{U}_k^*) \in \tilde{P}(\epsilon, \delta, n)\}}}{\#\left\{v \in \mathbf{T}_k^* \mid v \succeq \mathbf{U}_{N(\epsilon, \delta, n)}^* i\right\}} \right] &\leq \delta^{-\eta} \frac{n}{b_n} \cdot C \frac{\delta^{-\alpha-\eta}}{n} \cdot C \frac{\delta^{-\alpha-\eta}}{n} \cdot \delta^{(\gamma+\frac{1-\alpha}{2}-5\eta)m} \\ &\leq C \cdot \frac{\delta^{(\gamma+\frac{1-\alpha}{2}-5\eta)m-2\alpha-3\eta}}{n b_n}. \end{aligned}$$

Plugging the last display into (A.14) using the equality (A.15) and using (A.13) and (A.7), we then get that for small enough δ , for $\delta < \epsilon^{\frac{1}{\eta}}$ we have

$$\mathbb{P} \left(\sup_{v \in T_n} \left\{ \sum_{u \leq v} X_u \mathbf{1}_{\{\frac{\delta}{2} b_n < d_{T_n}^+(u) \leq \delta b_n\}} \right\} > \epsilon b_n^\gamma \right) \leq \delta^{(\gamma+\frac{1-\alpha}{2}-5\eta)m-2\alpha-4\eta},$$

which is what we wanted to prove.

A.4.2 Proof of Lemma A.7

Let us successively prove the three points of the lemma. First, let us remark that if $\delta \leq \frac{\delta}{b_n}$, then (i) and (iii) are trivial. Indeed, in that case the quantities on the left-hand-side are smaller than 1. On the other side, we can prove using the Potter bounds, that if we choose the constant large enough, the right-hand-side of that inequality is always greater than 1 for δ in that range. Hence when proving (i) and (iii), we can always assume that $\delta \geq \frac{\delta}{b_n}$ so that $\frac{\delta}{4} b_n \geq 2$.

Proof of (i) Let us work on the event $\{A(\mathbf{T}_k^*, \mathbf{U}_k^*) \in \tilde{P}(\epsilon, \delta, n)\}$. We denote by v_1, \dots, v_J the vertices of the form $\mathbf{U}_{N(\epsilon, \delta, n)}^* i$ for $i > \frac{\delta}{4} b_n$ that are not $\mathbf{U}_{N(\epsilon, \delta, n)+1}^*$. There is some number J of them where $1 \leq \frac{\delta}{4} b_n - 1 \leq J \leq \delta b_n$. For this proof, we define $\text{Cut}(\mathbf{T}_k^*, v_1, v_2, \dots, v_J)$ similarly as in (A.10) except this time we remove all the vertices that are strictly above all every one of the vertices v_1, v_2, \dots, v_J .

Now, note that the knowledge of the tree $\text{Cut}(\mathbf{T}_k^*, v_1, v_2, \dots, v_J)$ is enough to compute the quantity $\#\left\{v \in \mathbf{T}_k^* \mid v \succeq \mathbf{U}_{N(\epsilon, \delta, n)}^* i \text{ for } i \leq \frac{\delta}{4} b_n\right\}$, and that conditionally on the tree $\text{Cut}(\mathbf{T}_k^*, v_1, v_2, \dots, v_J)$, the subtrees τ_1, \dots, τ_J respectively above v_1, \dots, v_J are i.i.d. BGW_μ -distributed random trees. Then the total size of the whole tree \mathbf{T}_k^* is exactly n if and only if the total volume of those trees τ_1, \dots, τ_J is exactly n minus the number of vertices in the rest of \mathbf{T}_k^* . Setting

$$\Xi_n = \frac{\mathbf{1}_{\{A(\mathbf{T}_k^*, \mathbf{U}_k^*) \in \tilde{P}(\epsilon, \delta, n)\}}}{\#\left\{v \in \mathbf{T}_k^* \mid v \succeq \mathbf{U}_{N(\epsilon, \delta, n)}^* i \text{ for some } i \leq \frac{\delta}{4} b_n\right\}},$$

this (conditional) probability is bounded above as follows

$$\mathbb{P}(|\mathbf{T}_k^*| = n \mid \Xi_n) = \mathbb{E}[\mathbb{P}(|\mathbf{T}_k^*| = n \mid \text{Cut}(\mathbf{T}_k^*, v_1, v_2, \dots, v_J)) \mid \Xi_n]$$

$$\begin{aligned}
 &= \mathbb{E} \left[\mathbb{P} \left(\sum_{j=1}^J |\tau_j| = n + J - |\mathbf{T}_k^*| \mid \text{Cut}(\mathbf{T}_k^*, v_1, v_2, \dots, v_J) \right) \mid \Xi_n \right] \\
 &\leq \frac{C \cdot \delta^{-\alpha-\eta}}{n},
 \end{aligned}$$

using Lemma A.1(ii) and the fact that by our assumptions we have $J \geq \frac{\delta}{4}b_n - 1 \geq 1$.

Proof of (ii) For any $k \geq 1$, let us write $(X_i)_{0 \leq i \leq k} = (X_{U_i^*})$ to simplify the notation. We can write, using that $\epsilon \geq \delta^\eta$,

$$\begin{aligned}
 &\mathbb{P} \left(A(\tilde{\mathbf{T}}_k^*, \mathbf{U}_k^*) \in \tilde{P}(\epsilon, \delta, n) \right) \\
 &\leq \mathbb{P} \left(\sum_{i=1}^k X_i \mathbf{1}_{\{\frac{\delta}{2}b_n < D_i \leq \delta b_n\}} > \epsilon b_n^\gamma \right) \\
 &\leq \mathbb{P} \left(\sum_{i=1}^k \frac{X_i}{(\delta b_n)^\gamma} \mathbf{1}_{\{\frac{\delta}{2}b_n < D_i \leq \delta b_n\}} > \delta^{-\gamma+\eta} \mid \sum_{i \leq k} \mathbf{1}_{\{\frac{\delta}{2}b_n < D_i \leq \delta b_n\}} \leq \delta^{1-\alpha-3\eta} \right) \\
 &\quad + \mathbb{P} \left(\sum_{i \leq k} \mathbf{1}_{\{\frac{\delta}{2}b_n < D_i \leq \delta b_n\}} > \delta^{1-\alpha-3\eta} \right). \tag{A.17}
 \end{aligned}$$

The second term of the last display is always smaller than what we would get if we take the maximum value for k , i.e. $\delta^{-\eta} \frac{n}{b_n}$. Using Lemma A.1(i) we have for all $n \geq 1$,

$$\mathbb{P} \left(\frac{\delta}{2}b_n \leq D_i \leq \delta b_n \right) \leq C \cdot \delta^{1-\alpha-\eta} \cdot \frac{b_n}{n}.$$

Hence

$$\mathbb{P} \left(\sum_{i \leq k} \mathbf{1}_{\{\frac{\delta}{2}b_n < D_i \leq \delta b_n\}} > \delta^{1-\alpha-3\eta} \right) \leq \mathbb{P} \left(\text{Bin} \left(\delta^{-\eta} \frac{n}{b_n}, 1 \wedge (C \cdot \delta^{1-\alpha-\eta} \cdot \frac{b_n}{n}) \right) > \delta^{1-\alpha-3\eta} \right),$$

which decays exponentially in a negative power of δ . For the first term of (A.17), we use the fact that conditionally on the event $\{\sum_{i \leq k} \mathbf{1}_{\{\frac{\delta}{2}b_n < D_i \leq \delta b_n\}} \leq \delta^{1-\alpha-3\eta}\}$, all the random variables $\frac{X_i}{(\delta b_n)^\gamma}$ are independent with m -th moment bounded above uniformly by the same constant C ,

$$\sup_{0 \leq i \leq k} \mathbb{E} \left[\left(\frac{X_i}{(\delta b_n)^\gamma} \right)^m \right] < C.$$

Note that we have (for another constant C)

$$\sum_{i=1}^k \frac{\mathbb{E}[X_i]}{(\delta b_n)^\gamma} \mathbf{1}_{\{\frac{\delta}{2}b_n < D_i \leq \delta b_n\}} \leq C \cdot \sum_{i \leq k} \mathbf{1}_{\{\frac{\delta}{2}b_n < D_i \leq \delta b_n\}}. \tag{A.18}$$

We use Markov's inequality and a result of Petrov [30, Chapter III. Result 5.16], which applies thanks to the fact that $m \geq 2$. In the third line, we will use that $\gamma > \alpha - 1$ so that for δ small enough we have $\delta^{-\gamma+\eta} - C \cdot \delta^{1-\alpha-3\eta} \geq \delta^{-\gamma+2\eta}$. Defining the event

$$\mathcal{E}_n = \left\{ \sum_{i \leq k} \mathbf{1}_{\{\frac{\delta}{2}b_n < D_i \leq \delta b_n\}} \leq \delta^{1-\alpha-3\eta} \right\},$$

we have

$$\begin{aligned}
 & \mathbb{P} \left(\sum_{i=1}^k \frac{X_i}{(\delta b_n)^\gamma} \mathbf{1}_{\{\frac{\delta}{2} b_n < D_i \leq \delta b_n\}} > \delta^{-\gamma+\eta} \mid \mathcal{E}_n \right) \\
 & \stackrel{(A.18)}{\leq} \mathbb{P} \left(\sum_{i=1}^k \frac{X_i - \mathbb{E}[X_i]}{(\delta b_n)^\gamma} \mathbf{1}_{\{\frac{\delta}{2} b_n < D_i \leq \delta b_n\}} > \delta^{-\gamma+\eta} - C\delta^{1-\alpha-3\eta} \mid \mathcal{E}_n \right) \\
 & \stackrel{\text{Markov}}{\leq} \frac{\mathbb{E} \left[\left(\sum_{i=1}^k \frac{X_i - \mathbb{E}[X_i]}{(\delta b_n)^\gamma} \mathbf{1}_{\{\frac{\delta}{2} b_n < D_i \leq \delta b_n\}} \right)^m \mid \mathcal{E}_n \right]}{(\delta^{-\gamma+2\eta})^m} \\
 & \stackrel{\text{Petrov}}{\leq} C \cdot \frac{(\delta^{1-\alpha-3\eta})^{\frac{m}{2}}}{(\delta^{-\gamma+2\eta})^m} \\
 & \leq \delta^{(\gamma+\frac{1-\alpha}{2}-4\eta)m}.
 \end{aligned}$$

Taking the sum of the two terms in (A.17) ensures that for ϵ small enough and $\delta < \epsilon^{1/\eta}$,

$$\mathbb{P} \left(A(\tilde{\mathbf{T}}_k^*, \mathbf{U}_k^*) \in \tilde{P}(\epsilon, \delta, n) \right) \leq \delta^{(\gamma+\frac{1-\alpha}{2}-5\eta)m}.$$

Proof of (iii) Now let us reason conditionally on the event $\{A(\tilde{\mathbf{T}}_k^*, \mathbf{U}_k^*) \in \tilde{P}(\epsilon, \delta, n)\}$. On that event the quantity $\#\left\{v \in \mathbf{T}_k^* \mid v \succeq \mathbf{U}_{N(\epsilon, \delta, n)}^* i \text{ for } i \leq \frac{\delta}{4} b_n\right\}$ is at least the sum of the total size of $\frac{\delta}{4} b_n - 1$ independent BGW tree branching off of the spine. Hence

$$\begin{aligned}
 & \mathbb{E} \left[\frac{1}{\#\left\{v \in \mathbf{T}_k^* \mid v \succeq \mathbf{U}_{N(\epsilon, \delta, n)}^* i \text{ for } i \leq \frac{\delta}{4} b_n\right\}} \mid A(\mathbf{T}_k^*, \mathbf{U}_k^*) \in \tilde{P}(\epsilon, \delta, n) \right] \\
 & \leq \mathbb{E} \left[\frac{1}{\sum_{i=1}^{\frac{\delta}{4} b_n - 1} |\tau_i|} \right] \leq \frac{C\delta^{-\alpha-\eta}}{n},
 \end{aligned}$$

where the τ_i 's are i.i.d. under distribution BGW_μ . The last inequality is obtained using Lemma A.1(iii).

A.5 Proof of Lemma 3.4

Before going into the proof of Lemma 3.4, let us recall some general arguments concerning the Łukasiewicz path of BGW trees conditional on their total size. For T_n a BGW tree conditioned on having total size n , the law of the vector $(d_{T_n}^+(v_1) - 1, \dots, d_{T_n}^+(v_n) - 1; \nu_{v_1}(B_{v_1}), \dots, \nu_{v_n}(B_{v_n}))$ can be described as

$$\text{Law} \left((X_1, \dots, X_n; Z_1, \dots, Z_n) \mid \sum_{i=1}^n X_i = -1, \forall k \leq n-1, \sum_{i=1}^k X_i \geq 0 \right),$$

where the $(X_i, Z_i)_{1 \leq i \leq n}$ are i.i.d. with the distribution of $(D-1, \nu_D(B_D))$ where D follows the reproduction distribution.

Now using the so-called Vervaat transform, the above law can also be expressed as

$$\text{Law} \left((X_{U+1}, \dots, X_{U+n}; Z_{U+1}, \dots, Z_{U+n}) \mid \sum_{i=1}^n X_i = -1 \right),$$

where U is defined as $U := \min \left\{ 1 \leq k \leq n \mid \sum_{i=1}^k X_i = \min_{1 \leq k \leq n} \sum_{i=1}^k X_i \right\}$, and the indices in the last display are interpreted modulo n .

In what follows, we will also use an argument of absolute continuity. For any bounded function F we can write

$$\begin{aligned} & \mathbb{E} \left[F((X_1, \dots, X_{\lfloor \frac{3n}{4} \rfloor}; Z_1, \dots, Z_{\lfloor \frac{3n}{4} \rfloor})) \middle| \sum_{i=1}^n X_i = -1 \right] \\ &= \mathbb{E} \left[F((X_1, \dots, X_{\lfloor \frac{3n}{4} \rfloor}; Z_1, \dots, Z_{\lfloor \frac{3n}{4} \rfloor})) \frac{\mathbb{P} \left(\sum_{i=\lfloor \frac{3n}{4} \rfloor+1}^n X_i = -1 - \sum_{i=1}^{\lfloor \frac{3n}{4} \rfloor} X_i \middle| \sum_{i=1}^{\lfloor \frac{3n}{4} \rfloor} X_i \right)}{\mathbb{P} \left(\sum_{i=1}^n X_i = -1 \right)} \right] \end{aligned}$$

Using the local limit theorem [20, Theorem 4.2.1] and the fact that the α -stable density is bounded [41, Eq.(I20)], the term

$$\frac{\mathbb{P} \left(\sum_{i=\lfloor \frac{3n}{4} \rfloor+1}^n X_i = -1 - \sum_{i=1}^{\lfloor \frac{3n}{4} \rfloor} X_i \middle| \sum_{i=1}^{\lfloor \frac{3n}{4} \rfloor} X_i \right)}{\mathbb{P} \left(\sum_{i=1}^n X_i = -1 \right)}$$

appearing in the integral is bounded uniformly in $n \geq 4$. Using indicator functions for F , this ensures that any event that occurs with probability tending to 0 or 1 as $n \rightarrow \infty$ for the unconditioned model, also tend to 0 or 1 for the conditioned model. Note that we can apply the same argument for the vector $(X_{\lfloor \frac{n}{4} \rfloor}, \dots, X_n; Z_{\lfloor \frac{n}{4} \rfloor}, \dots, Z_n)$.

Proof of Lemma 3.4. We consider the case $\beta \leq \alpha$ and the case $\beta > \alpha$ in turn.

- Case $\beta \leq \alpha$. In this case, we want to show that there exists a sequence $(a_n)_{n \geq 1}$ so that $\frac{1}{a_n} \cdot (M_{\lfloor nt \rfloor})_{0 \leq t \leq 1} \xrightarrow[n \rightarrow \infty]{} (t)_{0 \leq t \leq 1}$ in probability for the uniform topology. First, let us show that it is enough to check this if we replace $(M_k)_{1 \leq k \leq n}$ with $(S_k)_{1 \leq k \leq n}$ where $S_k := \sum_{i=1}^k Z_i$. Indeed, if

$$\frac{1}{a_n} \cdot (S_{\lfloor nt \rfloor})_{0 \leq t \leq 1} \xrightarrow[n \rightarrow \infty]{} (t)_{0 \leq t \leq 1} \tag{A.19}$$

in probability, then we have

$$\frac{1}{a_n} \cdot (S_{\lfloor nt \rfloor})_{0 \leq t \leq \frac{3}{4}} \xrightarrow[n \rightarrow \infty]{} (t)_{0 \leq t \leq \frac{3}{4}} \quad \text{and} \quad \frac{1}{a_n} \cdot (S_{\lfloor nt \rfloor} - S_{\lfloor \frac{3}{4}n \rfloor})_{\frac{1}{4} \leq t \leq 1} \xrightarrow[n \rightarrow \infty]{} (t - \frac{1}{4})_{\frac{1}{4} \leq t \leq 1}$$

also in probability. By the absolute continuity argument, both those convergences also hold conditional on the event $\{\sum_{i=1}^n X_i = -1\}$, which ensures that $\frac{1}{a_n} \cdot (S_{\lfloor nt \rfloor})_{0 \leq t \leq 1} \xrightarrow[n \rightarrow \infty]{} (t)_{0 \leq t \leq 1}$ under $\mathbb{P}(\cdot \mid \sum_{i=1}^n X_i = -1)$. Then, this is enough to get that $\frac{1}{a_n} \cdot (S_{U+\lfloor nt \rfloor})_{0 \leq t \leq 1} \xrightarrow[n \rightarrow \infty]{} (t)_{0 \leq t \leq 1}$ as well. Since $(M_k)_{1 \leq k \leq n}$ is distributed as $S_{U+\lfloor nt \rfloor}$ under $\mathbb{P}(\cdot \mid \sum_{i=1}^n X_i = -1)$, we get $\frac{1}{a_n} \cdot (M_{\lfloor nt \rfloor})_{0 \leq t \leq 1} \xrightarrow[n \rightarrow \infty]{} (t)_{0 \leq t \leq 1}$ in probability as required.

In the end, we just have to find a sequence a_n such that (A.19) holds. In fact, by monotonicity, we just need to check that the convergence holds for $t = 1$. If $\mathbb{E}[Z] = \mathbb{E}[\nu_D(B_D)]$ is finite then it is easy to check that we can take $a_n := \mathbb{E}[\nu_D(B_D)] \cdot n$. This is always the case if $\beta < \alpha$ since, using the assumption of the lemma

$$\begin{aligned} \mathbb{E}[\nu_D(B_D)] &= \sum_{k=0}^{\infty} \mathbb{P}(D = k) \mathbb{E}[\nu_k(B_k)] \\ &\stackrel{\text{assum.}}{\leq} C \cdot \sum_{k=0}^{\infty} \mathbb{P}(D = k) k^\beta \leq C \cdot \mathbb{E}[D^\beta] < \infty, \end{aligned}$$

so we can apply the law of large numbers.

If $\mathbb{E}[Z]$ is infinite, which in our case can only happen if $\beta = \alpha$, we need to understand the tail behaviour of Z . We will show in that case that

$$\mathbb{P}(Z \geq x) \sim \mathbb{P}\left(D \geq x^{\frac{1}{\alpha}}\right) \cdot \mathbb{E}[\nu(B)] \quad \text{as } x \rightarrow \infty, \tag{A.20}$$

which in particular entails that Z has a regularly varying tail of index -1 as $x \rightarrow \infty$. Using then general results (e.g. [15, Theorem 3.7.2]) for sums of random variables with regularly varying tail yields the result.

First, for a given integer N we introduce Y_N^+ and Y_N^- whose distribution are defined in such a way that for all $x \geq 0$,

$$\mathbb{P}(Y_N^+ \geq x) = \sup_{k \geq N} \mathbb{P}\left(\frac{\nu_k(B_k)}{k^\alpha} \geq x\right) \quad \text{and} \quad \mathbb{P}(Y_N^- \geq x) = \inf_{k \geq N} \mathbb{P}\left(\frac{\nu_k(B_k)}{k^\alpha} \geq x\right).$$

Note that for both of them, we have the following bound thanks to our assumption and Markov's inequality

$$\mathbb{P}(Y_N^- \geq x) \leq \mathbb{P}(Y_N^+ \geq x) \leq \sup_{k \geq 1} \mathbb{P}\left(\frac{\nu_k(B_k)}{k^\alpha} \geq x\right) \leq \frac{\sup_{k \geq 1} \mathbb{E}\left[\left(\frac{\nu_k(B_k)}{k^\alpha}\right)^{1+\eta}\right]}{x^{1+\eta}}.$$

Since the quantity in the last display is integrable in x , and since $\mathbb{P}(Y_N^\pm \geq x) \rightarrow \mathbb{P}(\nu(\mathcal{B}) \geq x)$ as $N \rightarrow \infty$ thanks to assumption D1, we have $\mathbb{E}[Y_N^\pm] \rightarrow \mathbb{E}[\nu(\mathcal{B})]$ by dominated convergence. Also note that we have $\mathbb{E}\left[(Y_N^\pm)^{1+\eta/2}\right] < \infty$. In order to prove (A.20), we then just need to show that for every N , we have

$$\mathbb{P}(Z \geq x) \geq \mathbb{P}\left(D \geq x^{\frac{1}{\alpha}}\right) \cdot \mathbb{E}[Y_N^-] (1 + o(1))$$

and

$$\mathbb{P}(Z \geq x) \leq \mathbb{P}\left(D \geq x^{\frac{1}{\alpha}}\right) \cdot \mathbb{E}[Y_N^+] (1 + o(1)).$$

Let us fix N and write

$$\begin{aligned} \mathbb{P}(Z \geq x) &= \mathbb{P}(\nu_D(B_D) \geq x) = \mathbb{P}(\nu_D(B_D) \geq x, D \geq N) + \mathbb{P}(\nu_D(B_D) \geq x, D < N) \\ &= \mathbb{P}(\nu_D(B_D) \geq x, D \geq N) + O(x^{-1-\eta}). \end{aligned}$$

Now with a random variable Y_N^+ that is independent of D , we can write

$$\begin{aligned} \mathbb{P}(\nu_D(B_D) \geq x, D \geq N) &= \mathbb{E}\left[\mathbb{P}\left(\frac{\nu_D(B_D)}{D^\alpha} \geq \frac{x}{D^\alpha} \mid D\right) \mathbf{1}_{\{D \geq N\}}\right] \\ &\leq \mathbb{E}\left[\mathbb{P}\left(Y_N^+ \geq \frac{x}{D^\alpha}\right) \mathbf{1}_{\{D \geq N\}}\right] \\ &= \mathbb{P}(Y_N^+ D^\alpha \geq x, D > N) \\ &= \mathbb{P}(Y_N^+ D^\alpha \geq x) - \mathbb{P}(Y_N^+ D^\alpha \geq x, D \leq N) \\ &= \mathbb{P}(Y_N^+ D^\alpha \geq x) - O(x^{-1-\eta}). \end{aligned}$$

Now using [22, Proposition 1.1] this last display is equivalent to $\mathbb{P}\left(D \geq x^{\frac{1}{\alpha}}\right) \cdot \mathbb{E}[Y_N^+]$.

We also get the other inequality $\mathbb{P}(\nu_D(B_D) \geq x) \geq \mathbb{P}\left(D \geq x^{\frac{1}{\alpha}}\right) \cdot \mathbb{E}[Y_N^-] \cdot (1 + o(1))$, which finishes the proof.

• Case $\beta > \alpha$. Since the arguments used here are quite similar to what we did above, let us just sketch the proof. Let $\epsilon > 0$ and consider

$$R(n, \delta) := b_n^{-\beta} \cdot \sum_{i=1}^n \nu_{v_i}(B_{v_i}) \mathbf{1}_{\{d^+(v_i) \leq \delta b_n\}}.$$

Using the same type of arguments as above, we can first show that

$$b_n^{-\beta} \cdot \sum_{i=1}^{\lfloor \frac{3n}{4} \rfloor} Z_i \mathbf{1}_{\{X_{i+1} \leq \delta b_n\}} \quad \text{and} \quad b_n^{-\beta} \cdot \sum_{i=\lfloor \frac{n}{4} \rfloor}^n Z_i \mathbf{1}_{\{X_{i+1} \leq \delta b_n\}}$$

tend to 0 in probability under the unconditioned measure, uniformly in $n \geq 4$. Then using the absolute continuity argument together with the equality in distribution we get that

$$\limsup_{n \rightarrow \infty} \mathbb{P}(R(n, \delta) \geq \epsilon) \xrightarrow{\delta \rightarrow 0} 0, \tag{A.21}$$

This ensures that condition B3 is satisfied. □

References

- [1] Marie Albenque, Nina Holden, and Xin Sun, *Scaling limit of triangulations of polygons*, Electron. J. Probab. **25** (2020), 43 (English), Id/No 135. MR4171388
- [2] David Aldous, *The continuum random tree. III*, Ann. Probab. **21** (1993), no. 1, 248–289. MR1207226
- [3] E. Archer, *Random walks on decorated Galton-Watson trees*, arXiv:2011.07266 (2020).
- [4] Erich Baur, Grégory Miermont, and Gourab Ray, *Classification of scaling limits of uniform quadrangulations with a boundary*, Ann. Probab. **47** (2019), no. 6, 3397–3477 (English). MR4038036
- [5] Jean Bertoin, *Lévy processes*, Cambridge Tracts in Mathematics, vol. 121, Cambridge University Press, Cambridge, 1996. MR1406564
- [6] Jérémie Bettinelli, *Scaling limit of random planar quadrangulations with a boundary*, Ann. Inst. Henri Poincaré Probab. Stat. **51** (2015), no. 2, 432–477. MR3335010
- [7] N. H. Bingham, C. M. Goldie, and J. L. Teugels, *Regular variation*, Encyclopedia of Mathematics and its Applications, vol. 27, Cambridge University Press, Cambridge, 1989. MR1015093
- [8] Alessandra Caraceni, *The scaling limit of random outerplanar maps*, Ann. Inst. Henri Poincaré, Probab. Stat. **52** (2016), no. 4, 1667–1686 (English). MR3573291
- [9] Nicolas Curien, Bénédicte Haas, and Igor Kortchemski, *The CRT is the scaling limit of random dissections*, Random Structures Algorithms **47** (2015), no. 2, 304–327. MR3382675
- [10] Nicolas Curien and Igor Kortchemski, *Random stable looptrees*, Electron. J. Probab. **19** (2014), no. 108, 35. MR3286462
- [11] Nicolas Curien and Igor Kortchemski, *Percolation on random triangulations and stable looptrees*, Probab. Theory Related Fields **163** (2015), no. 1-2, 303–337. MR3405619
- [12] Thomas Duquesne, *A limit theorem for the contour process of conditioned Galton-Watson trees*, Ann. Probab. **31** (2003), no. 2, 996–1027. MR1964956
- [13] Thomas Duquesne and Jean-François Le Gall, *Random trees, Lévy processes and spatial branching processes*, Astérisque (2002), no. 281, vi+147. MR1954248
- [14] Thomas Duquesne and Jean-François Le Gall, *Probabilistic and fractal aspects of Lévy trees*, Probab. Theory Related Fields **131** (2005), no. 4, 553–603. MR2147221
- [15] Rick Durrett, *Probability: theory and examples*, fourth ed., Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge University Press, Cambridge, 2010. MR2722836
- [16] Jean-Francois Le Gall and Yves Le Jan, *Branching processes in Lévy processes: the exploration process*, The Annals of Probability **26** (1998), no. 1, 213–252. MR1617047
- [17] Christina Goldschmidt and Bénédicte Haas, *A line-breaking construction of the stable trees*, Electron. J. Probab. **20** (2015), no. 16, 24. MR3317158
- [18] Christina Goldschmidt, Bénédicte Haas, and Delphin Sénizergues, *Stable graphs: distributions and line-breaking construction*, arXiv:1811.06940 (2018). MR4526241
- [19] Ewain Gwynne and Jason Miller, *Convergence of the free Boltzmann quadrangulation with simple boundary to the Brownian disk*, Ann. Inst. Henri Poincaré, Probab. Stat. **55** (2019), no. 1, 551–589 (English). MR3901655

- [20] I. A. Ibragimov and Yu. V. Linnik, *Independent and stationary sequences of random variables*, Wolters-Noordhoff Publishing, Groningen, 1971, With a supplementary chapter by I. A. Ibragimov and V. V. Petrov, Translation from the Russian edited by J. F. C. Kingman. MR0322926
- [21] Svante Janson, *Simply generated trees, conditioned Galton-Watson trees, random allocations and condensation*, Probab. Surv. **9** (2012), 103–252. MR2908619
- [22] Yuji Kasahara, *A note on the product of independent random variables with regularly varying tails*, Tsukuba Journal of Mathematics **42** (2018), no. 2, 295–308. MR3934992
- [23] Igor Kortchemski, *Invariance principles for Galton-Watson trees conditioned on the number of leaves*, Stochastic Process. Appl. **122** (2012), no. 9, 3126–3172. MR2946438
- [24] Igor Kortchemski, *Sub-exponential tail bounds for conditioned stable Bienaymé-Galton-Watson trees*, Probab. Theory Related Fields **168** (2017), no. 1-2, 1–40. MR3651047
- [25] Igor Kortchemski and Cyril Marzouk, *Triangulating stable laminations*, Electron. J. Probab. **21** (2016), 31 pp. MR3485353
- [26] Igor Kortchemski and Loïc Richier, *The boundary of random planar maps via looptrees*, Ann. Fac. Sci. Toulouse, Math. (6) **29** (2020), no. 2, 391–430 (English). MR4150547
- [27] Jean-François Le Gall and Grégory Miermont, *Scaling limits of random trees and planar maps*, Probability and statistical physics in two and more dimensions. Proceedings of the Clay Mathematics Institute summer school and XIV Brazilian school of probability, Búzios, Brazil, July 11–August 7, 2010., Providence, RI: American Mathematical Society (AMS); Cambridge, MA: Clay Mathematics Institute, 2012, pp. 155–211 (English). MR3025391
- [28] Cyril Marzouk, *On the sizes of burnt and fireproof components for fires on a large Cayley tree*, Ann. Inst. Henri Poincaré, Probab. Stat. **52** (2016), no. 1, 355–375 (English). MR3449306
- [29] Cyril Marzouk, *On scaling limits of planar maps with stable face-degrees*, ALEA, Lat. Am. J. Probab. Math. Stat. **15** (2018), no. 2, 1089–1123 (English). MR3852246
- [30] V. V. Petrov, *Sums of independent random variables*, Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 82, Springer-Verlag, New York-Heidelberg, 1975, Translated from the Russian by A. A. Brown. MR0388499
- [31] Loïc Richier, *Limits of the boundary of random planar maps*, Probability Theory and Related Fields (2017). MR3877547
- [32] Delphin Sénizergues, *Random gluing of metric spaces*, The Annals of Probability **47** (2019), no. 6, 3812–3865. MR4038043
- [33] Delphin Sénizergues, *Geometry of weighted recursive and affine preferential attachment trees*, Electron. J. Probab. **26** (2021), Paper No. 80, 56. MR4269210
- [34] Delphin Sénizergues, *Growing random graphs with a preferential attachment structure*, ALEA, Lat. Am. J. Probab. Math. Stat. **19** (2022), no. 11, 259–309. MR4359793
- [35] Sigurdur Örn Stefánsson and Benedikt Stufler, *Geometry of large Boltzmann outerplanar maps*, Random Struct. Algorithms **55** (2019), no. 3, 742–771 (English). MR3997486
- [36] Benedikt Stufler, *Scaling limits of random outerplanar maps with independent link-weights*, Ann. Inst. Henri Poincaré, Probab. Stat. **53** (2017), no. 2, 900–915 (English). MR3634279
- [37] Benedikt Stufler, *Random enriched trees with applications to random graphs*, Electronic Journal of Combinatorics **25** (2018), no. 3. MR3853863
- [38] Benedikt Stufler, *Limits of random tree-like discrete structures*, Probab. Surv. **17** (2020), 318–477 (English). MR4132643
- [39] Cédric Villani, *Optimal transport, old and new*, Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 338, Springer-Verlag, Berlin, 2009. MR2459454
- [40] Ward Whitt, *Stochastic-process limits*, Springer Series in Operations Research, Springer-Verlag, New York, 2002, An introduction to stochastic-process limits and their application to queues. MR1876437
- [41] V. M. Zolotarev, *One-dimensional stable distributions*, Translations of Mathematical Monographs, vol. 65, American Mathematical Society, Providence, RI, 1986, Translated from the Russian by H. H. McFaden, Translation edited by Ben Silver. MR854867

Acknowledgments. We warmly thank the referee for the thorough reading of the manuscript and the helpful comments. We are grateful to Nicolas Curien for suggesting this collaboration, and to Eleanor Archer for discussing some earlier version of her project [3].

Electronic Journal of Probability

Electronic Communications in Probability

Advantages of publishing in EJP-ECP

- Very high standards
- Free for authors, free for readers
- Quick publication (no backlog)
- Secure publication (LOCKSS¹)
- Easy interface (EJMS²)

Economical model of EJP-ECP

- Non profit, sponsored by IMS³, BS⁴, ProjectEuclid⁵
- Purely electronic

Help keep the journal free and vigorous

- Donate to the IMS open access fund⁶ (click here to donate!)
- Submit your best articles to EJP-ECP
- Choose EJP-ECP over for-profit journals

¹LOCKSS: Lots of Copies Keep Stuff Safe <http://www.lockss.org/>

²EJMS: Electronic Journal Management System: <https://vtex.lt/services/ejms-peer-review/>

³IMS: Institute of Mathematical Statistics <http://www.imstat.org/>

⁴BS: Bernoulli Society <http://www.bernoulli-society.org/>

⁵Project Euclid: <https://projecteuclid.org/>

⁶IMS Open Access Fund: <https://imstat.org/shop/donation/>