



Simply Generated Unrooted Plane Trees

Leon Ramzews and Benedikt Stufler

Mathematisches Institut
Ludwig-Maximilians-Universität München
Theresienstraße 39,
80333 München, Germany.
E-mail address: ramzews@math.lmu.de
URL: <http://www.mathematik.uni-muenchen.de/~ramzews/>

Institut für Mathematik
Universität Zürich,
Winterthurerstrasse 190,
CH-8057 Zürich, Switzerland.
E-mail address: benedikt.stufler@math.uzh.ch
URL: <http://user.math.uzh.ch/stufler/>

Abstract. We study random unrooted plane trees with n vertices sampled according to the weights corresponding to the vertex-degrees. Our main result shows that if the generating series of the weights has positive radius of convergence, then this model of random trees may be approximated geometrically by a Galton–Watson tree conditioned on having a large random size. This implies that a variety of results for the well-studied planted case also hold for unrooted trees, including Gromov–Hausdorff–Prokhorov scaling limits, tail-bounds for the diameter, distributional graph limits, and limits for the maximum degree. Our work complements results by Wang (2016), who studied random unrooted plane trees whose diameter tends to infinity.

1. Introduction

The classical model of trees studied in probability theory is that of a *planted plane tree*. Such a tree has a distinguished vertex called the root, and each vertex may have a linearly ordered list of children. Surely the most prominent example for a model of random plane trees is that of a Galton–Watson tree. Of particular interest is the asymptotic behaviour when conditioning this model on producing a tree with a specific number of vertices, leaves, or more generally vertices with

Received by the editors September 13th, 2018; accepted February 8th, 2019.

2010 *Mathematics Subject Classification.* 05C05, 05C80, 60J80.

Key words and phrases. random trees, unrooted plane trees.

The second author is supported by the Swiss National Science Foundation grant number 200020_172515.

outdegree in a fixed set, or on having a specific height, and letting this parameter tend to infinity, cf. [Abraham and Delmas \(2014a,b\)](#); [Aldous \(1993\)](#); [Janson \(2012\)](#). In a certain sense, these models may be termed *simply generated*, as they fall under the more general setting of fixing a weight-sequence, assigning to each tree the product of weights corresponding to its vertex out-degrees, and sampling a tree with probability proportional to this weight from some set of plane trees.

Apart from planted plane trees, many other types of combinatorial trees have been studied from both an enumerative and probabilistic viewpoint. Such trees may be rooted or unrooted, ordered or unordered, and labelled or unlabelled. [Drmotá \(2009\)](#) gives an extensive account on the subject and since then further additions to the field have been made by [Marckert and Miermont \(2011\)](#); [Haas and Miermont \(2012\)](#); [Panagiotou and Stufler \(2018\)](#); [Stufler \(2014, 2018b\)](#); [Wang \(2016\)](#).

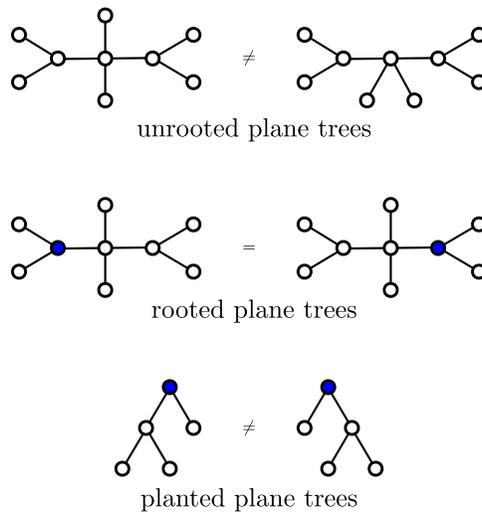


FIGURE 1.1. Different types of plane trees.

The present paper is concerned with simply generated unrooted plane trees as their number of vertices becomes large. In the language of planar maps, whose study has received considerable attention in recent literature, e.g. [Le Gall and Miermont \(2012\)](#), an *unrooted plane tree* is an (unrooted) planar map with only a single face. In combinatorial terms, we may describe it as an unlabelled tree where each vertex is endowed with a cyclic ordering of its neighbourhood. By distinguishing a vertex, we may form a *rooted plane tree*. Note that rooted plane trees differ from planted plane trees. In rooted plane trees, any vertex is endowed with a cyclic ordering of its neighbourhood, but in a planted plane tree it is a linear ordering on its offspring set. Figure 1.1 illustrates these subtle differences.

Our main result shows under minimal assumptions on the weights how we may geometrically approximate simply generated unrooted and rooted plane trees by the well studied model of simply generated planted plane trees. In fact, the only requirement we are going to make is that the generating series of the weights has positive radius of convergence. This level of generality is also one of the main difficulties in the proof. We comment on the non-analytic case in Remark 3.3 below.

The approximation is very accurate, so that *practically all* results available for Galton–Watson trees conditioned on having a fixed size carry over to simply generated unrooted plane trees. This includes a variety of graph limits such as Gromov–Hausdorff–Prokhorov scaling limits and Benjamini–Schramm limits, but also tail bounds for the diameter, a central limit theorem for the maximum degree, and results for other graph theoretic parameters.

A model of random unrooted plane trees has previously been studied by Wang (2016), who established powerful limit theorems as the diameter of these trees tends to infinity. There are some similarities in the approach of the present work and that of Wang (2016), although the focus in the latter is on random trees with a fixed diameter. In particular, Wang (2016, Sec. 5) decomposes such trees from their center and shows that typically there are no non-trivial central symmetries. This allows for an approximation via a pair of Galton–Watson trees conditioned to have specific heights. Uniform random unlabelled unrooted unordered trees were studied in Stuffer (2014) via the cycle-pointing technique, which we are not going to use in the present work. It could be applied to uniform random unrooted plane trees, but we aim for a much higher level of generality.

Outline of the paper

In Section 2 we present our main theorem. In Section 3 we describe the proof strategy and state our main lemmas. In Section 4 we state some applications of the main theorem. In particular, Subsection 4.1 provides scaling limits and tail-bounds for the diameter. Subsection 4.2 establishes a Benjamini–Schramm limit and a central limit theorem for the degree of a random vertex. Subsection 4.3 discusses a central limit theorem for the maximum degree and applications to other graph parameters. In Section 5 we recall the required combinatorial background to prepare for the proof of the geometric approximation. In Section 6 we present the proof of our main result. The proofs of our applications are collected in Section 7.

Notation

We let \mathbb{N} denote the set of positive integers and set $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$. The sets of positive and non-negative real numbers are denoted by $\mathbb{R}_{>0}$ and $\mathbb{R}_{\geq 0}$. Throughout, we assume that all considered random variables are defined on a common probability space. The *total variation distance* between two random variables X and Y with values in a countable state space S is defined by

$$d_{\text{TV}}(X, Y) = \sup_{\mathcal{E} \subset S} |\mathbb{P}(X \in \mathcal{E}) - \mathbb{P}(Y \in \mathcal{E})|.$$

A sequence of real-valued random variables $(X_n)_{n \geq 1}$ is *stochastically bounded*, if for each $\epsilon > 0$ there is a constant $M > 0$ with

$$\limsup_{n \rightarrow \infty} \mathbb{P}(|X_n| \geq M) \leq \epsilon.$$

We denote this by $X_n = O_p(1)$. Likewise, we write $X_n = o_p(1)$ if the sequence converges to 0 in probability. We use \xrightarrow{d} and \xrightarrow{p} to denote convergence in distribution and probability. A function

$$h : \mathbb{R}_{>0} \rightarrow \mathbb{R}_{>0}$$

is termed *slowly varying*, if for any fixed $t > 0$ it holds that

$$\lim_{x \rightarrow \infty} \frac{h(tx)}{h(x)} = 1.$$

For any power series $f(z)$, we let $[z^n]f(z)$ denote the coefficient of z^n . We also use the notation

$$[n] = \{1, \dots, n\}$$

for $n \geq 0$. In particular, $[0] = \emptyset$.

2. Main result

We let $\mathbf{w} = (\omega_i)_{i \geq 0}$ denote a fixed sequence of non-negative weights such that $\omega_0 > 0$ and $\omega_k > 0$ for at least one integer $k \geq 2$. The only restriction that we impose on this weight-sequence is that the series

$$\Phi(z) = \sum_{k \geq 0} \omega_k z^k$$

has positive radius of convergence $\rho_\Phi > 0$. We let \mathcal{U}_n denote the set of unlabelled unrooted plane trees with n vertices. Any tree $U \in \mathcal{U}_n$ receives the weight

$$\bar{\omega}(U) = \prod_{v \in U} \omega_{d_U(v)-1},$$

with the index v ranging over all vertices of U , and $d_U(\cdot)$ denoting the degree of a vertex. We let \mathcal{U}_n denote the random tree sampled from the set \mathcal{U}_n with probability proportional to its $\bar{\omega}$ -weight. Of course this is only possible when there are trees of size n with positive weight. Let $\text{span}(\mathbf{w})$ denote the greatest common divisor of all integers $i \geq 0$ satisfying $\omega_i > 0$. The following basic fact clarifies which values of n we may consider. It follows from Schur's lemma, see for example [Wilf \(2006, Thm. 3.15.2\)](#), by noting that $\sum_{v \in U} (d_U(v) - 1) = n - 2$ for $U \in \mathcal{U}_n$ and that all numbers in $\{i/\text{span}(\mathbf{w}) : \omega_i > 0, i \leq n\}$ are relatively prime for all n such that the set is non-empty.

Proposition 2.1. *If there is a tree in the set \mathcal{U}_n having positive weight, then it holds that*

$$n \equiv 2 \pmod{\text{span}(\mathbf{w})}.$$

Conversely, such a tree always exists if n is large enough and belongs to this congruence class.

Likewise, for any positive integer m we let \mathcal{T}_m denote the set of planted plane trees with m vertices. Any such tree T receives weight

$$\omega(T) = \prod_{v \in T} \omega_{d_T^+(v)},$$

with $d_T^+(\cdot)$ denoting the outdegree. We let \mathcal{T}_m denote the tree sampled from \mathcal{T}_m with probability proportional to its ω -weight. This is well-defined when

$$m \equiv 1 \pmod{\text{span}(\mathbf{w})}$$

is large enough.

Our main result reduces the study of the tree \mathcal{U}_n to that of the simply generated (planted) plane tree \mathcal{T}_m .

Theorem 2.2. *Suppose that $\rho_\Phi > 0$. Then there are constants $C, c > 0$ and a random planted plane tree \mathcal{V}_n , that is independent from the family $(\mathcal{T}_m)_m$ of simply generated planted plane trees and has stochastically bounded size*

$$K_n := |\mathcal{V}_n| = O_p(1),$$

such that the random tree \mathcal{S}_n , constructed by connecting the roots of \mathcal{T}_{n-K_n} and \mathcal{V}_n with an edge, satisfies

$$d_{\text{TV}}(\mathcal{U}_n, \mathcal{S}_n) \leq C \exp(-cn)$$

for all n . Furthermore, there is a Galton–Watson tree \mathcal{T} that is subcritical or critical such that

$$\mathcal{V}_n \xrightarrow{d} \mathcal{T}$$

as n becomes large. The offspring distribution of \mathcal{T} may be constructed from the weight-sequence \mathbf{w} in a canonical way, see Equation (4.2) below.

In other words, the random unrooted tree \mathcal{U}_n is very likely to look like the tree \mathcal{T}_{n-K_n} with a small tree attached to its root, to make up for the $K_n = O_p(1)$ missing vertices. This result proves that *almost every* asymptotic property known for simply generated plane trees also holds for simply generated unrooted plane trees. This includes Gromov–Hausdorff–Prokhorov scaling limits, as the attached tree is so small, that it does not change the global geometric shape. It includes Benjamini–Schramm limits, because for the relevant weight-sequences a random vertex is unlikely to fall into the small attached tree or anywhere near it. For the same reason, (central) limit theorems known for the degree of a random root carry over. The approximation also preserves limits for the maximum degree and any other graph-theoretic property, that does not get heavily perturbed by the small $O_p(1)$ -sized tree. As the total variational distance is exponentially small, tail bounds for the diameter and other parameters carry over as well. Our setting is also very general, since the only assumption we made on the weight-sequence is that its generating series $\Phi(z)$ has positive radius of convergence.

A similar approximation was constructed in [Stufler \(2014\)](#) in a different setting, where it was shown that random unlabelled unordered unrooted trees may be approximated by random unlabelled unordered rooted trees, and hence everything known (present and future) about rooted trees carries over *automatically* to the unrooted model.

3. Proof strategy

We let \mathcal{T}_n^* denote the random tree sampled from the set \mathcal{T}_n of all planted plane trees with probability proportional to its $\bar{\omega}$ -weight (as opposed to the ω -weight in the definition of the random tree \mathcal{T}_m). Our first main lemma states that \mathcal{T}_n^* is an excellent approximation of the random tree \mathcal{U}_n .

Lemma 3.1. *Suppose that $\rho_\Phi > 0$. There are constants $C, c > 0$ such that as unrooted trees it holds for all n*

$$d_{\text{TV}}(\mathcal{U}_n, \mathcal{T}_n^*) \leq C \exp(-cn).$$

In order to deduce Theorem 2.2, we consider the fringe subtree $\mathcal{T}_n^{(1)}$ at the first son of the root of \mathcal{T}_n^* and the remaining pruned tree $\mathcal{T}_n^{(2)}$. Recent results for convergent Gibbs partitions ([Stufler, 2018a](#)) show that the maximum size of the two trees belongs to $n + O_p(1)$.

Lemma 3.2. *Suppose that $\rho_\Phi > 0$. Let a, b be non-negative integers with $a + b = n$ such that the event $(|\mathcal{T}_n^{(1)}|, |\mathcal{T}_n^{(2)}|) = (a, b)$ has positive probability. Then the conditioned pair of trees*

$$((\mathcal{T}_n^{(1)}, \mathcal{T}_n^{(2)}) \mid |\mathcal{T}_n^{(1)}| = a, |\mathcal{T}_n^{(2)}| = b)$$

is distributed like the pair of an independent copy of \mathcal{T}_a and of \mathcal{T}_b . Moreover, if we let \mathcal{T}_n^{\max} denote the largest tree in the forest $\mathcal{T}_n^{(1)}, \mathcal{T}_n^{(2)}$, and \mathcal{T}_n^{\min} the smallest one, then there is a Galton–Watson tree \mathcal{T} that is critical or subcritical such that

$$\mathcal{T}_n^{\min} \xrightarrow{d} \mathcal{T} \tag{3.1}$$

as $n \equiv 2 \pmod{\text{span}(\mathbf{w})}$ becomes large. As the tree \mathcal{T} is almost surely finite, this implies that

$$|\mathcal{T}_n^{\max}| = n + O_p(1).$$

The offspring distribution of \mathcal{T} is made explicit in Equation (4.2).

Theorem 2.2 follows readily from this result. We are going to prove Lemmas 3.1 and 3.2 in Section 6.

Remark 3.3. The present work focuses on the case $\rho_\Phi > 0$. It is natural to wonder what happens when the radius of convergence ρ_Φ equals zero.

In fact, Lemma 3.2 still holds in this case with the offspring distribution ξ from Equation (4.2) being concentrated on 0 and hence \mathcal{T} consisting almost surely of a single root vertex. This may be proved in an analogous fashion as the case $\rho_\Phi > 0$ by using results for superexponential Gibbs partitions stated in Stuffer (2016, Lem. 6.17, Thm. 6.18) instead of results for the convergent case of Gibbs partitions in Stuffer (2018a, Lem. 3.3, Thm. 3.1).

We conjecture that Lemma 3.1 and hence also Theorem 2.2 still hold for $\rho_\Phi = 0$. The bounds in the proof of Stuffer (2016, Thm. 6.18) may prove helpful for this, but we did not go through any details.

4. Applications

In this Section we provide some applications of our main result. We collect their proofs in Section 7. In order to precisely state them we make use of a canonical choice for an offspring distribution ξ such that for any admissible integer m the simply generated tree \mathcal{T}_m is distributed like a ξ -Galton–Watson tree \mathcal{T} conditioned on having m vertices. We recall its construction as in Janson (2012, Thm. 7.1), but refer the reader to this source for detailed justifications.

We set

$$\Psi(t) := t\Phi'(t)/\Phi(t) \quad \text{and} \quad \nu := \lim_{t \nearrow \rho_\Phi} \Psi(t) \in]0, \infty]. \tag{4.1}$$

If $\nu \geq 1$, there is a unique finite number τ with $\Psi(\tau) = 1$. If $\nu < 1$, we set $\tau = \rho_\Phi$. Then the random non-negative integer ξ with distribution

$$\mathbb{P}(\xi = k) = \omega_k \tau^k / \Phi(\tau), \quad k \geq 0, \tag{4.2}$$

has mean

$$\mu := \min(\nu, 1) \tag{4.3}$$

and variance

$$\sigma^2 := \tau\Psi'(\tau). \tag{4.4}$$

By [Janson \(2012, Lem. 4.1\)](#) we have for any admissible integer m that the simply generated tree \mathcal{T}_m is distributed like the ξ -Galton–Watson tree \mathcal{T} conditioned on having m vertices.

4.1. Scaling limits and tail bounds for the diameter. Gromov–Hausdorff–Prokhorov scaling limits describe the asymptotic global geometric behaviour of a sequence of random geometric spaces. We refer the reader to the surveys by [Haas \(2018\)](#) and [Le Gall and Miermont \(2012\)](#) for an overview on scaling limits of random trees, and in-depth discussions of real-trees and the Gromov–Hausdorff–Prokhorov distance d_{GHP} on the space \mathbb{K} of equivalence classes of measured compact metric spaces. A brief introduction is provided below.

Theorem 4.1. *Let the random tree \mathcal{U}_n be endowed with the uniform measure ν_n on its leaves. Assume that $\mu = 1$.*

- (1) *Let \mathcal{T}_{Br} denote the (Brownian) continuum random tree introduced by [Aldous \(1993\)](#), and μ_{Br} its probability measure on its set of leaves. If the variance σ^2 is finite, then*

$$\left(\frac{\sigma \mathcal{U}_n}{2\sqrt{n}}, \nu_n \right) \xrightarrow{d} (\mathcal{T}_{\text{Br}}, \mu_{\text{Br}}) \quad (4.5)$$

in the space $(\mathbb{K}, d_{\text{GHP}})$ as n becomes large. Moreover, there are constants $C, c > 0$ such that the diameter $D(\mathcal{U}_n)$ of the tree \mathcal{U}_n satisfies for all n and $x \geq 0$

$$\mathbb{P}(D(\mathcal{U}_n) \geq x) \leq C \exp(-cx^2/n). \quad (4.6)$$

- (2) *Suppose that $\text{span}(\mathbf{w}) = 1$ and that the random variable ξ belongs to the domain of attraction of a stable-law with index $\alpha \in]1, 2[$. Let $(\mathcal{T}_\alpha, \mu_\alpha)$ denote the α -stable Lévy tree introduced by [Le Gall and Le Jan \(1998\)](#). Then there is a slowly varying sequence $g(n)$ such that*

$$\left(\frac{\mathcal{U}_n}{g(n)n^{1-1/\alpha}}, \nu_n \right) \xrightarrow{d} (\mathcal{T}_\alpha, \mu_\alpha) \quad (4.7)$$

in the space $(\mathbb{K}, d_{\text{GHP}})$ as n becomes large. Moreover, for all $\delta \in]0, \alpha[$ there are constants $C, c > 0$ such that for all n and $x \geq 0$

$$\mathbb{P}(D(\mathcal{U}_n) \geq x) \leq C \exp \left(-c \left(\frac{x}{g(n)n^{1-1/\alpha}} \right)^\delta \right). \quad (4.8)$$

Limits of this form have far-reaching consequences, see for example the recent work of [Addario-Berry et al. \(2018\)](#). Let us very briefly explain relevant notation used in [Theorem 4.1](#). For any metric space (Z, d_Z) and any real number $\beta > 0$ we let βZ denote the rescaled space $(Z, \beta d_Z)$. The *Hausdorff-distance* defines a metric on the collection of compact subsets of Z . It is given by

$$d_{\text{H}}(A, B) = \inf \{ \epsilon > 0 \mid A \subset U_\epsilon(B), B \subset U_\epsilon(A) \}$$

with A^ϵ denoting the ϵ -hull of the subset A . That is, the set of all points in Z with distance less than ϵ from A . The *Prokhorov-distance* metrizes weak convergence of probability measures on the Borel σ -field $\mathcal{B}(Z)$ of Z . For any two such measures \mathbb{P} and \mathbb{P}' it is given by

$$d_{\text{P}}(\mathbb{P}, \mathbb{P}') = \inf \{ \epsilon > 0 \mid \forall A \in \mathcal{B}(X) : \mathbb{P}(A) \leq \mathbb{P}'(A^\epsilon) + \epsilon, \mathbb{P}'(A) \leq \mathbb{P}(A^\epsilon) + \epsilon \}.$$

Let $(X, d_X), (Y, d_Y)$ be compact metric spaces endowed with Borel probability measures \mathbb{P}_X and \mathbb{P}_Y . Unless these spaces are subspaces of a common space, we cannot measure their distance using the Hausdorff or Prokhorov metric. The natural solution is to consider embeddings. The *Gromov–Hausdorff* distance of (X, d_X) and (Y, d_Y) is given by the minimal Hausdorff distance of isometric copies of X and Y in a common space. That is

$$d_{\text{GH}}(X, Y) = \inf_{\phi_X, \phi_Y} d_{\text{H}}(\phi_X(X), \phi_Y(Y))$$

with the indices ranging over all isometric embeddings $\phi_X : X \rightarrow Z$ and $\phi_Y : Y \rightarrow Z$ for all choices of metric spaces (Z, d_Z) . Letting $\phi_X^* \mathbb{P}_X$ and $\phi_Y^* \mathbb{P}_Y$ denote the push-forwards of the measures, the *Gromov–Hausdorff–Prokhorov* distance is defined by

$$d_{\text{GHP}}((X, \mathbb{P}_X), (Y, \mathbb{P}_Y)) = \inf_{\phi_X, \phi_Y} \min(d_{\text{H}}(\phi_X(X), \phi_Y(Y)), d_{\text{P}}(\phi_X^* \mathbb{P}_X, \phi_Y^* \mathbb{P}_Y)).$$

The GHP distance satisfies the axioms of a pre-metric and hence induces a metric on the collection of equivalence classes \mathbb{K} of measured compact metric spaces. Here two such spaces are equivalent if and only if there is a measure-preserving isometry between them. Here we admittedly have swept a small issue under the table, as these classes are not sets and hence technically their collection is not a well-defined object. But this issue is easily resolved by working with representatives instead, compare with [Burago et al. \(2001, Remark 7.2.5\)](#).

The metric space $(\mathbb{K}, d_{\text{GHP}})$ is known to be complete and separable according to [Abraham et al. \(2013, Thm. 2.3\)](#), and hence we use the classical notions of weak convergence of Borel-measures on Polish spaces by [Billingsley \(1971, Ch. 1\)](#).

4.2. Benjamini-Schramm limits and degree distribution. The term *Benjamini-Schramm limit* refers to the local weak limits of a sequence of finite random graphs with respect to an uniformly at random sampled point. The name is used in honour of [Benjamini and Schramm \(2001\)](#).

Theorem 4.2. *Suppose that $\mu = 1$ and $\sigma^2 < \infty$.*

- (1) *Let \mathcal{T}^\bullet denote the Benjamini-Schramm limit of the conditioned Galton–Watson trees $(\mathcal{T}_m)_{m \geq 1}$ established by [Aldous \(1991\)](#). Then \mathbb{U}_n converges also in the Benjamini-Schramm sense toward \mathcal{T}^\bullet .*
- (2) *If N_d denotes the number of vertices of degree $d \geq 1$ in the tree \mathbb{U}_n , then*

$$\frac{N_d - \mathbb{P}(\xi = d - 1)}{\sqrt{n}} \xrightarrow{d} \mathcal{N}(0, \sigma_d^2)$$

for some $0 < \sigma_d < \infty$.

The distribution of \mathcal{T}^\bullet is made explicit in [Janson \(2012, Remark 7.13, Remark 5.3\)](#). We very briefly recall the notion of local weak convergence used in [Theorem 4.2](#) and refer the reader to the notes of [Curien \(2016\)](#) for a more in-depth treatment. A *locally finite* graph may have infinitely many vertices, but each has finite degree. Any two such graphs G and H with distinguished vertices v_G and v_H are considered as *isomorphic*, if there is a bijection between their vertex sets that preserves the incidence relation and maps the root vertices to each other. This is denoted by $(G, v_G) \simeq (H, v_H)$. For any $\ell \geq 0$, the subgraph consisting of all vertices with graph distance at most ℓ from v_G is denoted by $U_\ell(G, v_G)$ and called the ℓ -*neighbourhood*.

We consider the ℓ -neighbourhood as a rooted graph. The *local distance* of the two rooted graphs H and G is defined by

$$d_L((G, v_G), (H, v_H)) = 1/(1 + \sup\{\ell \geq 0 \mid U_\ell(G, v_G) \simeq U_\ell(H, v_H)\}).$$

This yields a metric on the collection of (representatives of) isomorphism classes of rooted locally finite graphs, which is known to be complete and separable. A sequence of random rooted graphs (G_n, v_n) converges in the *local weak sense* toward a random limit graph (G, v) , if for each $\ell \geq 0$ and each rooted graph (Q, v_Q) it holds that

$$\mathbb{P}(U_\ell(G_n, v_n) \simeq (Q, v_Q)) \rightarrow \mathbb{P}(U_\ell(G, v) \simeq (Q, v_Q))$$

as n becomes large.

4.3. The maximum degree and other parameters. Theorem 2.2 reduces the study of the extremal vertex degree sizes of the unrooted tree \mathcal{U}_n to those of the random tree \mathcal{S}_n . It is clear that central limit theorems or laws of large numbers for the degree sizes in the simply generated tree \mathcal{T}_m also hold for the tree \mathcal{T}_{K_n} , since $K_n = n + O_p(1)$. The trees \mathcal{T}_{K_n} and \mathcal{S}_n differ only by a small tree with stochastically bounded size that gets attached to the root of \mathcal{T}_{K_n} . Consequently, available results for the maximum degree of $(\mathcal{T}_m)_{m \geq 1}$ automatically carry over to the random unrooted tree \mathcal{U}_n . In particular, we obtain by recent results of Kortchemski (2015, Thm. 1) a central limit theorem for the largest degree in the subcritical setting.

Corollary 4.3. *Suppose that $\mu < 1$ and*

$$\mathbb{P}(\xi = k) = f(k)k^{-\beta}$$

for some slowly varying function f and a parameter $\beta > 2$. Set

$$\alpha = \min(2, \beta - 1)$$

and let $(Y_t)_{t \geq 1}$ denote a spectrally positive Lévy process with Laplace exponent

$$\mathbb{E}[\exp(-\lambda Y_t)] = \exp(t\lambda^\alpha).$$

Then there exists a slowly varying function g such that the maximum degree $\Delta(\mathcal{U}_n)$ satisfies

$$\frac{(1 - \nu)n - \Delta(\mathcal{U}_n)}{g(n)n^{1/\alpha}} \xrightarrow{d} Y_1.$$

The second largest degree $\Delta_2(\mathcal{U}_n)$ satisfies

$$\Delta_2(\mathcal{U}_n) = O_p(g(n)n^{1/\alpha}).$$

Informally speaking, pretty much all “interesting” properties known for simply generated planted plane trees carry over to the random tree \mathcal{U}_n using the approximation in Theorem 2.2 (and possibly Lemma 3.2). See in particular the comprehensive survey of Janson (2012) for a wealth of further results that may be transferred.

5. Preliminaries

5.1. Subexponential power series. Subexponential sequences were studied by Chover et al. (1973), Embrechts (1983), and Embrechts and Omey (1984). Up to tilting and rescaling, these sequences correspond to subexponential densities of random variables with values in a lattice, and hence belong to the context of heavy-tailed and subexponential distributions, see the book by Foss et al. (2013).

Definition 5.1. Let $d \geq 1$ be an integer. A power series $g(x) = \sum_{k \geq 0} g_k x^k$ with non-negative coefficients and radius of convergence $\rho > 0$ is subexponential with span d , if $g_k = 0$ whenever k is not divisible by d , and

$$\frac{g_k}{g_{k+d}} \sim \rho^d, \quad \frac{1}{g_k} \sum_{i+j=k} g_i g_j \sim 2g(\rho) < \infty \quad (5.1)$$

as $k \equiv 0 \pmod{d}$ becomes large.

Following lemma is useful for deriving the behaviour of a randomly stopped random walk with i.i.d. steps (after centralizing the coefficients of $g(x)$ properly), and in our case it will prove handy as enumerative tool.

Lemma 5.2 (Foss et al., 2013, Thm. 4.8, 4.30). *If $g(x)$ is subexponential with span d and radius of convergence $\rho > 0$, and $f(x)$ is a non-constant power series with non-negative coefficients that is analytic at ρ , then $f(g(x))$ is subexponential with span d and radius of convergence ρ . Further, as $n \equiv 0 \pmod{d}$ becomes large,*

$$[x^n]f(g(x)) \sim f'(g(\rho))[x^n]g(x).$$

It was observed in Stufler (2018a) by building on results for simply generated trees (Janson, 2012), that power series satisfying a simple recursive relation are always subexponential up to a shift.

Lemma 5.3 (Stufler, 2018a, Lem. 3.3). *Let $g(x)$ and $f(x)$ be power series with non-negative coefficients such that*

$$g(x) = xf(g(x)).$$

If the series $f(x)$ has positive radius of convergence and satisfies $f(0) > 0$ and $[x^k]f(x) > 0$ for at least one $k \geq 2$, then $g(x)/x$ is subexponential with span d for some $d \geq 1$.

Lemma 5.4. *Let $g(x) = \sum_{k \geq 0} g_k x^k$ and $f(x) = \sum_{k \geq 0} f_k x^k$ be power series with non-negative coefficients. Suppose that the radii of convergence ρ_g and ρ_f of $g(x)$ and $f(x)$ satisfy $\rho_g > \rho_f > 0$. If $d \geq 1$ is an integer with*

$$\frac{f_n}{f_{n+d}} \sim \rho_f^d$$

as $n \equiv 0 \pmod{d}$ becomes large, then there exist constants $C, c > 0$ such that

$$\frac{g_n}{f_n} \leq C \exp(-cn)$$

holds for all $n \equiv 0 \pmod{d}$ with $f_n > 0$.

Proof: The assumption $f_n/f_{n+d} \sim \rho_f^d$ implies that the radius of convergence of $f(x)$ does not change if we restrict to coefficients f_n with index in the lattice $d\mathbb{N}_0$. It also implies that $f_{kd} > 0$ for all but finitely many integers $k \geq 0$. Hence without loss of generality we may assume that $d = 1$ and $f_n > 0$ for all $n \geq 0$.

Our assumption on $f(x)$ now reads $f_n/f_{n+1} \sim \rho_f$. Hence for any $\delta > 0$ there is an integer $n_0 \geq 0$ such that

$$\frac{f_{n_0}}{f_{n_0+k}} \leq (1 + \delta)^k \rho_f^k$$

for all $k \geq 0$. Thus

$$\begin{aligned} \sum_{n \geq 0} \frac{g_n}{f_n} (1 + \delta)^n &= \sum_{0 \leq n < n_0} \frac{g_n}{f_n} (1 + \delta)^n + (1 + \delta)^{n_0} \sum_{k \geq 0} \frac{g_{n_0+k}}{f_{n_0+k}} (1 + \delta)^k \\ &\leq \sum_{0 \leq n < n_0} \frac{g_n}{f_n} (1 + \delta)^n + \frac{(1 + \delta)^{n_0}}{f_{n_0}} \sum_{k \geq 0} g_{n_0+k} ((1 + \delta)^2 \rho_f)^k. \end{aligned}$$

Since $\rho_g > \rho_f$ we may choose $\delta > 0$ small enough such that $(1 + \delta)^2 \rho_f < \rho_g$. With this choice of δ it follows that

$$\sum_{n \geq 0} \frac{g_n}{f_n} (1 + \delta)^n < \infty.$$

Consequently, there are constants $C, c > 0$ such that

$$\frac{g_n}{f_n} \leq C \exp(-cn)$$

for all $n \geq 0$. □

5.2. Combinatorial Classes. In order to understand how symmetries of random unrooted plane trees typically behave, we will make use of some enumerative results. We aim to recall just enough combinatorial background for the non-specialist. A comprehensive survey of the theory is given in the books by [Flajolet and Sedgewick \(2009\)](#) and [Bergeron et al. \(1998\)](#).

A *weighted combinatorial class* is given by a set \mathcal{C}^ω of countably many objects equipped with a size function

$$|\cdot| : \mathcal{C}^\omega \rightarrow \mathbb{N}_0$$

and a weight function

$$\omega : \mathcal{C}^\omega \rightarrow \mathbb{R}_{\geq 0}.$$

Additionally, the subset $\mathcal{C}_n^\omega \subset \mathcal{C}^\omega$ containing all n -sized objects in \mathcal{C}^ω is required to satisfy

$$\sum_{C \in \mathcal{C}_n^\omega} \omega(C) < \infty$$

for all $n \in \mathbb{N}_0$. Each $C \in \mathcal{C}_n^\omega$ is said to be comprised of n *atoms*. There are two cases to be set apart:

- (1) Each atom bears a distinct label and the class is termed *labelled*. Then the (*weighted exponential*) *generating series* of \mathcal{C}^ω is the formal power series

$$\mathcal{C}^\omega(x) := \sum_{C \in \mathcal{C}^\omega} \frac{\omega(C)}{|C|!} x^{|C|}.$$

- (2) Whenever atoms are not distinguishable, we say that the class is *unlabelled*. For this section we denote such classes by $\tilde{\mathcal{C}}^\omega$ and the resulting (*weighted ordinary*) *generating series* is defined as the formal power series

$$\tilde{\mathcal{C}}^\omega(x) := \sum_{C \in \tilde{\mathcal{C}}^\omega} \omega(C) x^{|C|}.$$

Without loss of generality we will assume that any object $C \in \mathcal{C}_n^\omega$ has labels in the set $[n]$ and that \mathcal{C}^ω is defined in a coherent way such that it contains any possible relabelling of C with labels in $[n]$. For $\sigma : [n] \rightarrow [n]$ and $C \in \mathcal{C}_n^\omega$ we denote by $\sigma.C$ the object obtained by replacing the label $i_v \in [n]$ of the atom v of C by $\sigma(i_v)$ for all atoms $v \in C$.

We may alternatively view $\tilde{\mathcal{C}}^\omega$ as the set of equivalence classes under the relation which terms two objects $C, C' \in \mathcal{C}_n^\omega$ *isomorphic* if and only if one object is obtained by relabelling the other one, i.e. there exists a permutation $\sigma : [n] \rightarrow [n]$ such that $\sigma.C = C'$. Regarding this, it makes sense to impose that any labelled element in a equivalence class $\tilde{C} \in \tilde{\mathcal{C}}^\omega$ receives the same weight $\omega(\tilde{C})$.

5.2.1. *Cycle index sum.* Let $C \in \mathcal{C}^\omega$ and σ be a permutation on the label set of C such that $\sigma.C = C$. Then σ is called *automorphism* of C . Note that any object has at least one automorphism, namely the identity. The class of *symmetries* of \mathcal{C}^ω is defined as the collection of objects paired with an automorphism, that is

$$\text{Sym}(\mathcal{C}^\omega) := \{(C, \sigma) : C \in \mathcal{C}^\omega \text{ and } \sigma.C = C\}.$$

Any permutation σ can be decomposed into disjoint cycles and we denote by σ_i the number of cycles of length i in σ . In particular, σ_1 counts the number of fixpoints. The *cycle index sum* of a class \mathcal{C}^ω is then defined as the formal power series

$$Z_{\mathcal{C}^\omega}(s_1, s_2, \dots) := \sum_{k \geq 0} \sum_{(C, \sigma) \in \text{Sym}(\mathcal{C}_k^\omega)} \frac{\omega(C)}{k!} s_1^{\sigma_1} \cdots s_k^{\sigma_k}.$$

Considering symmetries is useful, as it provides a way of counting orbits.

Lemma 5.5 (Bodirsky et al., 2011, Lem. 1). *For any $\tilde{C} \in \tilde{\mathcal{C}}^\omega$ there are precisely $|C|!$ many symmetries (C, σ) in $\text{Sym}(\mathcal{C}^\omega)$ such that $C \in \tilde{C}$. Consequently,*

$$\tilde{\mathcal{C}}^\omega(x) = Z_{\mathcal{C}^\omega}(x, x^2, x^3, \dots).$$

Example 5.6. In the context of our intended applications, where we consider a weight sequence $\mathbf{w} = (\omega_i)_{i \geq 0}$ of non-negative real numbers, \mathcal{T}^ω is typically a class (of labelled or unlabelled, planted or plane, rooted or unrooted) trees and, for $T \in \mathcal{T}^\omega$,

$$\omega(T) := \prod_{v \in T} \omega_{f(v)}$$

for some arbitrary mapping f from the vertex set of T into \mathbb{N}_0 . For example $f(v) := d_T^+(v)$ may count the out-degree of a given vertex v . Furthermore, the size of a tree will be given by its number of vertices.

5.2.2. *Constructions.* In this section we present several constructions designed for obtaining more complex classes out of simpler ones. Let \mathcal{A}^ω and \mathcal{B}^ν be labelled classes and their unlabelled counterparts be denoted as before. For any arbitrary class \mathcal{C}^ω , we define the set $\mathcal{C}^\omega[P]$, which contains all objects in \mathcal{C}^ω of size $|P|$ relabelled canonically by a given finite set of labels $P \subset \mathbb{N}$.

Product. The *product* of objects $A \in \mathcal{A}^\omega$ and $B \in \mathcal{B}^\nu$ is the tuple (A, B) where each atom is relabelled according to labels in $|A| + |B|$. Formally,

$$\begin{aligned} \mathcal{A}^\omega \cdot \mathcal{B}^\nu &:= \{(A, B) : A \in \mathcal{A}^\omega[P_1], B \in \mathcal{B}^\nu[P_2], \\ &\quad (P_1, P_2) \text{ is a partition of } [|A| + |B|]\}. \end{aligned}$$

By discarding the labels, the product of the *unlabelled* classes $\tilde{\mathcal{A}}^\omega$ and $\tilde{\mathcal{B}}^\nu$ is given by the set theoretic product

$$\tilde{\mathcal{A}}^\omega \cdot \tilde{\mathcal{B}}^\nu := \{(A, B) : A \in \tilde{\mathcal{A}}^\omega, B \in \tilde{\mathcal{B}}^\nu\}.$$

In both cases, the size function is the canonical extension $|(A, B)| := |A| + |B|$ of the size functions of the underlying classes, such as the weight of (A, B) being given by $\omega(A) \cdot \nu(B)$. As a straight-forward consequence the generating series satisfy

$$(\mathcal{A}^\omega \cdot \mathcal{B}^\nu)(x) = \mathcal{A}^\omega(x) \cdot \mathcal{B}^\nu(x) \quad \text{and} \quad (\tilde{\mathcal{A}}^\omega \cdot \tilde{\mathcal{B}}^\nu)(x) = \tilde{\mathcal{A}}^\omega(x) \cdot \tilde{\mathcal{B}}^\nu(x). \quad (5.2)$$

Composition. For this construction we assume that $\mathcal{B}_0^\nu = \emptyset$. The composition $\mathcal{A}^\omega \circ \mathcal{B}^\nu$ is the set of objects obtained by picking $A \in \mathcal{A}^\omega$ and replacing every atom in A by an entire object from \mathcal{B}^ν relabelled properly according to the compound size. In other words, $\mathcal{A}^\omega \circ \mathcal{B}^\nu$ contains all sequences of the form

$$(A, B_1, \dots, B_k)_{\simeq}, \quad A \in \mathcal{A}_k^\omega, \quad B_i \in \mathcal{B}^\nu[P_i], \quad 1 \leq i \leq k,$$

where $k \geq 0$ and (P_1, \dots, P_k) is a partition of $[\sum_{1 \leq i \leq k} |B_i|]$. The relation “ \simeq ” terms two sequences (A, B_1, \dots, B_k) and (A', B'_1, \dots, B'_k) in $\mathcal{A}^\omega \circ \mathcal{B}^\nu$ isomorphic if $A = A'$ and for any permutation $\sigma : [k] \rightarrow [k]$ such that $\sigma.A = A$ it holds $B_{\sigma(i)} = B'_i$ for $1 \leq i \leq k$. Hence, any $M \in \mathcal{A}^\omega \circ \mathcal{B}^\nu$ possesses a *core structure* $A \in \mathcal{A}^\omega$ and *components* $B_1, \dots, B_{|A|} \in \mathcal{B}^\nu$. The size is then given by $|M| := \sum_{1 \leq i \leq |A|} |B_i|$ and the compound weight by $\omega(A) \prod_{1 \leq i \leq |A|} \nu(B_i)$. This implies that the generating series is

$$(\mathcal{A}^\omega \circ \mathcal{B}^\nu)(x) = \mathcal{A}^\omega(\mathcal{B}^\nu(x)).$$

The unlabelled composition $\tilde{\mathcal{A}}^\omega \circ \tilde{\mathcal{B}}^\nu$ is defined equivalently, except for skipping the part where the objects are relabelled. Due to possible symmetries appearing when removing the labels, the generating series satisfies a more complex formula, namely

$$(\tilde{\mathcal{A}}^\omega \circ \tilde{\mathcal{B}}^\nu)(x) = Z_{\mathcal{A}^\omega}(\tilde{\mathcal{B}}^\nu(x), \mathcal{B}^{\nu^2}(x^2), \mathcal{B}^{\nu^3}(x^3), \dots), \quad (5.3)$$

where \mathcal{B}^{ν^k} denotes the class with weight function $\nu(B)^k$ for $B \in \mathcal{B}^{\nu^k}$. This is stated in Bergeron et al. (1998, Prop. 11).

Cycle. The labelled class CYC^κ contains all bijections that are cycles and by convention the empty set is mapped onto itself. The size of a cycle is its length and the weight of a cycle $\tau \in \text{CYC}^\kappa$ depends only on the length, i.e. $\kappa(\tau) = \kappa_{|\tau|}$ for some non-negative real-valued sequence $(\kappa_k)_{k \geq 0}$ with $\kappa_0 = 1$. This follows from the fact that all cycles of the same length are isomorphic.

The symmetries of cycles may be described explicitly. It is elementary that any automorphism of a cycle τ with length $k \geq 1$ must be of the form τ^i for $0 \leq i \leq k - 1$. The disjoint cycles in τ^i all have the same length $k / \text{gcd}(i, k)$, and τ^i has $\text{gcd}(i, k)$ of them. There are precisely $(k - 1)!$ cycles on a fixed k -element set. This yields

$$Z_{\text{CYC}^\kappa}(s_1, s_2, \dots) = 1 + \sum_{k \geq 1} \frac{\kappa_k}{k} \sum_{i=0}^{k-1} s_{k/\text{gcd}(i,k)}^{\text{gcd}(i,k)}. \quad (5.4)$$

We let φ denote Euler’s totient function. That is, for each $d \geq 1$ the number $\varphi(d)$ counts the integers that are relatively prime to d . For each $k \geq 1$, there are precisely $\varphi(d)$ many integers $1 \leq i \leq k$ such that $k / \text{gcd}(i, k) = d$. Hence Equation (5.4) may be rephrased by

$$Z_{\text{CYC}^\kappa}(s_1, s_2, \dots) = 1 + \sum_{d \geq 1} \frac{\varphi(d)}{d} \sum_{j \geq 1} \frac{\kappa_{jd}}{j} s_d^j. \quad (5.5)$$

We denote by the composition $\text{CYC}^\kappa \circ \mathcal{B}^\nu$ the labelled class containing all cyclic orderings of objects in \mathcal{B}^ν relabelled properly according to their compound size. Formally, $\text{CYC}^\kappa \circ \mathcal{B}^\nu$ contains all sequences of the form

$$(C, B_1, \dots, B_k)_{\simeq}, \quad C \in \text{CYC}_k^\kappa, \quad B_i \in \mathcal{B}^\nu[P_i], \quad 1 \leq i \leq k,$$

where (P_1, \dots, P_k) is a partition of $[\sum_{1 \leq i \leq k} |B_i|]$ and the relation “ \simeq ” is as before. The size of such an object is given by $\sum_{1 \leq i \leq k} |B_i|$ and the weight by $\kappa(C) \prod_{1 \leq i \leq k} \nu(B_i)$.

5.3. *Boltzmann distribution.* Given the (labelled) weighted combinatorial class \mathcal{C}^ω and a parameter y such that $0 < \mathcal{C}^\omega(y) < \infty$ we define the corresponding Boltzmann probability measure by

$$\mathbb{P}(\mathcal{C} = C) := \frac{\omega(C)y^{|C|}}{|C|! \mathcal{C}^\omega(y)}, \quad C \in \mathcal{C}^\omega.$$

The random variable \mathcal{C} is taking values in the entire space \mathcal{C}^ω and when conditioning on having size n we obtain the object \mathcal{C}_n drawn proportional to its ω -weight from all objects in \mathcal{C}^ω of size n . That is

$$\mathbb{P}(\mathcal{C}_n = C) = \frac{w(C)}{\sum_{C' \in \mathcal{C}_n^\omega} w(C')}, \quad C \in \mathcal{C}_n^\omega.$$

Let \mathcal{X}^2 denote the unique class consisting of a single ordered pair of atoms that receives weight 1 and \mathcal{B}^ν be a labelled class such that $\mathcal{B}_0^\nu = \emptyset$. Denote by \mathcal{X}_n the $(\mathcal{X}^2 \circ \mathcal{B}^\nu)_n$ -valued random variable drawn proportional to its compound weight. Later we will need the distribution of the remainder \mathcal{R}_n , which is obtained by removing “the” largest component of \mathcal{X}_n (if the two components of \mathcal{X}_n are equally sized, we pick an arbitrary one). Therefore, we state a far simplified version of [Stuffer \(2018a, Thm. 3.4\)](#) (here \mathcal{X}^2 is replaced by an arbitrary combinatorial class).

Lemma 5.7. *Suppose that $x^{-1} \mathcal{B}^\nu(x)$ is subexponential with radius of convergence ρ and span d . Then*

$$\mathcal{R}_n \xrightarrow{d} \mathcal{R}$$

as $n \equiv 2 \pmod d$ becomes large, where \mathcal{R} is given by

$$\mathbb{P}(\mathcal{R} = B) = \frac{\nu(B)\rho^{|B|}}{|B|! \mathcal{B}^\nu(\rho)}, \quad B \in \mathcal{B}^\nu.$$

This implies that, letting n tend to infinity on a properly chosen lattice, there emerges one giant component in \mathcal{X}_n containing all but a stochastically bounded number of atoms.

6. Proof of the main result

We briefly recapitulate the assumptions as well as the notation needed for the entire proof.

Our only restriction is that a given weight sequence $\mathbf{w} = (\omega_k)_{k \geq 0}$ is such that $\omega_0 > 0$ and $\omega_k > 0$ for some $k \in \mathbb{N}$. Further, $\Phi(z) := \sum_{k \geq 0} \omega_k z^k$ is assumed to have finite radius of convergence $\rho_\Phi > 0$.

For the remaining proof section it will prove handy to transition into the language of combinatorial classes introduced in Section 5.2. More precisely, we denote by

$L(\mathcal{U})^\omega$ and $L(\mathcal{T})^{\bar{\omega}}$ the classes of labelled unrooted plane trees and of labelled planted trees, respectively, where ω and $\bar{\omega}$ are the weight functions given by

$$\omega(T) = \prod_{v \in T} \omega_{d_T^+(v)} \quad \text{or} \quad \bar{\omega}(T) = \prod_{v \in T} \omega_{d_T(v)-1}$$

for any arbitrary tree T . With this at hand we define the unlabelled counterparts $\mathcal{U}^\omega = \widetilde{L(\mathcal{U})^\omega}$ and $\mathcal{T}^{\bar{\omega}} = \widetilde{L(\mathcal{T})^{\bar{\omega}}}$.

For an arbitrary class (labelled or unlabelled) \mathcal{C}^ν we denote by $\mathcal{C}_n \sim \mathcal{C}^\nu$ the random variable drawn from all objects in \mathcal{C}^ν of size n proportional to its ν -weight, i.e.

$$\mathbb{P}(\mathcal{C}_n = C) = \frac{\nu(C)}{\sum_{C' \in \mathcal{C}_n^\nu} \nu(C')}, \quad C \in \mathcal{C}_n^\nu.$$

In this fashion let

- (1) $\mathcal{U}_n \sim \mathcal{U}^{\bar{\omega}}$ be drawn proportional to its $\bar{\omega}$ -weight from all unlabelled unrooted plane trees of size n ,
- (2) $\mathcal{T}_n \sim \mathcal{T}^{\bar{\omega}}$ proportional to its $\bar{\omega}$ -weight and
- (3) $\mathcal{T}_n^* \sim \mathcal{T}^\omega$ proportional to its ω -weight from all unlabelled planted trees of size n .

6.1. *Galton-Watson trees and tilting.* Recall the notation from Equations (4.1)–(4.4). For any plane tree T it holds that

$$\mathbb{P}(\mathcal{T} = T) = \frac{\omega(T) \rho_{\mathcal{T}^\omega}^{|T|}}{\mathcal{T}^\omega(\rho_{\mathcal{T}^\omega})} \quad (6.1)$$

with $\rho_{\mathcal{T}^\omega} = \tau/\Phi(\tau)$ denoting the radius of convergence of the unique power series $\mathcal{T}^\omega(x)$ satisfying

$$\mathcal{T}^\omega(x) = x\Phi(\mathcal{T}^\omega(x)). \quad (6.2)$$

See Janson (2012, Remark 3.2 and Remark 7.5) for a justification of these facts. As we shall see in the proof of Lemma 3.2, the random tree \mathcal{T} is also the distributional limit in (3.1).

Remark 6.1. The preceding text implies that we may without loss of generality assume that \mathbf{w} is a probability weight sequence and $\rho_{\mathcal{T}^\omega} = 1$ whenever we sample objects conditioned on having a certain size. This is due to the fact that, for $a, b > 0$, the rescaled sequence $\tilde{\mathbf{w}} := (ab^k \omega_k)_{k \geq 0}$ yields for any tree $T \in \mathcal{T}_n^\omega$

$$\tilde{\omega}(T) := \prod_{v \in T} \tilde{\omega}_{d^+(v)} = a^n b^{\sum_{v \in T} d^+(v)} \omega(T) = a^n b^{n-1} \omega(T).$$

Then, both sequences induce the same conditioned probability

$$\mathbb{P}(\tilde{\mathcal{T}} = T \mid |\tilde{\mathcal{T}}| = n) = \frac{\tilde{\omega}(T)}{\sum_{T' \in \mathcal{T}_n^{\tilde{\omega}}} \tilde{\omega}(T')} = \frac{\omega(T)}{\sum_{T' \in \mathcal{T}_n^\omega} \omega(T')} = \mathbb{P}(\mathcal{T} = T \mid |\mathcal{T}| = n).$$

Further, choosing $a = \Phi(\tau)^{-1}$ and $b = \tau$ we obtain

$$\begin{aligned} \mathcal{T}^{\tilde{\omega}}(x) &= \sum_{T \in \mathcal{T}_n^{\tilde{\omega}}} \tilde{\omega}(T) x^{|T|} = \sum_{T \in \mathcal{T}_n^\omega} \Phi(\tau)^{-|T|} \tau^{|T|-1} \omega(T) x^{|T|} \\ &= \tau^{-1} \sum_{T \in \mathcal{T}_n^\omega} \rho_{\mathcal{T}^\omega}^{|T|} \omega(T) x^{|T|} = \frac{\mathcal{T}^\omega(\rho_{\mathcal{T}^\omega} x)}{\mathcal{T}^\omega(\rho_{\mathcal{T}^\omega})} \end{aligned}$$

and we readily deduce that $\rho_{\mathcal{T}^{\tilde{\omega}}} = 1$.

6.2. *The geometric approximation (Proof of Lemma 3.2).* Our first aim is to prove Lemma 3.2. Therefore, we observe that there exists a weight-preserving bijection between trees in $\mathcal{T}^{\bar{\omega}}$ and ordered pairs of trees in \mathcal{T}^{ω} , i.e. the random object \mathcal{T}_n^* can be decomposed into $(\mathcal{T}_{n_1}, \mathcal{T}_{n_2})$ such that $n_1 + n_2 = n$. To see that we simply split up $T \in \mathcal{T}^{\bar{\omega}}$ into $(T^{(1)}, T^{(2)})$, where $T^{(1)}$ is the fringe subtree at the first son of the root of T and $T^{(2)}$ is the remaining pruned tree. Clearly, this guarantees that the roots of $T^{(1)}$ and $T^{(2)}$ receive weights according to their outdegree, by which we obtain trees in \mathcal{T}^{ω} . (Note that for any non-rooted node v in an arbitrary rooted tree T it holds that $d_T(v) - 1 = d_T^+(v)$.) On the other hand, given $(T^{(1)}, T^{(2)}) \in \mathcal{T}^{\omega} \cdot \mathcal{T}^{\omega}$, connect the root of $T^{(1)}$ to the root of $T^{(2)}$ as the leftmost son to obtain an object in $\mathcal{T}^{\bar{\omega}}$. By this procedure the neighbourhood of both roots increases by one, leading to the proper weighting. Formally, this identity allows us to interpret \mathcal{T}_n^* as a

$$\mathcal{T}^{\omega} \cdot \mathcal{T}^{\omega} \simeq \mathcal{X}^2 \circ \mathcal{T}^{\omega} \tag{6.3}$$

composite object, with \mathcal{X}^2 denoting the unique class consisting of a single ordered pair of atoms that receives weight 1. Let

$$\mathcal{T}^{\omega}(x) := \sum_{T \in \mathcal{T}^{\omega}} \omega(T)x^{|T|}$$

be the ordinary generating function of \mathcal{T}^{ω} , which is known to satisfy the equation

$$\mathcal{T}^{\omega}(x) = x\Phi(\mathcal{T}^{\omega}(x)),$$

see for example Drmota (2009, Sec. 3.1.4). Consequently, the assumptions of Lemma 5.3 are fulfilled and $\mathcal{T}^{\omega}(z)/z$ is subexponential with $d = \text{span}(\mathbf{w})$.

The class \mathcal{T}^{ω} is *asymmetric* in the sense that each object only admits the trivial automorphism. Hence, there are exactly $n!$ ways to label objects of size n and it makes no difference whether we draw an object from \mathcal{T}_n^{ω} proportional to its ω -weight or any of the $n!$ labelled counterparts from the set of all labelled objects in $L(\mathcal{T}_n^{\omega})$ proportional to their ω -weights. Further, the ordinary and exponential generating series coincide in this case. Keeping that comment in mind, the assumptions of Lemma 5.7 are fulfilled. Denote by \mathcal{T}_n^{\min} and \mathcal{T}_n^{\max} “the” smallest and the largest tree corresponding to \mathcal{T}_n^* . Whenever both trees have the same size, we pick an arbitrary order. Then, Lemma 5.7 states that upon removing the largest component of $(\mathcal{T}_n^{\min}, \mathcal{T}_n^{\max})$ (which is \mathcal{T}_n^{\max} per definition), the remainder converges in total variation to the random unlabelled planted tree \mathcal{T} given by

$$\mathbb{P}(\mathcal{T} = T) = \frac{\omega(T)\rho_{\mathcal{T}^{\omega}}^{|T|}}{\mathcal{T}^{\omega}(\rho_{\mathcal{T}^{\omega}})}, \quad T \in \mathcal{T}^{\omega}.$$

By Equation (6.1) we know that \mathcal{T} follows the distribution of a Galton-Watson tree with offspring distribution given in Equation (4.2). This fully proves Lemma 3.2.

6.3. *Approximating unlabelled by labelled trees (Proof of Lemma 3.1).* Let $\mathcal{L}_n \sim L(\mathcal{U})^{\bar{\omega}}$ denote the random tree drawn proportional to its $\bar{\omega}$ -weight from all labelled unrooted plane trees of size n . The crucial step towards proving Lemma 3.1 is showing that, for large n , the unlabelled object \mathcal{U}_n can be approximated by drawing \mathcal{L}_n and dropping the labels afterwards. For any labelled tree T , let \tilde{T} be its unlabelled counterpart.

Lemma 6.2. *There exist constants $C, c > 0$ such that for all n*

$$d_{\text{TV}}(\mathcal{U}_n, \tilde{\mathcal{L}}_n) \leq C \exp(-cn).$$

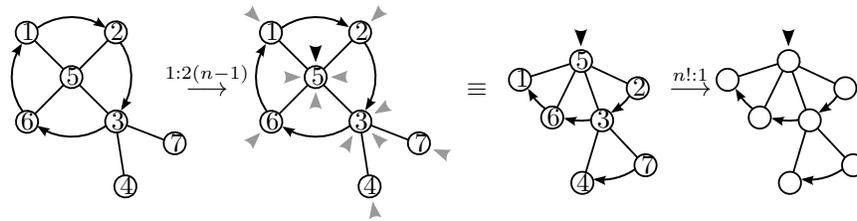


FIGURE 6.2. Given a labelled plane tree with n vertices (first), there are $2(n - 1)$ possibilities to appoint a corner as root (second). The position of the corner induces a natural ordering of the children of the root (third) and, as there exist no symmetries, we find $n!$ corresponding unlabelled trees (fourth).

Any labelled unrooted plane tree of size n has exactly $2(n - 1)$ corners and by determining one corner as root and ordering the descendants in a canonical way we obtain a labelled planted tree. Further, as mentioned before, planted trees do only allow the trivial automorphism, by which we deduce that there are exactly $n!$ possibilities to obtain an unlabelled planted tree, as depicted in Figure 6.2. All in all, this means that

$$\tilde{\mathcal{L}}_n \stackrel{(d)}{=} \tilde{\mathcal{T}}_n^*$$

and Lemma 3.1 follows readily.

Proof of Lemma 6.2: Let $\text{Sym}(L(\mathcal{U})^{\bar{\omega}})$ be the set of all pairs (L, σ) , where L is a labelled unrooted plane tree and σ is an automorphism of L , that is $\sigma.L = L$. For any such object we define its $\bar{\omega}$ -weight by

$$\bar{\omega}((L, \sigma)) := \bar{\omega}(L).$$

Further, the (random) pair $(\mathcal{L}_n, \sigma_n)$ is drawn proportional to its $\bar{\omega}$ -weight from all symmetries of size n in $\text{Sym}(L(\mathcal{U})^{\bar{\omega}})$. For $(L, \sigma) \in \text{Sym}(L(\mathcal{U})^{\bar{\omega}})$ we denote by $(L, \sigma)^\sim$ the tree which is obtained by removing the labelling of L and discarding the automorphism. By the previous notation this means $(L, \sigma)^\sim := \tilde{L}$. It is a well known fact that every unlabelled structure in \mathcal{U}_n induces $n!$ symmetries in $\text{Sym}(L(\mathcal{U})^{\bar{\omega}})$ in the sense that for any $U \in \mathcal{U}_n$ there are $n!$ symmetries (L, σ) such that $(L, \sigma)^\sim = U$, see Lemma 5.5. This implies

$$\mathcal{U}_n \stackrel{(d)}{=} (\mathcal{L}_n, \sigma_n)^\sim.$$

There is a canonical way of rooting any tree at its geometric center by pruning its leaves until there is only one edge or one node left. We call such trees *edge-centered* and *vertex-centered*, respectively. With this comment in mind, there appear three different scenarios for any (L, σ) in $\text{Sym}(L(\mathcal{U})^{\bar{\omega}})$:

- (1) $\sigma = \text{id}$.
- (2) L is vertex-centered and $\sigma \neq \text{id}$.
- (3) L is edge-centered and $\sigma \neq \text{id}$.

Symmetries being part of the first case can equally be seen as regularly labelled objects without an automorphism.

Whenever the tree of a symmetry (L, σ) is vertex-centered and the automorphism is not trivial, there is a unique way of decomposing L , cf. Figure 6.3. Consider the

cyclically ordered list of trees spreading from the center v of L and appoint each of the nodes connected to v by an edge as root of the tree it is contained in. Let us call these planted trees T_1, \dots, T_k . As planted trees do only allow the trivial automorphism and $\sigma \neq \text{id}$ each node in T_1 needs to be transported into some T_ℓ for $\ell \neq 1$ along σ . Hence, in order to maintain the linear ordering within T_1 , *all* of its vertices are sent to T_ℓ and consequently T_ℓ is an identical copy of T_1 . Similarly, all vertices of T_ℓ are mapped into either T_1 or some other identical tree. Continuing this procedure, we receive $d \geq 2$ identical trees, which interchange their entire set of vertices among each other under σ . By repeating the previous steps successively we may eventually partition $\{T_1, \dots, T_k\}$ into $j \in [k]$ subsets of sizes d_1, \dots, d_j , where each subset contains identical trees interchanging their vertex-sets under σ and $d_1 + \dots + d_j = k$. Recall that the roots of the T_1, \dots, T_k are cyclically ordered around the center v , which immediately implies that $d_1 = \dots = d_j$ and $j = k/d_1$ needs to be a divisor of k . Further, the only fixpoint in σ sends the center v to itself.

On the other hand, given the product of a single vertex v and a cyclically ordered list of planted trees T_1, \dots, T_k together with a non-id automorphism $\tilde{\sigma}$ (as before), we may reconstruct a vertex-centered tree by connecting v to the roots of the planted trees. This is due to the fact that the longest path of two identical copies attached to v by two edges is always odd. Then, by extending $\tilde{\sigma}$ with a fixpoint at v we obtain a symmetry of the second case.

This decomposition guarantees that the set containing all symmetries of the second case, denoted by \mathcal{R}_v , can be expressed as

$$\mathcal{R}_v = \{(L, \sigma) : L \in \mathcal{X} \cdot (\text{CYC}^\kappa \circ L(\mathcal{T})^\omega), \sigma.L = L, \sigma \neq \text{id}\},$$

where $\kappa(\tau) = \omega_{k-1}$ for $\tau \in \text{CYC}_k^\kappa$ and \mathcal{X} is the unique class consisting of a single vertex receiving weight 1 accounting for the center v . As described before, we deduce that in this case any automorphism σ contains only *one* single fixpoint (sending the center v to itself) as for otherwise *all* cycles in σ were fixpoints violating the assumption $\sigma \neq \text{id}$. The cycle index sum $Z_{\text{CYC}^\kappa \circ L(\mathcal{T})^\omega}(0, x^2, x^3, \dots)$ counts symmetries of $\text{CYC}^\kappa \circ L(\mathcal{T})^\omega$ without any fixpoints and hence, according to (5.2) and (5.5),

$$\mathcal{R}_v(x) := x Z_{\text{CYC}^\kappa}(Z_{L(\mathcal{T})^\omega}(0, x^2, x^3, \dots), Z_{L(\mathcal{T})^\omega}{}^2(x^2, x^4, x^6, \dots), \dots)$$

counts exactly the symmetries in \mathcal{R}_v . Per definition

$$Z_{L(\mathcal{T})^\omega}(0, s_2, s_3, \dots) = \sum_{k \geq 0} \frac{1}{k!} \sum \omega(T) s_2^{\sigma_2} \dots s_k^{\sigma_k},$$

where the sum is conducted over all $(T, \sigma) \in \text{Sym}(L(\mathcal{T})^\omega)$ such that σ has no fixpoint. On the other hand, labelled planted trees do only allow the trivial automorphism and thus $Z_{L(\mathcal{T})^\omega}(0, x^2, x^3, \dots) = 0$. This shows

$$x^{-1} \mathcal{R}_v(x) = Z_{\text{CYC}^\kappa}(0, \mathcal{T}^{\omega^2}(x^2), \mathcal{T}^{\omega^3}(x^3), \dots) = \sum_{d \geq 2} \frac{\varphi(d)}{d} \sum_{j \geq 1} \frac{\omega_{jd-1}}{j} \left(\mathcal{T}^{\omega^d}(x^d) \right)^j. \tag{6.4}$$

Note that we used $\widetilde{L(\mathcal{T})^{\omega^i}}(x) = \mathcal{T}^{\omega^i}(x)$ in the latter identity.

For the third case consider (L, σ) such that L is edge-centered and $\sigma \neq \text{id}$. By similar reasoning as for the vertex-centered case, L consists of two copies of

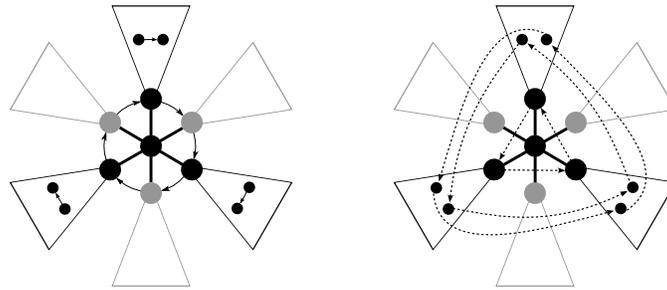


FIGURE 6.3. On the left: a vertex-centered plane tree, where solid arrows indicate the ordering of vertices within the tree; on the right: dashed arrows indicate a non-id permutation which preserves the ordering and sends all vertices of each tree to their respective clones

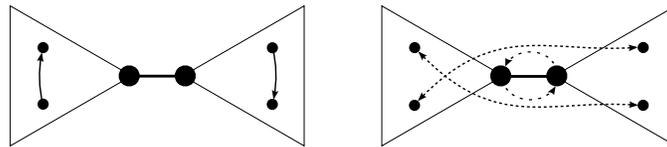


FIGURE 6.4. On the left: an edge-centered plane tree, where solid arrows indicate the ordering of vertices within the tree; on the right: dashed arrows indicate a non-id permutation which preserves the ordering and sends all vertices of each tree to their respective clones

the same planted tree (having half the size of L) connected by an edge and the automorphism σ sends each vertex in one tree to its clone in the other tree, see Figure 6.4. Clearly, the respective generating series is

$$\mathcal{R}_e(x) := \mathcal{F}^{\omega^2}(x^2).$$

Concluding we have that

$$Z_{\mathcal{U}^{\bar{\omega}}}(x, x^2, x^3, \dots) = L(\mathcal{U})^{\bar{\omega}}(x) + \mathcal{R}_v(x) + \mathcal{R}_e(x),$$

where $L(\mathcal{U})^{\bar{\omega}}(x)$ is the exponential generating series of labelled plane trees. Define $\mathcal{R}(x) = \mathcal{R}_v(x) + \mathcal{R}_e(x)$ and assume $n \equiv 2 \pmod{\text{span}(\mathbf{w})}$. Given the random structure (L_n, σ_n) drawn proportional to its $\bar{\omega}$ -weight from $\text{Sym}(L(\mathcal{U})^{\bar{\omega}})$ we first observe

$$\mathbb{P}(\sigma_n \neq \text{id}_n) = \frac{[x^n]\mathcal{R}(x)}{[x^n](L(\mathcal{U})^{\bar{\omega}}(x) + \mathcal{R}(x))} \leq \frac{[x^n]\mathcal{R}(x)}{[x^n]L(\mathcal{U})^{\bar{\omega}}(x)}.$$

Again we make use of the fact that there is an $n! : 2(n-1)$ -relation between labelled plane trees and unlabelled planted trees to obtain

$$[x^n]L(\mathcal{U})^{\bar{\omega}}(x) = \frac{1}{2(n-1)}[x^n]\mathcal{F}^{\bar{\omega}}(x) = \frac{1}{2(n-1)}[x^n]\mathcal{F}^{\omega}(x)^2,$$

where the latter identity is due to Equation (6.3). By Lemma 5.3 the shifted generating series $x^{-1}\mathcal{F}^{\omega}(x)$ is subexponential for $d = \text{span}(\mathbf{w})$. Hence, Lemma 5.2

gives us the existence of a constant $A > 0$ such that

$$[x^n]x^{-2} \mathcal{T}^\omega(x)^2 \geq A \cdot [x^n]x^{-1} \mathcal{T}^\omega(x)$$

for all $n \equiv 0 \pmod d$. We conclude

$$\frac{[x^n]\mathcal{R}(x)}{[x^n]L(\mathcal{U})^{\overline{\omega}}(x)} = 2(n-1) \frac{[x^{n-2}]x^{-2}\mathcal{R}(x)}{[x^{n-2}]x^{-2}\mathcal{T}^\omega(x)} \leq A \cdot 2(n-1) \frac{[x^{n-2}]x^{-1}\mathcal{R}(x)}{[x^{n-2}]x^{-1}\mathcal{T}^\omega(x)}$$

for all $n \equiv 2 \pmod d$. If we assume that the radius of convergence of $x^{-2}\mathcal{R}(x)$ is strictly greater than the one of $x^{-2}\mathcal{T}^\omega(x)$, Lemma 5.4 gives us the existence of $C, c > 0$ such that, for $n \equiv 2 \pmod d$,

$$\frac{[x^{n-2}]x^{-2}\mathcal{R}(x)}{[x^{n-2}]x^{-2}\mathcal{T}^\omega(x)} \leq C \exp(-cn).$$

Consequently, for some $C', c' > 0$,

$$\mathbb{P}(\sigma_n \neq \text{id}_n) \leq A \cdot 2(n-1) \frac{[x^{n-2}]x^{-2}\mathcal{R}(x)}{[x^{n-2}]x^{-2}\mathcal{T}^\omega(x)} \leq C' \exp(-c'n)$$

and Lemma 6.2 is proven.

Hence, to finish the proof, it suffices to show that

$$\rho_{\mathcal{T}^\omega} < \min(\rho_{\mathcal{R}_e}, \rho_{\mathcal{R}_v}). \tag{6.5}$$

Therefore, we follow closely the presentation in Janson (2012, Sec. 7) and apply the related results given there. Define

$$\Phi_d(x) := \sum_{k \geq 0} \omega_k^d x^k$$

together with its radius of convergence ρ_d for $d \geq 2$. Then the recursive structure of planted trees guarantees $\mathcal{T}^{\omega^d}(x) = x\Phi_d(\mathcal{T}^{\omega^d}(x))$. Next, let Ψ and ν be defined as in (4.1). We recall that there appear two cases: if $\nu \geq 1$, let τ be the unique number in $[0, \rho]$ such that $\Psi(\tau) = 1$, and set $\tau := \rho$ otherwise. Likewise, we define ν_d and τ_d with respect to Φ_d . Alternatively, in both cases τ can be characterized as the unique maximum point of $\Phi(t)$ on the interval $[0, \rho]$ and an equivalent characterization holds for τ_d as well, cf. Janson (2012, Remark 7.4). The radii of convergence of $\mathcal{T}^\omega(x)$ and $\mathcal{T}^{\omega^d}(x)$ fulfil

$$\rho_{\mathcal{T}^\omega} = \frac{\tau}{\Phi(\tau)} \quad \text{and} \quad \rho_{\mathcal{T}^{\omega^d}} = \frac{\tau_d}{\Phi_d(\tau_d)}$$

as well as

$$\tau = \mathcal{T}^\omega(\rho_{\mathcal{T}^\omega}) \quad \text{and} \quad \tau_d = \mathcal{T}^{\omega^d}(\rho_{\mathcal{T}^{\omega^d}}), \quad d \geq 2, \tag{6.6}$$

as stated in (6.1) and Janson (2012, Remark 7.5). By standard tilting arguments, we may assume that $(\omega_k)_{k \geq 0}$ is a probability weight sequence (which implies that $\rho \geq 1$) and $\rho_{\mathcal{T}^\omega} = 1$, see Remark 6.1. We know that ρ can be computed by

$$\rho = 1 / \limsup_{k \rightarrow \infty} \omega_k^{1/k}.$$

Hence,

$$\rho_d = 1 / \limsup_{k \rightarrow \infty} \omega_k^{d/k} = \rho^d$$

and $\rho_d \geq 1$ for all $d \geq 2$. Consequently, the choice of τ_d yields

$$\rho_{\mathcal{T}^{\omega^d}} \geq \Phi_d(1)^{-1},$$

giving us that the radius of convergence of $\mathcal{T}^{\omega^d}(x^d)$ satisfies

$$\rho_{\mathcal{T}^{\omega^d}}^{1/d} \geq \Phi_d(1)^{-1/d} = \|\omega\|_d^{-1} \geq \|\omega\|_2^{-1} > 1 = \rho_{\mathcal{T}^\omega}. \quad (6.7)$$

Since $\mathcal{R}_e(x) = \mathcal{T}^{\omega^2}(x^2)$ has radius of convergence $\rho_{\mathcal{T}^{\omega^2}}^{1/2}$, this implies

$$\rho_{\mathcal{T}^\omega} < \rho_{\mathcal{R}_e}. \quad (6.8)$$

In order to complete the verification of Claim (6.5) it remains to show that $\rho_{\mathcal{R}_v} \leq \rho_{\mathcal{R}_e}$. By Inequality (6.7) there exist $\varepsilon > 0$ and $0 < \lambda < 1$ such that

$$\rho_{\mathcal{T}^\omega} + \varepsilon \leq \lambda \rho_{\mathcal{T}^{\omega^d}}^{1/d}$$

and consequently with (6.6)

$$\mathcal{T}^{\omega^d}((\rho_{\mathcal{T}^\omega} + \varepsilon)^d) \leq \mathcal{T}^{\omega^d}(\lambda^d \rho_{\mathcal{T}^{\omega^d}}) \leq \lambda^d \mathcal{T}^{\omega^d}(\rho_{\mathcal{T}^{\omega^d}}) = \lambda^d \tau_d \leq \lambda^d \rho_d = \lambda^d \rho^d$$

uniformly in $d \geq 2$. Thus,

$$\begin{aligned} (\rho_{\mathcal{T}^\omega} + \varepsilon)^{-1} \mathcal{R}_v(\rho_{\mathcal{T}^\omega} + \varepsilon) &\leq \sum_{d \geq 2} \frac{\varphi(d)}{d} \sum_{j \geq 1} \frac{\omega_{jd-1}}{j} \lambda^{jd} \rho^{jd} \\ &\leq \sum_{m \geq 2} \omega_{m-1} \rho^m \frac{\lambda^m}{m} \sum_{jd=m} \varphi(d) \\ &\leq \sum_{m \geq 2} \omega_{m-1} \rho^m m \lambda^m < \infty, \end{aligned}$$

by which it is proven that

$$\rho_{\mathcal{T}^\omega} < \rho_{\mathcal{R}_v}. \quad (6.9)$$

Using Inequality (6.8) it follows that Claim (6.5) holds and the proof is complete. \square

7. Proofs of the applications

7.1. *Scaling limits and tail bounds for the diameter.* We will use the following easy observation, which is inspired by a result for the Gromov–Hausdorff metric in Burago et al. (2001, Thm. 7.3.25).

Lemma 7.1. *Let (X, d_X, μ_X) be a measured compact metric space. Then for all $\beta > 0$ it holds that*

$$d_{\text{GHP}}((X, \mu_X), (\beta X, \mu_X)) \leq |1 - \beta| \text{D}(X) / 2.$$

Proof: In order to avoid confusion, set $(Y, d_Y) = (X, \beta d_X)$. For each $x \in X$ let x' denote the corresponding point in Y . Consider the metric on the disjoint union $X \sqcup Y$ that extends the metrics on X and Y , and satisfies for all $x_1 \in X$ and $x'_2 \in Y$

$$d(x_1, x'_2) = |1 - \beta| \text{D}(X) / 2 + \min(d_X(x_1, x_2), d_Y(x'_1, x'_2)).$$

It is elementary to verify that d satisfies the axioms of a metric, and that

$$d_{\text{P}}(\mu_X, \mu_Y) = d_{\text{H}}(X, Y) = |1 - \beta| \text{D}(X) / 2. \quad \square$$

We are now ready to present the proof for the scaling limit and diameter tail-bounds for the random tree \mathcal{U}_n .

Proof of Theorem 4.1: In the finite variance case, the simply generated planted plane tree \mathcal{T}_m satisfies a scaling limit as in (4.5) by results of Aldous (1993, Thm. 23). (The case with periodic offspring distributions was stated in Le Gall, 2010, Thm 6.1.) It satisfies a tail bound as in (4.6) by results of Addario-Berry et al. (2013, Thm. 1.2).

To verify the tail bound (4.6), let $C_1, c_1 > 0$ be constants such that for all m and $x \geq 0$

$$\mathbb{P}(\mathcal{T}_m \geq x) \leq C_1 \exp(-c_1 x^2/m). \tag{7.1}$$

Lemma 3.1 yields that there are constants $C_2, c_2 > 0$ such that

$$\mathbb{P}(\mathbf{D}(\mathcal{U}_n) \geq x) \leq C_2 \exp(-c_2 n) + \mathbb{P}(\mathbf{D}(\mathcal{T}_n^*) \geq x).$$

If $x > n$, then the left-hand side of this inequality equals zero. If $x \leq n$, then it clearly holds that $\exp(-c_2 n) \leq \exp(-c_2 x^2/n)$. Hence we may write

$$\mathbb{P}(\mathbf{D}(\mathcal{U}_n) \geq x) \leq C_2 \exp(-c_2 x^2/n) + \mathbb{P}(\mathbf{D}(\mathcal{T}_n^*) \geq x). \tag{7.2}$$

The diameter of the tree \mathcal{T}^* may be bounded by twice its height. If \mathcal{T}^* has large height $\text{He}(\mathcal{T}^*) \geq x/2$, then it holds that $\text{He}(\mathcal{T}^{(1)}) \geq x/2 - 1$ or $\text{He}(\mathcal{T}^{(2)}) \geq x/2 - 1$. Both events have the same probability. Setting $\bar{x} = x/2 - 1$, it follows from Lemma 3.2 and Inequality (7.1) that

$$\begin{aligned} \mathbb{P}(\mathbf{D}(\mathcal{T}_n^*) \geq x) &\leq 2\mathbb{P}(\text{He}(\mathcal{T}_n^{(1)}) \geq \bar{x}) \\ &= 2 \sum_{0 \leq k \leq n} \mathbb{P}(|\mathcal{T}_n^{(1)}| = k) \mathbb{P}(\text{He}(\mathcal{T}_k) \geq \bar{x}) \\ &\leq 2C_1 \sum_{0 \leq k \leq n} \mathbb{P}(|\mathcal{T}_n^{(1)}| = k) \exp(-c_1 \bar{x}^2/k) \\ &\leq 2C_1 \exp(-c_1 \bar{x}^2/n). \end{aligned}$$

Hence the bound (4.6) follows from Inequality (7.2).

We proceed with establishing the scaling limit for the random tree \mathcal{U}_n . For each m , let μ_m denote the uniform measure on the leaves of \mathcal{T}_m . By Theorem 2.2 it suffices to show convergence for the random tree \mathcal{S}_n with the uniform measure η_n on its leaves. Clearly $K_n = n + O_p(1)$ implies that

$$\left(\frac{\sigma \mathcal{T}_{K_n}}{2\sqrt{K_n}}, \mu_{K_n} \right) \xrightarrow{d} (\mathcal{T}_{\text{Br}}, \mu_{\text{Br}}) \tag{7.3}$$

in the Gromov–Hausdorff–Prokhorov sense as n becomes large. Hence it suffices to show that

$$d_{\text{GHP}} \left(\left(\frac{\sigma \mathcal{S}_n}{2\sqrt{n}}, \eta_n \right), \left(\frac{\sigma \mathcal{T}_{K_n}}{2\sqrt{n}}, \mu_{K_n} \right) \right) \xrightarrow{p} 0 \tag{7.4}$$

and

$$d_{\text{GHP}} \left(\left(\frac{\sigma \mathcal{T}_{K_n}}{2\sqrt{n}}, \mu_{K_n} \right), \left(\frac{\sigma \mathcal{T}_{K_n}}{2\sqrt{K_n}}, \mu_{K_n} \right) \right) \xrightarrow{p} 0. \tag{7.5}$$

We start with (7.4). Note that the Hausdorff distance $d_{\text{GH}}(\mathcal{S}_n, \mathcal{T}_{K_n})$ between \mathcal{S}_n and \mathcal{T}_{K_n} is bounded by $n - K_n = O_p(1)$. Thus,

$$d_{\text{H}} \left(\frac{\sigma \mathcal{S}_n}{2\sqrt{n}}, \frac{\sigma \mathcal{T}_{K_n}}{2\sqrt{n}} \right) = \frac{\sigma}{2\sqrt{n}} d_{\text{H}}(\mathcal{S}_n, \mathcal{T}_{K_n}) = \frac{\sigma}{2\sqrt{n}} O_p(1) \xrightarrow{d} 0 \tag{7.6}$$

as n becomes large. The Prokhorov distance $d_P(\cdot, \cdot)$ is not homogeneous, so we have to argue differently. For each tree T let $L(T)$ denote its set of leaves. If the tree is rooted, then it is custom to never count the root as a leaf, unless it's the only vertex of the tree. If we regard η_n and μ_{K_n} as measures on the set of vertices of the tree \mathcal{S}_n , then for each subset $A \subset \mathcal{S}_n$

$$\eta_n(A) = |L(\mathcal{S}_n) \cap A|/|L(\mathcal{S}_n)| \quad \text{and} \quad \mu_{K_n}(A) = |L(\mathcal{T}_{K_n}) \cap A|/|L(\mathcal{T}_{K_n})|.$$

The nominator of the first quotient differs from the nominator of the second by at most $n - K_n$, and the same clearly holds for the denominator. It is elementary, that consequently the total variational distance of the random measures η_n and μ_{K_n} may be bounded by

$$d_{\text{TV}}(\eta_n, \mu_{K_n}) \leq 2(n - K_n)/|L(\mathcal{T}_{K_n})|.$$

Janson (2012, Thm. 7.11) showed that

$$L(\mathcal{T}_m)/m \xrightarrow{p} \mathbb{P}(\xi = 0) > 0$$

as m becomes large. As $n - K_n = O_p(1)$, it follows that

$$d_P \left(\left(\frac{\sigma \mathcal{S}_n}{2\sqrt{n}}, \eta_n \right), \left(\frac{\sigma \mathcal{T}_{K_n}}{2\sqrt{K_n}}, \mu_{K_n} \right) \right) \leq d_{\text{TV}}(\eta_n, \mu_{K_n}) \leq \frac{2(n - K_n)}{K_n(\mathbb{P}(\xi = 0) + o_p(1))}.$$

This bound clearly converges in probability to 0 as n becomes large. Together with (7.6), the limit (7.4) readily follows.

In order to verify (7.5), note that by the limit (7.3) it holds that

$$D(\mathcal{T}_{K_n})/\sqrt{K_n} = O_p(1).$$

It follows by Lemma 7.1 that

$$\begin{aligned} d_{\text{GHP}} \left(\left(\frac{\sigma \mathcal{T}_{K_n}}{2\sqrt{n}}, \mu_{K_n} \right), \left(\frac{\sigma \mathcal{T}_{K_n}}{2\sqrt{K_n}}, \mu_{K_n} \right) \right) &\leq \frac{\sigma D(\mathcal{T}_{K_n})}{2} \left(\frac{1}{\sqrt{n}} - \frac{1}{\sqrt{K_n}} \right) \\ &\leq O_p(1) \left(\sqrt{\frac{K_n}{n}} - 1 \right). \end{aligned}$$

Since $K_n = n + O_p(1)$, this bound clearly converges in probability to 0 as n becomes large. Thus (7.5) holds.

In the infinite variance setting of claim 2., the scaling limit (4.7) holds for the simply generated planted plane tree \mathcal{T}_m by results of Duquesne (2003) (see also Kortchemski, 2013 and Haas and Miermont, 2012), and the tail bound (4.8) by Kortchemski (2017). This allows us to deduce the tail bound and convergence for \mathcal{U}_n in an identical way as for the finite variance case. In order to avoid redundancy, we will not make this explicit. \square

7.2. Benjamini-Schramm limits and degree distribution.

Proof of Theorem 4.2: We start with the graph limit. By Theorem 2.2, it suffices to show convergence for the random tree \mathcal{S}_n . Let v_n be a uniformly at random drawn vertex of the tree \mathcal{S}_n , and likewise u_{K_n} a uniformly sampled node of \mathcal{T}_{K_n} . Aldous (1991) showed that

$$(\mathcal{T}_m, u_m) \xrightarrow{d} \mathcal{T}^\bullet$$

with respect to the local metric d_L as m becomes large. Since $K_n = n + O_p(1)$, it follows that also

$$(\mathcal{T}_{K_n}, u_{K_n}) \xrightarrow{d} \mathcal{T}^\bullet.$$

Thus it suffices to verify that

$$d_L((\mathcal{S}_n, v_n), (\mathcal{T}_{K_n}, u_{K_n})) \xrightarrow{p} 0 \quad (7.7)$$

as n tends to infinity. Since the compliment of the tree \mathcal{T}_{K_n} in \mathcal{S}_n has stochastically bounded size, it follows that the random vertex v_n lies in the tree \mathcal{T}_{K_n} with probability tending to 1 as n becomes large. Conditioned on belonging to \mathcal{T}_{K_n} , the vertex v_n is distributed like the uniform node u_{K_n} . For all $\ell \geq 0$ let $Z_\ell(\mathcal{T}_m)$ denote the number of vertices with distance ℓ from the root in \mathcal{T}_m . Janson (2006, Thm. 1.13) showed that for each $r \geq 1$ there is a constant $C > 0$ with

$$\mathbb{E}[Z_\ell(\mathcal{T}_m)^r] \leq C\ell^r$$

for all m and ℓ . In particular for each fixed ℓ it holds that $U_\ell(\mathcal{T}_m)$ is stochastically bounded as m becomes large. Consequently, the same holds for $U_\ell(\mathcal{T}_{K_n})$. Hence the random vertex u_{K_n} lies outside of $U_\ell(\mathcal{T}_{K_n})$ with probability tending to 1 as n becomes large. Whenever this is the case, it follows that

$$U_\ell(\mathcal{S}_n, u_{K_n}) = U_\ell(\mathcal{T}_{K_n}, u_{K_n}).$$

Consequently,

$$d_{\text{TV}}(U_\ell(\mathcal{S}_n, v_n), U_\ell(\mathcal{T}_{K_n}, u_{K_n})) \rightarrow 0$$

as n becomes large. As this holds for arbitrarily large fixed ℓ , the limit (7.7) follows.

It remains to prove the central limit theorem for the number N_d of nodes with degree d in \mathcal{U}_n . For each m , let $\bar{N}_d(\mathcal{T}_m)$ denote the number of nodes with out-degree d in the tree \mathcal{T}_m . Kolchin (1986, Thm. 2.3.1) showed that

$$\frac{\bar{N}_d(\mathcal{T}_m) - \mathbb{P}(\xi = d)m}{\sqrt{m}} \xrightarrow{d} \mathcal{N}(0, \sigma_d^2)$$

as m becomes large for some $\sigma_d > 0$. Since $K_n = n + O_p(1)$, it follows that

$$\frac{\bar{N}_d(\mathcal{T}_{K_n}) - \mathbb{P}(\xi = d)n}{\sqrt{n}} = \frac{\bar{N}_d(\mathcal{T}_{K_n}) - \mathbb{P}(\xi = d)K_n}{\sqrt{K_n}} + o_p(1) \xrightarrow{d} \mathcal{N}(0, \sigma_d^2).$$

As the size of the small tree attached to the root of \mathcal{T}_{K_n} in \mathcal{S}_n is bounded, it follows that

$$|N_d(\mathcal{U}_n) - \bar{N}_{d-1}(\mathcal{T}_{K_n})| = O_p(1).$$

Consequently,

$$\frac{N_d(\mathcal{U}_n) - \mathbb{P}(\xi = d-1)n}{\sqrt{n}} \xrightarrow{d} \mathcal{N}(0, \sigma_{d-1}^2).$$

□

Acknowledgements

We thank the referee for carefully reading our manuscript.

References

- R. Abraham and J.-F. Delmas. Local limits of conditioned Galton-Watson trees: the condensation case. *Electron. J. Probab.* **19**, no. 56, 29 (2014a). [MR3227065](#).
- R. Abraham and J.-F. Delmas. Local limits of conditioned Galton-Watson trees: the infinite spine case. *Electron. J. Probab.* **19**, no. 2, 19 (2014b). [MR3164755](#).
- R. Abraham, J.-F. Delmas and P. Hoscheit. A note on the Gromov-Hausdorff-Prokhorov distance between (locally) compact metric measure spaces. *Electron. J. Probab.* **18**, no. 14, 21 (2013). [MR3035742](#).
- L. Addario-Berry, O. Angel, G. Chapuy, É. Fusy and C. Goldschmidt. Voronoi tessellations in the CRT and continuum random maps of finite excess. In *Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 933–946. SIAM, Philadelphia, PA (2018). [MR3775848](#).
- L. Addario-Berry, L. Devroye and S. Janson. Sub-Gaussian tail bounds for the width and height of conditioned Galton-Watson trees. *Ann. Probab.* **41** (2), 1072–1087 (2013). [MR3077536](#).
- D. Aldous. Asymptotic fringe distributions for general families of random trees. *Ann. Appl. Probab.* **1** (2), 228–266 (1991). [MR1102319](#).
- D. Aldous. The continuum random tree. III. *Ann. Probab.* **21** (1), 248–289 (1993). [MR1207226](#).
- I. Benjamini and O. Schramm. Recurrence of distributional limits of finite planar graphs. *Electron. J. Probab.* **6**, no. 23, 13 pp. (electronic) (2001). [MR1873300](#).
- F. Bergeron, G. Labelle and P. Leroux. *Combinatorial species and tree-like structures*, volume 67 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge (1998). ISBN 0-521-57323-8. [MR1629341](#).
- P. Billingsley. *Weak convergence of measures: Applications in probability*. Society for Industrial and Applied Mathematics, Philadelphia, Pa. (1971). [MR0310933](#).
- M. Bodirsky, É. Fusy, M. Kang and S. Vigerske. Boltzmann samplers, Pólya theory, and cycle pointing. *SIAM J. Comput.* **40** (3), 721–769 (2011). [MR2810913](#).
- D. Burago, Y. Burago and S. Ivanov. *A course in metric geometry*, volume 33 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI (2001). ISBN 0-8218-2129-6. [MR1835418](#).
- J. Chover, P. Ney and S. Wainger. Functions of probability measures. *J. Analyse Math.* **26**, 255–302 (1973). [MR0348393](#).
- N. Curien. Notes provisoires du cours de M2 “graphes aléatoires” (2016). URL <http://www.math.u-psud.fr/~curien/cours-RG-V2.pdf>.
- M. Drmota. *Random trees. An interplay between combinatorics and probability*. SpringerWienNewYork, Vienna (2009). ISBN 978-3-211-75355-2. [MR2484382](#).
- T. Duquesne. A limit theorem for the contour process of conditioned Galton-Watson trees. *Ann. Probab.* **31** (2), 996–1027 (2003). [MR1964956](#).
- P. Embrechts. The asymptotic behaviour of series and power series with positive coefficients. *Med. Konink. Acad. Wetensch. België* **45** (1), 41–61 (1983). [MR714482](#).
- P. Embrechts and E. Omey. Functions of power series. *Yokohama Math. J.* **32** (1-2), 77–88 (1984). [MR772907](#).
- P. Flajolet and R. Sedgewick. *Analytic combinatorics*. Cambridge University Press, Cambridge (2009). ISBN 978-0-521-89806-5. [MR2483235](#).

- S. Foss, D. Korshunov and S. Zachary. *An introduction to heavy-tailed and subexponential distributions*. Springer Series in Operations Research and Financial Engineering. Springer, New York, second edition (2013). ISBN 978-1-4614-7100-4; 978-1-4614-7101-1. [MR3097424](#).
- B. Haas. Scaling limits of Markov-branching trees and applications: lecture notes of the XII Simposio de Probabilidad y Procesos Estocásticos 16–20 Novembre 2015, Mérida, Yucatán. In *XII Symposium of Probability and Stochastic Processes*, volume 73 of *Progr. Probab.*, pages 3–50. Birkhäuser/Springer, Cham (2018). [MR3835788](#).
- B. Haas and G. Miermont. Scaling limits of Markov branching trees with applications to Galton-Watson and random unordered trees. *Ann. Probab.* **40** (6), 2589–2666 (2012). [MR3050512](#).
- S. Janson. Random cutting and records in deterministic and random trees. *Random Structures Algorithms* **29** (2), 139–179 (2006). [MR2245498](#).
- S. Janson. Simply generated trees, conditioned Galton-Watson trees, random allocations and condensation. *Probab. Surv.* **9**, 103–252 (2012). [MR2908619](#).
- V. F. Kolchin. *Random mappings*. Translation Series in Mathematics and Engineering. Optimization Software, Inc., Publications Division, New York (1986). ISBN 0-911575-16-2. [MR865130](#).
- I. Kortchemski. A simple proof of Duquesne’s theorem on contour processes of conditioned Galton-Watson trees. In *Séminaire de Probabilités XLV*, volume 2078 of *Lecture Notes in Math.*, pages 537–558. Springer, Cham (2013). [MR3185928](#).
- I. Kortchemski. Limit theorems for conditioned non-generic Galton-Watson trees. *Ann. Inst. Henri Poincaré Probab. Stat.* **51** (2), 489–511 (2015). [MR3335012](#).
- I. Kortchemski. Sub-exponential tail bounds for conditioned stable Bienaymé-Galton-Watson trees. *Probab. Theory Related Fields* **168** (1-2), 1–40 (2017). [MR3651047](#).
- J.-F. Le Gall. Itô’s excursion theory and random trees. *Stochastic Process. Appl.* **120** (5), 721–749 (2010). [MR2603061](#).
- J.-F. Le Gall and Y. Le Jan. Branching processes in Lévy processes: the exploration process. *Ann. Probab.* **26** (1), 213–252 (1998). [MR1617047](#).
- J.-F. Le Gall and G. Miermont. Scaling limits of random trees and planar maps. In *Probability and statistical physics in two and more dimensions*, volume 15 of *Clay Math. Proc.*, pages 155–211. Amer. Math. Soc., Providence, RI (2012). [MR3025391](#).
- J.-F. Marckert and G. Miermont. The CRT is the scaling limit of unordered binary trees. *Random Structures Algorithms* **38** (4), 467–501 (2011). [MR2829313](#).
- K. Panagiotou and B. Stuffer. Scaling limits of random Pólya trees. *Probab. Theory Related Fields* **170** (3-4), 801–820 (2018). [MR3773800](#).
- B. Stuffer. The continuum random tree is the scaling limit of unlabelled unrooted trees. *ArXiv Mathematics e-prints* (2014). [arXiv: 1412.6333](#).
- B. Stuffer. Limits of random tree-like discrete structures. *ArXiv Mathematics e-prints* (2016). [arXiv: 1612.02580](#).
- B. Stuffer. Gibbs partitions: The convergent case. *Random Structures & Algorithms* **53** (3), 537–558 (2018a). DOI: [10.1002/rsa.20771](#).
- B. Stuffer. Random enriched trees with applications to random graphs. *Electron. J. Combin.* **25** (3), Paper 3.11, 81 (2018b). [MR3853863](#).

- M. Wang. Scaling limits for a family of unrooted trees. *ALEA Lat. Am. J. Probab. Math. Stat.* **13** (2), 1039–1067 (2016). [MR3573447](#).
- H. S. Wilf. *generatingfunctionology*. A K Peters, Ltd., Wellesley, MA, third edition (2006). ISBN 978-1-56881-279-3; 1-56881-279-5. [MR2172781](#).