Fringe trees for random trees with given vertex degrees

Gabriel Berzunza Ojeda
Department of Mathematical Sciences, University of Liverpool, United Kingdom.

Cecilia Holmgren
Department of Mathematics, Uppsala University, Sweden.

Svante Janson
Department of Mathematics, Uppsala University, Sweden.

Abstract

We prove that the number of fringe subtrees, isomorphic to a given tree, in uniformly random trees with given vertex degrees, asymptotically follows a normal distribution. As an application, we establish the same asymptotic normality for random simply generated trees (conditioned Galton-Watson trees). Our approach relies on an extension of Gao and Wormald's (2004) theorem to the multivariate setting.

2012 ACM Subject Classification Mathematics of computing → Probabilistic and statistics

Keywords and phrases Conditioned Galton-Watson trees; fringe trees; simply generated trees; uniformly random trees with given vertex degrees.

Digital Object Identifier 10.4230/LIPIcs.CVIT.2016.23

Related Version *Full Version*: `arXiv:2312.04243` [3]

Funding *Cecilia Holmgren*: Supported by the Knut and Alice Wallenberg Foundation; Ragnar Söderberg's Foundation; the Swedish Research Council.
Svante Janson: Supported by the Knut and Alice Wallenberg Foundation; the Swedish Research Council.

1 Introduction and main results

In this paper, we consider fringe trees of random plane trees with given vertex statistics, i.e., a given number of vertices of each degree. As an application, we also give corresponding result for random simply generated trees (or conditioned Galton–Watson trees). The main results are laws of large numbers and central limit theorems for the number of fringe trees of a given type.

Let $\mathbb{T}$ be the set of all (finite) plane rooted trees (also called ordered rooted trees); see e.g., [9]. Denote the size, i.e. the number of vertices, of a tree T by $|T|$. The (out)degree of a vertex $v \in T$, denoted $d_T(v)$, is its number of children in T; thus leaves have degree 0 and all other vertices have strictly positive degree. The *degree statistic* of a rooted tree T is the sequence $\mathbf{n}_T = (n_T(i))_{i\geq 0}$, where $n_T(i) := |\{v \in T : d_T(v) = i\}|$ is the number of vertices of T with i children. We have

$$|T| = \sum_{i\geq 0} n_T(i) = 1 + \sum_{i\geq 0} i n_T(i). \tag{1}$$

A sequence $\mathbf{n} = (n(i))_{i\geq 0}$ is the degree statistic of some tree if and only if $\sum_{i\geq 0} n(i) = 1 + \sum_{i\geq 0} in(i)$. For such sequences, we let $|\mathbf{n}| := \sum_{i\geq 0} n(i)$ be the size of $\mathbf{n}$, and we write $\mathbb{T}_{\mathbf{n}}$ for the set of plane rooted trees with degree statistic $\mathbf{n}$. We let $\mathcal{T}_{\mathbf{n}}$ be a uniformly random element of the set $\mathbb{T}_{\mathbf{n}}$, and we denote this by $\mathcal{T}_{\mathbf{n}} \sim \mathrm{Unif}(\mathbb{T}_{\mathbf{n}})$. It is also well known that the total number of plane rooted trees with degree statistic $\mathbf{n}$ is given by (see [23, Exercise 6.2.1])

$$|\mathbb{T}_{\mathbf{n}}| = \frac{1}{|\mathbf{n}|}\binom{|\mathbf{n}|}{\mathbf{n}} = \frac{1}{|\mathbf{n}|}\frac{|\mathbf{n}|!}{\prod_{i\geq 0} n(i)!}. \tag{2}$$

For $T \in \mathbb{T}$ and a vertex $v \in T$, let T_v be the subtree of T rooted at v consisting of v and all its descendants. We call T_v a fringe (sub)tree of T. We regard T_v as an element of $\mathbb{T}$ and let, for $T, T' \in \mathbb{T}$,

$$N_{T'}(T) := |\{v \in T : T_v = T'\}| = \sum_{v\in T} \mathbf{1}_{\{T_v = T'\}}, \tag{3}$$

i.e., the number of fringe subtrees of T that are equal (i.e., isomorphic to) to T'. A random fringe subtree T^{fr} of $T \in \mathbb{T}$ is the random rooted tree obtained by taking the fringe subtree T_v at a uniform random vertex $v \in T$. Thus, the distribution of T^{fr} is given by

$$\mathbb{P}(T^{\text{fr}} = T') = \frac{N_{T'}(T)}{|T|}, \quad \text{for } T' \in \mathbb{T}. \tag{4}$$

We prove an asymptotic result on the distribution of a random fringe subtree in a random rooted plane tree with a given degree statistic. In order to state the theorem, we need a little more terminology. (See also Section 1.2 for some notation.) For a degree statistic $\mathbf{n}$, denote by $\mathbf{p}(\mathbf{n}) = (p_i(\mathbf{n}))_{i\geq 0}$ its (empirical) degree distribution, i.e.,

$$p_i(\mathbf{n}) := \frac{n(i)}{|\mathbf{n}|}, \quad \text{for } i \geq 0. \tag{5}$$

In this paper, we assume for convenience the following condition.

▶ **Condition 1.** *$\mathbf{n}_\kappa = (n_\kappa(i))_{i\geq 0}$, $\kappa \geq 1$, are degree statistics such that as $\kappa \to \infty$:*

(i) $|\mathbf{n}_\kappa| \to \infty$,

(ii) *For every $i \geq 0$, we have $p_i(\mathbf{n}_\kappa) \to p_i$, where $\mathbf{p} = (p_i)_{i\geq 0}$ is a probability distribution on $\mathbb{N}_0$.*

▶ Remark 2. The condition that $\mathbf{p}$ is a probability distribution is no restriction. In fact, the degree distribution $\mathbf{p}(\mathbf{n}_\kappa)$ has mean

$$\sum_{i\geq 0} i p_i(\mathbf{n}_\kappa) = \frac{1}{|\mathbf{n}_\kappa|}\sum_{i\geq 0} i n_\kappa(i) = \frac{|\mathbf{n}_\kappa| - 1}{|\mathbf{n}_\kappa|} < 1, \tag{6}$$

and thus the sequence of distributions $\mathbf{p}(\mathbf{n}_\kappa)$ is always tight. Hence, if $p_i(\mathbf{n}_\kappa) \to p_i$, for every $i \geq 0$, then $\mathbf{p} = (p_i)_{i\geq 0}$ is a probability distribution. Note also that (ii) says that $\mathbf{p}(\mathbf{n}_\kappa)$ converges weakly to $\mathbf{p}$, as $\kappa \to \infty$. (As is well known, this is equivalent to convergence in total variation.)

By (6) and Fatou's lemma, if Condition 1 holds, then $\sum_{i\geq 0} i p_i \leq 1$. Conversely, it is easily seen that any such probability distribution $\mathbf{p}$ is the limit of $\mathbf{p}(\mathbf{n}_\kappa)$ for some sequence of degree statistics $\mathbf{n}_\kappa$. In other words, the set of probability distributions $\mathbf{p}$ that can appear as limits in Condition 1 is precisely the set of probability distributions $\mathbf{p}$ on $\mathbb{N}_0$ with mean $\sum_{i\geq 0} i p_i \leq 1$; we denote this set by $\mathcal{P}_1(\mathbb{N}_0)$.

For a probability distribution $\mathbf{p} = (p_i)_{i\geq 0} \in \mathcal{P}_1(\mathbb{N}_0)$, let $\mathcal{T}_{\mathbf{p}}$ be a Galton–Watson tree with offspring distribution $\mathbf{p}$, and define $\pi_{\mathbf{p}}$ as the distribution of $\mathcal{T}_{\mathbf{p}}$, i.e., (with $0^0 := 1$ as usual)

$$\pi_{\mathbf{p}}(T) := \mathbb{P}(\mathcal{T}_{\mathbf{p}} = T) = \prod_{i\geq 0} p_i^{n_T(i)} = \prod_{i\in\mathcal{D}(T)} p_i^{n_T(i)}, \qquad \text{for } T \in \mathbb{T}, \tag{7}$$

where

$$\mathcal{D}(T) := \{i : n_T(i) > 0\} = \{d_T(v) : v \in T\}, \tag{8}$$

the set of degrees that appear in T. Note that $\pi_{\mathbf{p}}(T) = 0 \iff p_i = 0$ for some $i \in \mathcal{D}(T)$. In particular, if $\mathbf{n}_\kappa$ and $\mathbf{p}$ are as in Condition 1, then $\pi_{\mathbf{p}}(T) = 0$ if and only if $n_\kappa(i) = o(|\mathbf{n}_\kappa|)$ for some $i \in \mathcal{D}(T)$.

We first give a law of large numbers for the number of fringe trees of a given type in a random rooted plane tree with a given degree statistic. The proofs of this and the following theorem are given in later sections.

▶ **Theorem 3.** *Let $\mathbf{n}_\kappa$, $\kappa \geq 1$, be some degree statistics that satisfy Condition 1, and let $\mathcal{T}_{\mathbf{n}_\kappa} \sim \mathrm{Unif}(\mathbb{T}_{\mathbf{n}_\kappa})$. For every fixed $T \in \mathbb{T}$, as $\kappa \to \infty$:*

(i) *(Annealed version)* $\mathbb{P}(\mathcal{T}^{\mathrm{fr}}_{\mathbf{n}_\kappa} = T) = \dfrac{\mathbb{E}[N_T(\mathcal{T}_{\mathbf{n}_\kappa})]}{|\mathbf{n}_\kappa|} \to \pi_{\mathbf{p}}(T)$.

(ii) *(Quenched version)* $\mathbb{P}(\mathcal{T}^{\mathrm{fr}}_{\mathbf{n}_\kappa} = T \mid \mathcal{T}_{\mathbf{n}_\kappa}) = \dfrac{N_T(\mathcal{T}_{\mathbf{n}_\kappa})}{|\mathbf{n}_\kappa|} \to \pi_{\mathbf{p}}(T)$ *in probability.*

In other words, the random fringe tree converges in distribution as $\kappa \to \infty$: (i) *says $\mathcal{T}^{\mathrm{fr}}_{\mathbf{n}_\kappa} \xrightarrow{\mathrm{d}} \mathcal{T}_{\mathbf{p}}$, or equivalently $\mathcal{L}(\mathcal{T}^{\mathrm{fr}}_{\mathbf{n}_\kappa}) \to \mathcal{L}(\mathcal{T}_{\mathbf{p}})$, and* (ii) *is the conditional version $\mathcal{L}\big(\mathcal{T}^{\mathrm{fr}}_{\mathbf{n}_\kappa} \mid \mathcal{T}_{\mathbf{n}_\kappa}\big) \xrightarrow{\mathrm{p}} \mathcal{L}(\mathcal{T}_{\mathbf{p}})$.*

▶ Remark 4. Similar results are known for several other models of random trees. In particular, a version of Theorem 3 was proved by Aldous [2] for conditioned Galton–Watson trees with finite offspring variance; this was extended to general simply generated trees in [19, Theorem 7.12]. In those cases, the degree statistic is random, but Condition 1 holds in probability, with a non-random limiting probability distribution $\mathbf{p}$. We return to simply generated trees in Section 5. Another standard example is family trees of Crump–Mode–Jagers branching processes (which includes e.g. random recursive trees, binary search trees and preferential attachment trees); see e.g. [2] and [17, Theorem 5.14].

Theorem 3 is thus a law of large numbers for the number of fringe trees of a given type. In this work, we also study the fluctuations and prove a central limit theorem for this number; we furthermore show that this holds jointly for different types of fringe trees.

For a probability distribution $\mathbf{p} = (p_i)_{i\geq 0} \in \mathcal{P}_1(\mathbb{N}_0)$ and $T, T' \in \mathbb{T}$, let

$$\eta_{\mathbf{p}}(T, T') := (|T| - 1)(|T'| - 1) - \sum_{i\geq 0} \frac{n_T(i) n_{T'}(i)}{p_i}, \tag{9}$$

where we interpret $0/0 := 0$, and, for $T \neq T'$,

$$\gamma_{\mathbf{p}}(T, T) := \pi_{\mathbf{p}}(T) + \eta_{\mathbf{p}}(T, T)(\pi_{\mathbf{p}}(T))^2, \tag{10}$$

$$\gamma_{\mathbf{p}}(T, T') := N_{T'}(T)\pi_{\mathbf{p}}(T) + N_T(T')\pi_{\mathbf{p}}(T') + \eta_{\mathbf{p}}(T, T')\pi_{\mathbf{p}}(T)\pi_{\mathbf{p}}(T'). \tag{11}$$

Note that $\eta_{\mathbf{p}}(T, T') = -\infty$ if $p_i = 0$ for some $i \in \mathcal{D}(T) \cap \mathcal{D}(T')$. In this case, $\pi_{\mathbf{p}}(T) = \pi_{\mathbf{p}}(T') = 0$, and we interpret $\infty \cdot 0 := 0$ in (10)–(11); thus $\gamma_{\mathbf{p}}(T, T')$ is always finite.

▶ **Theorem 5.** *Let $\mathbf{n}_\kappa$, $\kappa \geq 1$, be some degree statistics that satisfy Condition 1 and let $\mathcal{T}_{\mathbf{n}_\kappa} \sim \mathrm{Unif}(\mathbb{T}_{\mathbf{n}_\kappa})$. For a fixed $m \geq 1$, let $T_1, \ldots, T_m \in \mathbb{T}$ be a fixed sequence of rooted plane trees. Then, as $\kappa \to \infty$,*

$$\mathbb{E}\, N_{T_i}(\mathcal{T}_{\mathbf{n}_\kappa}) = \pi_{\mathbf{p}}(T_i)|\mathbf{n}_\kappa| + o(|\mathbf{n}_\kappa|), \tag{12}$$

$$\mathrm{Var}(N_{T_i}(\mathcal{T}_{\mathbf{n}_\kappa})) = \gamma_{\mathbf{p}}(T_i, T_i)|\mathbf{n}_\kappa| + o(|\mathbf{n}_\kappa|), \tag{13}$$

$$\mathrm{Cov}\big(N_{T_i}(\mathcal{T}_{\mathbf{n}_\kappa}), N_{T_j}(\mathcal{T}_{\mathbf{n}_\kappa})\big) = \gamma_{\mathbf{p}}(T_i, T_j)|\mathbf{n}_\kappa| + o(|\mathbf{n}_\kappa|), \tag{14}$$

for $1 \le i, j \le m$*, and*

$$\left(\frac{N_{T_j}(\mathcal{T}_{\mathbf{n}_\kappa}) - \mathbb{E}[N_{T_j}(\mathcal{T}_{\mathbf{n}_\kappa})]}{\sqrt{|\mathbf{n}_\kappa|}}\right)_{j=1}^{m} \xrightarrow{\mathrm{d}} \mathrm{N}(0, \Gamma_{\mathbf{p}}), \tag{15}$$

where the covariance matrix is defined by $\Gamma_{\mathbf{p}} := (\gamma_{\mathbf{p}}(T_i, T_j))_{i,j=1}^{m}$*. Furthermore, in* (15)*, we can replace* $\mathbb{E}[N_{T_i}(\mathcal{T}_{\mathbf{n}_\kappa})]$ *by* $|\mathbf{n}_\kappa| \pi_{\mathbf{p}(\mathbf{n}_\kappa)}(T_i)$*.*

If $T \in \mathbb{T}$ *with* $\pi_{\mathbf{p}}(T) > 0$ *and* $|T| > 1$*, then* $\gamma_{\mathbf{p}}(T, T) > 0$ *and thus* (13) *and* (15) *(with* $m = 1$*) show that* $N_T(\mathcal{T}_{\mathbf{n}_\kappa})$ *is asymptotically normal, with*

$$\frac{N_T(\mathcal{T}_{\mathbf{n}_\kappa}) - \mathbb{E}[N_T(\mathcal{T}_{\mathbf{n}_\kappa})]}{\sqrt{\mathrm{Var}(N_T(\mathcal{T}_{\mathbf{n}_\kappa}))}} \xrightarrow{\mathrm{d}} \mathrm{N}(0, 1), \qquad \kappa \to \infty. \tag{16}$$

The case $|T| = 1$ is trivial, with $N_T(\mathcal{T}_{\mathbf{n}_\kappa}) = n_\kappa(0)$ non-random. Ignoring this case, Theorem 5 shows that $N_T(\mathcal{T}_{\mathbf{n}_\kappa})$ is asymptotically normal when $\pi_{\mathbf{p}}(T) > 0$. On the other hand, if $\pi_{\mathbf{p}}(T) = 0$, then also $\gamma_{\mathbf{p}}(T, T) = 0$, and the theorems above do not give precise information on the asymptotic distribution of $N_T(\mathcal{T}_{\mathbf{n}_\kappa})$. In this case, [3, Theorem 1.7] in the full version is more precise.

In the case of critical conditioned Galton–Watson trees with finite offspring variance, (joint) normal convergence of the subtree counts in analogy to (15) was proved in [20, Corollary 1.8] (together with convergence of mean and variance). Indeed, [20, Theorem 1.5] proved, more generally, asymptotic normality of additive functionals that are defined via toll functions (under some conditions); see [3, Section 8] in the full version for further discussion on additive functionals.

▶ Remark 6. Results on asymptotic normality for fringe tree counts have also been proved earlier for several other classes of random trees. For example, for binary search trees see [7], [8], [6], [12], [16]; for random recursive trees see [11], [16]; for increasing trees see [13]; for m-ary search trees and preferential attachment trees see [18]; for random tries see [21].

Our approach relies on a multivariate version of the Gao–Wormald theorem [14, Theorem 1]; see [3, Theorem A.1]. The original Gao–Wormald theorem [14] provides a way to show asymptotic normality by analysing the behaviour of sufficiently high factorial moments. (Typically, factorial moments are more convenient than standard moments in combinatorics.) The multivariate version [3, Theorem A.1] extends this by considering joint factorial moments. In our framework, this is very convenient since we can precisely compute the joint factorial moments of the subtree counts in (3) for random trees with given degree statistics. (Another, closely related, multivariate version of the Gao–Wormald theorem has independently been shown recently by Hitczenko and Wormald [15].)

The (one dimensional) Gao–Wormald theorem has been used before by Cai and Devroye [5] to study large fringe trees in critical conditioned Galton–Watson trees with finite offspring variance. Indeed, they considered fringe subtree counts of a sequence of trees instead of a fixed tree. In particular, they showed that asymptotic normality still holds in some regimes, while in others there is a Poisson limit. In a forthcoming work, we will study the case of not fixed fringe trees in the framework of random trees with given degrees.

1.1 Organization of the paper

In Section 2 we provide exact formulas for factorial moments of $N_T(\mathcal{T}_{\mathbf{n}})$. These formulas are then used in Sections 3–4 to prove our main results. An application to simply generated trees is given in Sections 5.

1.2 Some notation

In addition to the notation introduced above, we use the following standard notation.

We let $\mathbb{Z} := \{\dots,-1,0,1,\dots\}$, $\mathbb{N} := \{1,2,\dots\}$, $\mathbb{N}_0 := \{0,1,2,\dots\}$. We let 0 denote also vectors and matrices with all elements 0 (the dimension will be clear from the context). We use standard o and O notation, for sequences and functions of a real variable.

$\mathbf{1}_{\mathcal{E}}$ is the indicator function of an event $\mathcal{E}$, and $\delta_{ij} := \mathbf{1}_{\{i=j\}}$ is Kronecker's delta.

For $x \in \mathbb{R}$ and $q \in \mathbb{N}_0$, we let $(x)_q := x(x-1)\cdots(x-q+1)$ denote the qth falling factorial of x. (Here $(x)_0 := 1$. Note that $(x)_q = 0$ whenever $x \in \mathbb{N}_0$ and $x-q+1 \le 0$.)

We interpret $0/0 = 0$ and $0 \cdot \infty = 0$.

We use $\xrightarrow{\mathrm{d}}$ for convergence in distribution, and $\xrightarrow{\mathrm{p}}$ for convergence in probability, for a sequence of random variables in some metric space. Also, $\mathcal{L}(X)$ denotes the distribution of X, and $\overset{\mathrm{d}}{=}$ means equal in distribution. We write $\mathrm{N}(0,\Gamma)$ for the multivariate normal distribution with mean vector 0 and covariance matrix $\Gamma := (\gamma_{ij})_{i,j=1}^m$, for $m \in \mathbb{N}$. (This includes the case $\Gamma = 0$; in this case $X \sim \mathrm{N}(0,\Gamma)$ means that $X = 0 \in \mathbb{R}^m$ a.s.)

Unspecified limits are as $\kappa \to \infty$.

2 Moment computations

In this section, we compute the joint factorial moments of $N_{T_1}(\mathcal{T}_{\mathbf{n}}),\dots,N_{T_m}(\mathcal{T}_{\mathbf{n}})$, for $m \ge 1$ and a sequence of distinct rooted plane trees $T_1,\dots,T_m \in \mathbb{T}$, where $\mathcal{T}_{\mathbf{n}}$ is a uniformly random tree of $\mathbb{T}_{\mathbf{n}}$, for a degree statistic $\mathbf{n}$. Before that, we need to introduce some notation. For $1 \le i,j \le m$, let

$$\tau_{ij} := N_{T_i}(T_j)\mathbf{1}_{\{i\neq j\}} \tag{17}$$

be the number of proper fringe subtrees of T_j that are equal to T_i. (Note that many of these terms are 0. In particular, if we order $T_1,\dots,T_m$ according to their sizes, the matrix $(\tau_{ij})_{i,j=1}^m$ is strictly triangular.)

For $q_1,\dots,q_m \in \mathbb{N}_0$, note that the product $(N_{T_1}(\mathcal{T}_{\mathbf{n}}))_{q_1}\cdots(N_{T_m}(\mathcal{T}_{\mathbf{n}}))_{q_m}$ is the number of sequences of $q := q_1+\dots+q_m$ distinct fringe subtrees of $\mathcal{T}_{\mathbf{n}}$, where the first q_1 are copies of T_1, the next q_2 are copies of T_2, and so on. Given such a sequence of fringe subtrees, we say that these fringe subtrees are *marked*. Furthermore, for each such sequence of marked fringe subtrees of $\mathcal{T}_{\mathbf{n}}$, say that a tree in the sequence is *bound* if it is a fringe subtree of another tree in the sequence; otherwise it is *free*. Note that the free trees are disjoint. Furthermore, each bound tree in the sequence is a fringe subtree of exactly one free tree. For a sequence $b = (b_1,\dots,b_m) \in \mathbb{N}_0^m$, let $S_b(\mathcal{T}_{\mathbf{n}})$ be the number of such sequences of q fringe trees such that exactly b_i of the fringe trees T_i are bound, for $1 \le i \le m$. We thus have

$$\mathbb{E}\big[(N_{T_1}(\mathcal{T}_{\mathbf{n}}))_{q_1}\cdots(N_{T_m}(\mathcal{T}_{\mathbf{n}}))_{q_m}\big] = \sum_{b\in\mathbb{N}_0^m}\mathbb{E}[S_b(\mathcal{T}_{\mathbf{n}})]. \tag{18}$$

The sum is really only over $b = (b_1,\dots,b_m) \in \mathbb{N}_0^m$ such that $0 \le b_i \le q_i$ for $1 \le i \le m$, since otherwise $S_b(\mathcal{T}_{\mathbf{n}}) = 0$. This sum can be computed by the following lemma.

▶ **Lemma 7.** *Let $\mathbf{n}$ be a degree statistic and let $\mathcal{T}_{\mathbf{n}} \sim \mathrm{Unif}(\mathbb{T}_{\mathbf{n}})$. For $m \ge 1$ and $q_1,\dots,q_m \in \mathbb{N}$, let $T_1,\dots,T_m \in \mathbb{T}$ be a sequence of distinct rooted plane trees such that $|\mathbf{n}| \ge \sum_{j=1}^m (q_j - b_j)(|T_j|-1)+1$. Then $\mathbb{E}[S_b(\mathcal{T}_{\mathbf{n}})]$ is equal to*

$$\frac{|\mathbf{n}|}{(|\mathbf{n}|)_{1+\sum_{j=1}^m (q_j-b_j)(|T_j|-1)}}\prod_{i\ge 0}(n(i))_{\sum_{j=1}^m (q_j-b_j)n_{T_j}(i)}\prod_{j=1}^m \frac{(q_j)_{b_j}\left(\sum_{k=1}^m (q_k-b_k)\tau_{jk}\right)_{b_j}}{b_j!}, \tag{19}$$

for every $b = (b_1, \dots, b_m) \in \mathbb{N}_0^m$ *such that* $0 \le b_i \le q_i$*, for* $1 \le i \le m$.

Proof. If $\sum_{j=1}^m (q_j - b_j) n_{T_j}(i) > n(i)$ for some $i \ge 0$, then both $\mathbb{E}[S_b(\mathcal{T}_\mathbf{n})]$ and (19) are 0. We may thus assume that $\sum_{j=1}^m (q_j - b_j) n_{T_j}(i) \le n(i)$ for all $i \ge 0$.

First, let us consider the case when all fringe trees are free, that is, the case $b = 0 = (0, \dots, 0) \in \mathbb{N}_0^m$. Replace each marked fringe subtree in $\mathcal{T}_\mathbf{n}$ by a single leaf; moreover, mark this leaf and order all marked leaves into a sequence, corresponding to the order of the fringe subtrees. This yields another tree $\widetilde{\mathcal{T}}$, which we call a *reduced tree*, with a sequence of q marked leaves. Since $\mathcal{T}_\mathbf{n}$ has $n(i)$ vertices of degree i, for $i \ge 0$, and we have replaced q_j copies of T_j by leaves, the degree statistic $\tilde{\mathbf{n}} = (\tilde{n}(i))_{i \ge 0}$ of $\widetilde{\mathcal{T}}$ is given by

$$\tilde{n}(i) \coloneqq \begin{cases} n(i) - \sum_{j=1}^m q_j n_{T_j}(i), & i \ge 1, \\ n(0) - \sum_{j=1}^m q_j n_{T_j}(0) + \sum_{j=1}^m q_j, & i = 0, \end{cases} \tag{20}$$

and has size

$$|\tilde{\mathbf{n}}| \coloneqq \sum_{i \ge 0} \tilde{n}(i) = |\mathbf{n}| - \sum_{j=1}^m q_j (|T_j| - 1). \tag{21}$$

There is a one-to-one correspondence between trees in $\mathbb{T}_\mathbf{n}$ with a sequence of marked fringe subtrees as above, and reduced trees with the degree statistic (20) and a sequence of q marked leaves. If we ignore the marks, the number of possible reduced trees is given by (2) with the degree statistic $\tilde{\mathbf{n}}$ in (20). In each unmarked reduced tree, the number of ways to choose sequences of marked leaves is $(\tilde{n}(0))_{q_1 + \dots + q_m}$. Thus, the number of trees in $\mathbb{T}_\mathbf{n}$ with marked sequences of free fringe subtrees is the product of these numbers, i.e.,

$$\frac{(|\tilde{\mathbf{n}}| - 1)!}{\prod_{i \ge 0} \tilde{n}(i)!} (\tilde{n}(0))_{\sum_{j=1}^m q_j} = \frac{(|\tilde{\mathbf{n}}| - 1)!}{\prod_{i \ge 0} (n(i) - \sum_{j=1}^m q_j n_{T_j}(i))!}. \tag{22}$$

By dividing with $|\mathbb{T}_\mathbf{n}|$, which is given by (2), and using (21), we find

$$\mathbb{E}[S_0(\mathcal{T}_\mathbf{n})] = \frac{1}{(|\mathbf{n}| - 1)_{\sum_{j=1}^m q_j (|T_j| - 1)}} \prod_{i \ge 0} (n(i))_{\sum_{j=1}^m q_j n_{T_j}(i)}. \tag{23}$$

Now consider the general case with a sequence $b = (b_1, \dots, b_m)$ telling the number of bound fringe subtrees. There are thus $q_j - b_j$ free trees of type T_j. The number of ways to choose the positions of the bound trees in the sequences of fringe trees is $\prod_{j=1}^m \binom{q_j}{b_j}$, and for each choice of free trees, there are $\sum_{k=1}^m (q_k - b_k) \tau_{jk}$ possible bound trees of type T_j; thus the number of choices of the bound trees is

$$\prod_{j=1}^m \frac{(q_j)_{b_j} \left(\sum_{k=1}^m (q_k - b_k) \tau_{jk}\right)_{b_j}}{b_j!}. \tag{24}$$

The number of trees in $\mathbb{T}_\mathbf{n}$ with sequences of $q_j - b_j$ free trees T_j, for $1 \le j \le m$, is given by replacing q_j by $q_j - b_j$ in (20)–(22). Hence, we obtain (19), extending (23). ◀

We record two important special cases of Lemma 7 (see the proof of [3, Lemma 3.3] in the full version for details).

▶ **Lemma 8.** *Let* $\mathbf{n}$ *be a degree statistic and let* $\mathcal{T}_\mathbf{n} \sim \mathrm{Unif}(\mathbb{T}_\mathbf{n})$.

(i) *For* $q \in \mathbb{N}$ *and* $T \in \mathbb{T}$ *such that* $|\mathbf{n}| \ge q|T| - q + 1$,

$$\mathbb{E}[(N_T(\mathcal{T}_\mathbf{n}))_q] = \frac{|\mathbf{n}|}{(|\mathbf{n}|)_{q|T| - q + 1}} \prod_{i \ge 0} (n(i))_{q n_T(i)}. \tag{25}$$

(ii) *For distinct* $T, T' \in \mathbb{T}$ *such that* $|\mathbf{n}| \geq |T| + |T'| - 1$,

$$\mathbb{E}[N_T(\mathcal{T}_\mathbf{n})N_{T'}(\mathcal{T}_\mathbf{n})] = N_T(T')\,\mathbb{E}[N_{T'}(\mathcal{T}_\mathbf{n})] + N_{T'}(T)\,\mathbb{E}[N_T(\mathcal{T}_\mathbf{n})] + \frac{|\mathbf{n}|}{(|\mathbf{n}|)_{|T|+|T'|-1}} \prod_{i\geq 0} (n(i))_{n_T(i)+n_{T'}(i)}. \tag{26}$$

3 Proof of Theorems 3

In this section we prove Theorem 3. In what follows we will frequently use the following well-known estimate (see for example, [3, Lemma 4.1]).

▶ **Lemma 9.** *If* $x \geq 1$ *is a real number and* $0 \leq k \leq x/2$ *is an integer, then*

$$(x)_k = x^k \exp\left(-\frac{k(k-1)}{2x} + O\left(\frac{k^3}{x^2}\right)\right). \tag{27}$$

We start by proving the following theorem.

▶ **Theorem 10.** *Let* $T \in \mathbb{T}$ *be a fixed tree. Then, uniformly for all degree statistics* $\mathbf{n} = (n(i))_{i\geq 0}$,

$$\mathbb{E}\,N_T(\mathcal{T}_\mathbf{n}) = |\mathbf{n}|\pi_{\mathbf{p}(\mathbf{n})}(T) + O(1), \tag{28}$$

$$\operatorname{Var} N_T(\mathcal{T}_\mathbf{n}) = |\mathbf{n}|\gamma_{\mathbf{p}(\mathbf{n})}(T,T) + O(1). \tag{29}$$

More generally, if $T, T' \in \mathbb{T}$, *then*

$$\operatorname{Cov}\big(N_T(\mathcal{T}_\mathbf{n}), N_{T'}(\mathcal{T}_\mathbf{n})\big) = |\mathbf{n}|\gamma_{\mathbf{p}(\mathbf{n})}(T,T') + O(1). \tag{30}$$

Proof. Note first the trivial bound

$$N_T(\mathcal{T}_\mathbf{n}) \leq \frac{n(i)}{n_T(i)} \leq n(i), \qquad i \in \mathcal{D}(T), \tag{31}$$

since the copies of T in $\mathcal{T}_\mathbf{n}$ are distinct. Furthermore, by (7) and (5),

$$|\mathbf{n}|\pi_{\mathbf{p}(\mathbf{n})}(T) \leq |\mathbf{n}|p_i(\mathbf{n}) = n(i), \qquad i \in \mathcal{D}(T). \tag{32}$$

Hence, (28) is trivial if $n(i) = O(1)$ for some $i \in \mathcal{D}(T)$. In particular, we may in the sequel assume $n(i) \geq 2n_T(i)$ for every $i \geq 0$, and thus $|\mathbf{n}| \geq 2|T|$. Then, by (25) (with $q = 1$) and Lemma 9,

$$\begin{aligned}
\mathbb{E}\,N_T(\mathcal{T}_\mathbf{n}) &= |\mathbf{n}|^{1-|T|} \prod_{i\in\mathcal{D}(T)} n(i)^{n_T(i)} \\
&\qquad \times \exp\left(\frac{|T|(|T|-1)}{2|\mathbf{n}|} - \sum_{i\in\mathcal{D}(T)} \frac{n_T(i)(n_T(i)-1)}{2n(i)} + O\Big(\sum_{i\in\mathcal{D}(T)} \frac{1}{n(i)^2}\Big)\right) \\
&= |\mathbf{n}|\pi_{\mathbf{p}(\mathbf{n})}(T) \\
&\qquad \times \exp\left(\frac{|T|(|T|-1)}{2|\mathbf{n}|} - \sum_{i\in\mathcal{D}(T)} \frac{n_T(i)(n_T(i)-1)}{2n(i)} + O\Big(\sum_{i\in\mathcal{D}(T)} \frac{1}{n(i)^2}\Big)\right),
\end{aligned} \tag{33}$$

which implies (28) by (32).

Similarly, taking $q = 2$ in (25), and now assuming as we may $n(i) \geq 4n_T(i)$ for every $i \geq 0$,

$$
\begin{aligned}
\mathbb{E}(N_T(\mathcal{T}_{\mathbf{n}}))_2 &= \frac{|\mathbf{n}|}{(|\mathbf{n}|)_{2|T|-1}} \prod_{i\in\mathcal{D}(T)} (n(i))_{2n_T(i)} \\
&= |\mathbf{n}|^{2-2|T|} \prod_{i\in\mathcal{D}(T)} n(i)^{2n_T(i)} \\
&\qquad \times \exp\left(\frac{(2|T|-1)(2|T|-2)}{2|\mathbf{n}|} - \sum_{i\in\mathcal{D}(T)} \frac{2n_T(i)(2n_T(i)-1)}{2n(i)} + O\Big(\sum_{i\in\mathcal{D}(T)} \frac{1}{n(i)^2}\Big)\right) \\
&= \big(|\mathbf{n}|\pi_{\mathbf{p}(\mathbf{n})}(T)\big)^2 \\
&\qquad \times \exp\left(\frac{(2|T|-1)(|T|-1)}{|\mathbf{n}|} - \sum_{i\in\mathcal{D}(T)} \frac{n_T(i)(2n_T(i)-1)}{n(i)} + O\Big(\sum_{i\in\mathcal{D}(T)} \frac{1}{n(i)^2}\Big)\right),
\end{aligned}
\tag{34}
$$

Hence, using also (33),

$$
\mathbb{E}(N_T(\mathcal{T}_{\mathbf{n}}))_2 = \big(\mathbb{E}\, N_T(\mathcal{T}_{\mathbf{n}})\big)^2 \times \exp\left(\frac{(|T|-1)^2}{|\mathbf{n}|} - \sum_{i\in\mathcal{D}(T)} \frac{n_T(i)^2}{n(i)} + O\Big(\sum_{i\in\mathcal{D}(T)} \frac{1}{n(i)^2}\Big)\right). \tag{35}
$$

Consequently, using (28) and noting that $\mathbb{E}\, N_T(\mathcal{T}_{\mathbf{n}}) = O(n(i))$ for $i \in \mathcal{D}(T)$ by (28) and (32),

$$
\begin{aligned}
\operatorname{Var}[N_T(\mathcal{T}_{\mathbf{n}})] &= \mathbb{E}(N_T(\mathcal{T}_{\mathbf{n}}))_2 + \mathbb{E}\, N_T(\mathcal{T}_{\mathbf{n}}) - \big(\mathbb{E}\, N_T(\mathcal{T}_{\mathbf{n}})\big)^2 \\
&= \big(\mathbb{E}\, N_T(\mathcal{T}_{\mathbf{n}})\big)^2 \left(\frac{(|T|-1)^2}{|\mathbf{n}|} - \sum_{i\in\mathcal{D}(T)} \frac{n_T(i)^2}{n(i)}\right) + \mathbb{E}\, N_T(\mathcal{T}_{\mathbf{n}}) + O(1) \\
&= \big(|\mathbf{n}|\pi_{\mathbf{p}(\mathbf{n})}(T)\big)^2 \left(\frac{(|T|-1)^2}{|\mathbf{n}|} - \sum_{i\in\mathcal{D}(T)} \frac{n_T(i)^2}{n(i)}\right) + |\mathbf{n}|\pi_{\mathbf{p}(\mathbf{n})}(T) + O(1),
\end{aligned}
\tag{36}
$$

which yields (29) by the definitions (10), (9) and (5).

For the proof of (30) we use (26). The first two terms are handled by (28), and the final term is treated as in (34)–(36) with mainly notational differences; we omit the details. ◀

Proof of Theorem 3. By Condition 1, we have $p_i(\mathbf{n}_\kappa) \to p_i$ for every $i \geq 0$, and thus $\pi_{\mathbf{p}(\mathbf{n}_\kappa)}(T) \to \pi_{\mathbf{p}}(T)$. Hence, (i) follows from (28).

Moreover, it follows from (9)–(10) that $\gamma_{\mathbf{p}(\mathbf{n}_\kappa)}(T,T) = O(1)$ (for a fixed T), and thus (29) yields $\operatorname{Var} N_T(\mathcal{T}_{\mathbf{n}_\kappa}) = O(|\mathbf{n}_\kappa|)$. Therefore, (ii) follows from (i) and Chebyshev's inequality. ◀

4 Proof of Theorems 5

We have now all the ingredients to prove Theorem 5.

Proof of Theorem 5. First note that Condition 1 implies

$$
\pi_{\mathbf{p}(\mathbf{n}_\kappa)}(T_i) \to \pi_{\mathbf{p}}(T_i) \quad \text{and} \quad \gamma_{\mathbf{p}(\mathbf{n}_\kappa)}(T_i, T_j) \to \gamma_{\mathbf{p}}(T_i, T_j), \qquad \text{for } 1 \leq i, j \leq m. \tag{37}
$$

Hence, (12)–(14) follow from (28)–(30) in Theorem 10.

We next prove the asymptotic normality result in (15). Note first that (28) implies that it does not matter whether we use $\mathbb{E}[N_{T_i}(\mathcal{T}_{\mathbf{n}_\kappa})]$ or

$$\mu_{\mathbf{n}_\kappa}(T) := |\mathbf{n}_\kappa|\pi_{\mathbf{p}(\mathbf{n}_\kappa)}(T) = |\mathbf{n}_\kappa| \prod_{i\geq 0} p_i(\mathbf{n}_\kappa)^{n_T(i)} = |\mathbf{n}_\kappa| \prod_{i\in\mathcal{D}(T)} p_i(\mathbf{n}_\kappa)^{n_T(i)}. \tag{38}$$

in (15).

If $\pi_{\mathbf{p}}(T_i) = 0$, for some $1 \leq i \leq m$, then it follows from (10) that $\gamma_{\mathbf{p}}(T_i, T_i) = 0$, and thus (13) yields $\mathrm{Var}[N_{T_i}(\mathcal{T}_{\mathbf{n}_\kappa})] = o(|\mathbf{n}_\kappa|)$; consequently, (28) and Chebyshev's inequality yield, as $\kappa \to \infty$,

$$\frac{N_{T_i}(\mathcal{T}_{\mathbf{n}_\kappa}) - \mathbb{E}[N_{T_i}(\mathcal{T}_{\mathbf{n}_\kappa})]}{\sqrt{|\mathbf{n}_\kappa|}} \xrightarrow{\mathrm{p}} 0. \tag{39}$$

Hence, convergence of the i-th component in (15) is trivial in this case. Furthermore, $\pi_{\mathbf{p}}(T_i) = 0$ also implies $\gamma_{\mathbf{p}}(T_i, T_j) = 0$ for every $1 \leq j \leq m$ by (11), noting that if $N_{T_i}(T_j) > 0$ then also $\pi_{\mathbf{p}}(T_j) = 0$. Thus, we may ignore all i in (15) with $\pi_{\mathbf{p}}(T_i) = 0$ and show (joint) convergence for the remaining ones, because then (15) in general will follow from [4, Theorem 3.9 in Chapter 1]. Consequently, we henceforth assume that $\pi_{\mathbf{p}}(T_i) > 0$ for all $1 \leq i \leq m$. Equivalently, $p_k > 0$ for every $k \in \bigcup_{i=1}^m \mathcal{D}(T_i)$. We may also assume that $T_1, \ldots, T_m$ are distinct.

To see the main idea of the proof, we consider only the univariate case $m = 1$. The general case follows similarly by a multidimensional version of the Gao–Wormald theorem [3, Theorem A.1] in the full version. The main complication in the multivariate case is the possibility that fringe trees of type T_j may contain fringe trees of type T_k for some $1 \leq j, k \leq m$; we thus use the decomposition in (18) and estimate the terms separately; we refer to the proof of [3, Theorem 1.5] in the full version for details.

We then consider $m = 1$ and omit the index 1 and write T instead of T_1. In this case, we can use the Gao–Wormald theorem [14, Theorem 1] and the following estimate. For any $q_\kappa = O(|\mathbf{n}_\kappa|^{1/2})$, (25) and Lemma 9 yield, recalling the definitions (5), (7), (9), (10), and (38) of $p_i(\mathbf{n})$, $\pi_{\mathbf{p}}(T)$, $\eta_{\mathbf{p}}(T,T)$, $\gamma_{\mathbf{p}}(T,T)$, and $\mu_{\mathbf{n}_\kappa}(T)$,

$$\begin{aligned}
\mathbb{E}[(N_T(\mathcal{T}_{\mathbf{n}_\kappa}))_{q_\kappa}] &= \frac{\prod_{i\geq 0} n_\kappa(i)^{q_\kappa n_T(i)}}{|\mathbf{n}_\kappa|^{q_\kappa(|T|-1)}} \exp\left(\frac{\left(q_\kappa(|T|-1)\right)^2}{2|\mathbf{n}_\kappa|} - \sum_{i\geq 0} \frac{\left(q_\kappa n_T(i)\right)^2}{2n_\kappa(i)} + o(1) \right) \\
&= |\mathbf{n}_\kappa|^{q_\kappa} \prod_{i\geq 0} p_i(\mathbf{n}_\kappa)^{q_\kappa n_T(i)} \exp\left(\frac{\left(q_\kappa(|T|-1)\right)^2}{2|\mathbf{n}_\kappa|} - \sum_{i\geq 0} \frac{\left(q_\kappa n_T(i)\right)^2}{2n_\kappa(i)} + o(1) \right) \\
&= \left(|\mathbf{n}_\kappa|\pi_{\mathbf{p}(\mathbf{n}_\kappa)}(T)\right)^{q_\kappa} \exp\left(\frac{q_\kappa^2}{2|\mathbf{n}_\kappa|}\eta_{\mathbf{p}(\mathbf{n}_\kappa)}(T,T) + o(1) \right) \\
&= \mu_{\mathbf{n}_\kappa}(T)^{q_\kappa} \exp\left(\frac{(\gamma_{\mathbf{p}(\mathbf{n}_\kappa)}(T,T) - \pi_{\mathbf{p}(\mathbf{n}_\kappa)}(T))|\mathbf{n}_\kappa|}{2\mu_{\mathbf{n}_\kappa}(T)^2} q_\kappa^2 + o(1) \right) \\
&= \mu_{\mathbf{n}_\kappa}(T)^{q_\kappa} \exp\left(\frac{\gamma_{\mathbf{p}}(T,T)|\mathbf{n}_\kappa| - \mu_{\mathbf{n}_\kappa}(T)}{2\mu_{\mathbf{n}_\kappa}(T)^2} q_\kappa^2 + o(1) \right).
\end{aligned} \tag{40}$$

If $\gamma_{\mathbf{p}}(T,T) > 0$, we may now apply the Gao–Wormald theorem [14, Theorem 1] with $\mu_\kappa := \mu_{\mathbf{n}_\kappa}(T)$ and $\sigma_\kappa^2 := \gamma_{\mathbf{p}}(T,T)|\mathbf{n}_\kappa|$ and conclude (16), which by (13) is equivalent to (15) (with $m = 1$). The case $\gamma_{\mathbf{p}}(T,T) = 0$ is trivial, since then (13) implies (39). Alternatively, for any $\gamma_{\mathbf{p}}(T,T)$, we may take the same μ_κ but $\sigma_\kappa^2 := |\mathbf{n}_\kappa|$ in the case $m = 1$ of our version [3, Theorem A.1] of the Gao–Wormald theorem. ◀

5 Application to simply generated trees

Let $\mathbb{T}_n$ denote the (finite) subset of all plane rooted trees of size $n \in \mathbb{N}$. Let $\mathbf{w} = (w_i)_{i\geq 0}$ be a sequence of non-negative real weights with $w_0 > 0$ and $w_i > 0$ for at least one $i \geq 2$. For a finite rooted plane tree $T \in \mathbb{T}$, we define the weight of T to be

$$w(T) := \prod_{v\in T} w_{d_T(v)} = \prod_{i\geq 0} w_i^{n_T(i)}. \tag{41}$$

For $n \in \mathbb{N}$, let $Z_n(\mathbf{w}) = \sum_{T\in\mathbb{T}_n} w(T)$. If $Z_n(\mathbf{w}) > 0$, then we define the random tree $\mathcal{T}_{\mathbf{w},n}$ by picking an element of $\mathbb{T}_n$ at random with probability proportional to its weight, i.e.,

$$\mathbb{P}(\mathcal{T}_{\mathbf{w},n} = T) = \frac{w(T)}{Z_n(\mathbf{w})}, \quad \text{for } T \in \mathbb{T}_n. \tag{42}$$

The random tree $\mathcal{T}_{\mathbf{w},n}$ is called simply generated tree of size n and weight sequence $\mathbf{w}$; see e.g. [9] and [19]. If $\mathbf{w}$ is a probability distribution (i.e., $\sum_{i\geq 0} w_i = 1$), then $\mathcal{T}_{\mathbf{w},n}$ is a Galton–Watson tree with offspring distribution $\mathbf{w}$ conditioned to have n vertices.

Let $\Phi_{\mathbf{w}}(z) = \sum_{i\geq 0} w_i z^i$ be the generating function of the weight sequence $\mathbf{w}$, and let $\rho_{\mathbf{w}} \in [0,\infty]$ be its radius of convergence. For $0 \leq s < \rho_{\mathbf{w}}$, we let

$$\Psi_{\mathbf{w}}(s) := \frac{s\Phi'_{\mathbf{w}}(s)}{\Phi_{\mathbf{w}}(s)} = \frac{\sum_{i\geq 0} i w_i s^i}{\sum_{i\geq 0} w_i s^i}. \tag{43}$$

Furthermore, if $\Phi_{\mathbf{w}}(\rho_{\mathbf{w}}) < \infty$, we define also $\Psi_{\mathbf{w}}(\rho_{\mathbf{w}})$ by (43); if $\Phi_{\mathbf{w}}(\rho_{\mathbf{w}}) = \infty$ then we define $\Psi_{\mathbf{w}}(\rho_{\mathbf{w}}) := \lim_{s\uparrow\rho_{\mathbf{w}}} \Psi_{\mathbf{w}}(s)$; the limit exists by [19, Lemma 3.1 (i)]. Let $\nu_{\mathbf{w}} := \Psi_{\mathbf{w}}(\rho_{\mathbf{w}}) \in [0,\infty]$, and define

$$\tau_{\mathbf{w}} = \begin{cases} \rho_{\mathbf{w}} & \text{if } \nu_{\mathbf{w}} < 1, \\ \Psi_{\mathbf{w}}^{-1}(1) & \text{if } \nu_{\mathbf{w}} \geq 1. \end{cases} \tag{44}$$

It follows from [19, Lemma 3.1] that

$$\rho_{\mathbf{w}} > 0 \iff \nu_{\mathbf{w}} > 0 \iff \tau_{\mathbf{w}} > 0. \tag{45}$$

The following result from [19] shows that simply generated trees satisfy Condition 1 in probability.

▶ **Theorem 11** ([19, Theorem 7.1 and Theorem 7.11]). *Let $\mathbf{w}$ be a sequence of non-negative real weights with $w_0 > 0$ and $w_i > 0$ for at least one $i \geq 2$. Define*

$$\theta_i(\mathbf{w}) = \frac{w_i \tau_{\mathbf{w}}^i}{\Phi_{\mathbf{w}}(\tau_{\mathbf{w}})}, \quad \textit{for } i \geq 0. \tag{46}$$

Then, $\theta(\mathbf{w}) = (\theta_i(\mathbf{w}))_{i\geq 0}$ is a probability distribution with expectation $\mu_{\mathbf{w}} = \min(1,\nu_{\mathbf{w}})$ and variance $\sigma_{\mathbf{w}}^2 = \tau_{\mathbf{w}}\Psi'_{\mathbf{w}}(\tau_{\mathbf{w}}) \in [0,\infty]$. Moreover, for $n \in \mathbb{N}$ with $Z_n(\mathbf{w}) > 0$, let $\mathcal{T}_{\mathbf{w},n}$ be a simply generated tree of size n and weight sequence $\mathbf{w}$. Then, the (empirical) degree distribution $\mathbf{p}(\mathbf{n}_{\mathcal{T}_{\mathbf{w},n}})$ of $\mathcal{T}_{\mathbf{w},n}$ satisfies, for every $i \geq 0$, $p_i(\mathbf{n}_{\mathcal{T}_{\mathbf{w},n}}) \xrightarrow{\text{p}} \theta_i(\mathbf{w})$, as $n \to \infty$ (along integers n such that $Z_n(\mathbf{w}) > 0$).

Note that if $\rho_{\mathbf{w}} = 0$, then $\theta_0(\mathbf{w}) = 1$ and $\theta_i(\mathbf{w}) = 0$ for $i \geq 1$; otherwise, $\tau_{\mathbf{w}} > 0$ and (46) shows that $\theta_i(\mathbf{w}) > 0 \iff w_i > 0$ for $i \geq 0$.

Using Theorem 11, we can show that Theorem 5 implies the following version for conditioned Galton–Watson trees. The asymptotic normality (49) was proved in case (i) by different methods in [20, Corollary 1.8]; (ii) and (iii) are new.

▶ **Theorem 12** (partly [20]). *Let $\mathbf{w}$ be a sequence of non-negative real weights with $w_0 > 0$ and $w_i > 0$ for at least one $i \geq 2$. Moreover, for $n \in \mathbb{N}$ with $Z_n(\mathbf{w}) > 0$, let $\mathcal{T}_{\mathbf{w},n}$ be a simply generated tree of size n and weight sequence $\mathbf{w}$. For fixed $m \geq 1$, let $T_1, \dots, T_m \in \mathbb{T}$ be a fixed sequence of rooted plane trees. Then, as $n \to \infty$ (along integers n such that $Z_n(\mathbf{w}) > 0$),*

$$\left(\frac{N_{T_j}(\mathcal{T}_{\mathbf{w},n}) - \mathbb{E}[N_{T_j}(\mathcal{T}_{\mathbf{w},n}) \mid \mathbf{n}_{\mathcal{T}_{\mathbf{w},n}}]}{\sqrt{n}} \right)_{j=1}^{m} \xrightarrow{\text{d}} \mathrm{N}(0, \Gamma_{\theta(\mathbf{w})}), \tag{47}$$

where the covariance matrix $\Gamma_{\theta(\mathbf{w})}$ is defined by (10)–(11), and for $1 \leq j \leq m$,

$$\mathbb{E}[N_{T_j}(\mathcal{T}_{\mathbf{w},n}) \mid \mathbf{n}_{\mathcal{T}_{\mathbf{w},n}}] = \frac{n}{(n)_{|T_j|}} \prod_{i \geq 0} (n_{\mathcal{T}_{\mathbf{w},n}}(i))_{n_{T_j}(i)}. \tag{48}$$

Furthermore, suppose that the weight sequence $\mathbf{w}$ satisfies one of the following conditions:

(i) *$\nu_{\mathbf{w}} \geq 1$ and $\sigma^2_{\mathbf{w}} \in (0, \infty)$.*
(ii) *$\nu_{\mathbf{w}} \geq 1$, $\sigma^2_{\mathbf{w}} = \infty$ and $\theta(\mathbf{w})$ belongs to the domain of attraction of a stable law of index $\alpha \in (1, 2]$. (The last condition is equivalent to that there exists a slowly varying function $L : \mathbb{R}_+ \to \mathbb{R}_+$ such that $\sum_{i=0}^{k} i^2 \theta_i(\mathbf{w}) = k^{2-\alpha} L(k)$, as $k \to \infty$ [10, Theorem XVII.5.2].)*
(iii) *$0 < \nu_{\mathbf{w}} < 1$ and $\theta_i(\mathbf{w}) = c i^{-\beta} + o(i^{-\beta})$, as $i \to \infty$, with fixed $c > 0$ and $\beta > 2$.*

Then, as $n \to \infty$ (along integers n such that $Z_n(\mathbf{w}) > 0$),

$$\left(\frac{N_{T_j}(\mathcal{T}_{\mathbf{w},n}) - n\pi_{\theta(\mathbf{w})}(T_j)}{\sqrt{n}} \right)_{j=1}^{m} \xrightarrow{\text{d}} \mathrm{N}(0, \widetilde{\Gamma}_{\theta(\mathbf{w})}), \tag{49}$$

where the covariance matrix $\widetilde{\Gamma}_{\theta(\mathbf{w})} = (\widetilde{\gamma}_{\theta(\mathbf{w})}(T_i, T_j))_{i,j=1}^{m}$ is given by, for $T, T' \in \mathbb{T}$ such that $T \neq T'$,

$$\widetilde{\gamma}_{\theta(\mathbf{w})}(T, T) = \pi_{\theta(\mathbf{w})}(T) - \left(2|T| - 1 + \varsigma_{\mathbf{w}}^{-2}\right) (\pi_{\theta(\mathbf{w})}(T))^2, \tag{50}$$

$$\widetilde{\gamma}_{\theta(\mathbf{w})}(T, T') = N_{T'}(T)\pi_{\theta(\mathbf{w})}(T) + N_T(T')\pi_{\theta(\mathbf{w})}(T') - \left(|T| + |T'| - 1 + \varsigma_{\mathbf{w}}^{-2}\right) \pi_{\theta(\mathbf{w})}(T)\pi_{\theta(\mathbf{w})}(T'), \tag{51}$$

with $\varsigma^2_{\mathbf{w}} = \sigma^2_{\mathbf{w}}$ in case (i)*, and $\varsigma^2_{\mathbf{w}} = \infty$ in cases* (ii) *and* (iii)*.*

▶ Remark 13. Recall that for any weight sequence $\mathbf{w}$ and any constants $a, b > 0$, the weight sequence $\widehat{\mathbf{w}} = (\widehat{w}_i)_{i \geq 0}$ with $\widehat{w}_i := ab^i w_i$ is equivalent to $\mathbf{w}$, i.e., it satisfies that $\mathcal{T}_{\mathbf{w},n} \overset{\text{d}}{=} \mathcal{T}_{\widehat{\mathbf{w}},n}$, for all n for which either (and thus both) of the random trees are defined; this is a consequence of (42). In the setting of Theorem 11, if $\rho_{\mathbf{w}} > 0$, then the weight sequence $\mathbf{w}$ is equivalent to the weight sequence $\theta(\mathbf{w}) = (\theta_i(\mathbf{w}), i \geq 0)$, which is a probability distribution with mean $\mu_{\mathbf{w}} = \min(1, \nu_{\mathbf{w}})$; see further [19, Section 7]. Thus, if $\rho_{\mathbf{w}} > 0$ we can regard $\mathcal{T}_{\mathbf{w},n}$ as a Galton–Watson tree $\mathcal{T}_{\theta(\mathbf{w}),n}$ with offspring distribution $\theta(\mathbf{w})$ conditioned to have n vertices. This explains the appearance of $\theta(\mathbf{w})$ in Theorem 12, and it shows that there is no real loss of generality to consider (as is often done) only the case $\tau_{\mathbf{w}} = 1$ when $\theta(\mathbf{w}) = \mathbf{w}$. Note that the conditioned Galton–Watson tree $\mathcal{T}_{\theta(\mathbf{w}),n}$ is critical if $\nu_{\mathbf{w}} \geq 1$, and subcritical if $0 < \nu_{\mathbf{w}} < 1$.

The complete proof of Theorem 12 is given in [3, Section 7] of the full version. Here, we only comment on the main ideas. Indeed, for any fixed degree statistic $\mathbf{n}$ with $\mathbb{P}(\mathbf{n}_{\mathcal{T}_{\mathbf{w},n}} = \mathbf{n}) > 0$, (42) implies that conditionally given $\mathbf{n}_{\mathcal{T}_{\mathbf{w},n}} = \mathbf{n}$, $\mathcal{T}_{\mathbf{w},n} \sim \mathrm{Unif}(\mathbb{T}_{\mathbf{n}})$; see e.g., [1, Proposition 8]. By the Skorohod coupling theorem [22, Theorem 4.30], we can assume that the convergence in Theorem 11 holds a.s.; in other words, Condition 1 holds a.s. for the degree statistics $\mathbf{n}_{\mathcal{T}_{\mathbf{w},n}}$, with $\mathbf{p} = \theta(\mathbf{w})$. Moreover, e.g. by resampling $\mathcal{T}_{\mathbf{w},n}$ conditioned on $\mathbf{n}_{\mathcal{T}_{\mathbf{w},n}}$, we may assume that also conditioned on the entire sequence of degree statistics $(\mathbf{n}_{\mathcal{T}_{\mathbf{w},n}})_{n=1}^{\infty}$, the random trees $\mathcal{T}_{\mathbf{w},n}$, $n \geq 1$, have the (conditional) distributions $\mathrm{Unif}(\mathbb{T}_{\mathbf{n}_{\mathcal{T}_{\mathbf{w},n}}})$. It follows that we may apply Theorem 5 conditioned on the sequence of degree statistics $(\mathbf{n}_{\mathcal{T}_{\mathbf{w},n}})_{n=1}^{\infty}$; this shows that (47) holds conditioned on $(\mathbf{n}_{\mathcal{T}_{\mathbf{w},n}})_{n=1}^{\infty}$. Then, (47) also holds unconditionally by the dominated convergence theorem. Furthermore, (48) follows from Lemma 8 (with $q = 1$). On the other hand, the central idea to obtain the unconditional limit (49) is by combining the conditional limit (47) with a limit result for the conditional expectations in (48). For this, one uses a theorem on asymptotic normality of the degree statistics, which is proved in [20] and [24] (see also [3, Theorem 7.6] for a different approach).

Theorem 12 gives a partial solution to [19, Problem 21.4], but the general case remains open.

▶ **Problem 14.** *Does (49) in Theorem 12 hold for any weight sequence* $\mathbf{w}$*, with some covariance matrix* $\widetilde{\Gamma}_{\theta(\mathbf{w})} = (\widetilde{\gamma}_{\theta(\mathbf{w})}(T_i, T_j))_{i,j=1}^{m}$*?*

References

1. Louigi Addario-Berry and Serte Donderwinkel. Random trees have height $O(\sqrt{n})$. *arXiv e-prints*, January 2022. arXiv:2201.11773.
2. David Aldous. Asymptotic fringe distributions for general families of random trees. *Ann. Appl. Probab.*, 1(2):228–266, 1991.
3. Gabriel Berzunza Ojeda, Cecilia Holmgren, and Svante Janson. Fringe trees for random trees with given vertex degrees. *arXiv e-prints*, December 2023. arXiv:2312.04243.
4. Patrick Billingsley. *Convergence of probability measures*. Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons, Inc., New York, second edition, 1999. A Wiley-Interscience Publication. doi:10.1002/9780470316962.
5. Xing Shi Cai and Luc Devroye. A study of large fringe and non-fringe subtrees in conditional Galton-Watson trees. *ALEA Lat. Am. J. Probab. Math. Stat.*, 14(1):579–611, 2017. doi:10.30757/alea.v14-29.
6. Huilan Chang and Michael Fuchs. Limit theorems for patterns in phylogenetic trees. *J. Math. Biol.*, 60(4):481–512, 2010. doi:10.1007/s00285-009-0275-6.
7. Luc Devroye. Limit laws for local counters in random binary search trees. *Random Structures Algorithms*, 2(3):303–315, 1991. doi:10.1002/rsa.3240020305.
8. Luc Devroye. Limit laws for sums of functions of subtrees of random binary search trees. *SIAM J. Comput.*, 32(1):152–171, 2002/03. doi:10.1137/S0097539701383923.
9. Michael Drmota. *Random trees, An interplay between combinatorics and probability*. Springer-WienNewYork, Vienna, 2009. doi:10.1007/978-3-211-75357-6.
10. William Feller. *An introduction to probability theory and its applications. Vol. II.* John Wiley & Sons, Inc., New York-London-Sydney, second edition, 1971.
11. Qunqiang Feng and Hosam M. Mahmoud. On the variety of shapes on the fringe of a random recursive tree. *J. Appl. Probab.*, 47(1):191–200, 2010. doi:10.1239/jap/1269610825.
12. Philippe Flajolet, Xavier Gourdon, and Conrado Martínez. Patterns in random binary search trees. *Random Structures Algorithms*, 11(3):223–244, 1997. doi:10.1002/(SICI)1098-2418(199710)11:3<223::AID-RSA2>3.0.CO;2-2.
13. Michael Fuchs. Limit theorems for subtree size profiles of increasing trees. *Combin. Probab. Comput.*, 21(3):412–441, 2012. doi:10.1017/S096354831100071X.
14. Zhicheng Gao and Nicholas C. Wormald. Asymptotic normality determined by high moments, and submap counts of random maps. *Probab. Theory Related Fields*, 130(3):368–376, 2004. doi:10.1007/s00440-004-0356-9.
15. Pawel HItczenko and Nick Wormald. Multivariate asymptotic normality determined by high moments. *arXiv e-prints*, December 2023. arXiv:2312.04246.

16 Cecilia Holmgren and Svante Janson. Limit laws for functions of fringe trees for binary search trees and random recursive trees. *Electron. J. Probab.*, 20:no. 4, 51, 2015. `doi:10.1214/EJP.v20-3627`.

17 Cecilia Holmgren and Svante Janson. Fringe trees, Crump-Mode-Jagers branching processes and m-ary search trees. *Probab. Surv.*, 14:53–154, 2017. `doi:10.1214/16-PS272`.

18 Cecilia Holmgren, Svante Janson, and Matas Šileikis. Multivariate normal limit laws for the numbers of fringe subtrees in m-ary search trees and preferential attachment trees. *Electron. J. Combin.*, 24(2):Paper No. 2.51, 49, 2017. `doi:10.37236/6374`.

19 Svante Janson. Simply generated trees, conditioned Galton-Watson trees, random allocations and condensation. *Probab. Surv.*, 9:103–252, 2012. `doi:10.1214/11-PS188`.

20 Svante Janson. Asymptotic normality of fringe subtrees and additive functionals in conditioned Galton-Watson trees. *Random Structures Algorithms*, 48(1):57–101, 2016. `doi:10.1002/rsa.20568`.

21 Svante Janson. Central limit theorems for additive functionals and fringe trees in tries. *Electron. J. Probab.*, 27:Paper No. 47, 1–63, 2022. `doi:10.1214/22-ejp776`.

22 Olav Kallenberg. *Foundations of modern probability*. Probability and its Applications (New York). Springer-Verlag, New York, second edition, 2002. `doi:10.1007/978-1-4757-4015-8`.

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

24 Paul Thévenin. Vertices with fixed outdegrees in large Galton-Watson trees. *Electron. J. Probab.*, 25:Paper No. 64, 25, 2020. `doi:10.1214/20-ejp465`.