Fringe trees for random trees with given vertex degrees

Gabriel Berzunza Ojeda[*], Cecilia Holmgren[†] and Svante Janson[‡]

6 December, 2023

Abstract

We prove asymptotic normality for the number of fringe subtrees isomorphic to any given tree in uniformly random trees with given vertex degrees. As applications, we also prove corresponding results for random labelled trees with given vertex degrees, for random simply generated trees (or conditioned Galton–Watson trees), and for additive functionals.

The key tool for our work is an extension to the multivariate setting of a theorem by Gao and Wormald (2004), which provides a way to show asymptotic normality by analysing the behaviour of sufficiently high factorial moments.

Key words and phrases: Additive functionals; conditioned Galton-Watson trees; fringe trees; random labelled trees; simply generated trees; toll functions.
MSC 2020 Subject Classifications: 60C05; 05C05; 60F05.

1 Introduction and main results

In this paper, we consider fringe trees of three types of random trees. In the main parts of the paper, we consider random plane trees with given vertex statistics, i.e., a given number of vertices of each degree. As applications of these results, we also give corresponding results for random labelled trees with given vertex degrees, and 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., [8]. 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.

Supported by the Knut and Alice Wallenberg Foundation; Ragnar Söderberg's Foundation; the Swedish Research Council.
[*]Department of Mathematical Sciences, University of Liverpool, United Kingdom. E-mail: gabriel.berzunza-ojeda@liverpool.ac.uk URL: https://www.liverpool.ac.uk/mathematical-sciences/staff/gabriel-berzunza-ojeda/
[†]Department of Mathematics, Uppsala University, Sweden. E-mail: cecilia.holmgren@math.uu.se https://katalog.uu.se/empinfo/?id=N5-824
[‡]Department of Mathematics, Uppsala University, Sweden. E-mail: svante.janson@math.uu.se URL: http://www2.math.uu.se/ svante/

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

$$|T| = \sum_{i\geq 0} n_T(i) = 1 + \sum_{i\geq 0} i n_T(i). \tag{1.1}$$

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 $\mathcal{T}_{\mathbf{n}} \sim \mathrm{Unif}(\mathbb{T}_{\mathbf{n}})$.

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 $T, T' \in \mathbb{T}$,

$$N_{T'}(T) := |\{v \in T : T_v = T'\}| = \sum_{v \in T} \mathbf{1}_{\{T_v = T'\}}, \tag{1.2}$$

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

$$\mathbb{P}(T^{\mathrm{fr}} = T') = \frac{N_{T'}(T)}{|T|}, \quad \text{for } T' \in \mathbb{T}. \tag{1.3}$$

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.,

$$p_i(\mathbf{n}) := \frac{n(i)}{|\mathbf{n}|}, \quad \text{for } i \geq 0. \tag{1.4}$$

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

Condition 1.1. $\mathbf{n}_\kappa = (n_\kappa(i))_{i\geq 0}$, $\kappa \geq 1$, *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 1.2. The condition that $\mathbf{p}$ is a probability distribution is no restriction. In fact, the degree distribution $\mathbf{p}(\mathbf{n}_\kappa)$ has mean

$$\sum_{i\geq 0} i p_i(\mathbf{n}_\kappa) = \frac{1}{|\mathbf{n}_\kappa|} \sum_{i\geq 0} i n_\kappa(i) = \frac{|\mathbf{n}_\kappa| - 1}{|\mathbf{n}_\kappa|} < 1, \tag{1.5}$$

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 (1.5) and Fatou's lemma, if Condition 1.1 holds, then $\sum_{i\geq 0} ip_i \leq 1$. 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 1.1 is precisely the set of probability distributions $\mathbf{p}$ on $\mathbb{N}_0$ with mean $\sum_{i\geq 0} ip_i \leq 1$; 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)

$$\pi_\mathbf{p}(T) := \mathbb{P}(\mathcal{T}_\mathbf{p} = T) = \prod_{i\geq 0} p_i^{n_T(i)} = \prod_{i\in\mathcal{D}(T)} p_i^{n_T(i)}, \qquad \text{for } T \in \mathbb{T}, \tag{1.6}$$

where

$$\mathcal{D}(T) := \{i : n_T(i) > 0\} = \{d_T(v) : v \in T\}, \tag{1.7}$$

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 1.1, 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 theorems are given in later sections.

Theorem 1.3. *Let $\mathbf{n}_\kappa$, $\kappa \geq 1$, be some degree statistics that satisfy Condition 1.1, 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)* $\quad \mathbb{P}(\mathcal{T}_{\mathbf{n}_\kappa}^{\mathrm{fr}} = T) = \dfrac{\mathbb{E}[N_T(\mathcal{T}_{\mathbf{n}_\kappa})]}{|\mathbf{n}_\kappa|} \to \pi_\mathbf{p}(T).$

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

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

Remark 1.4. Similar results are known for several other models of random trees. In particular, a version of Theorem 1.3 was proved by Aldous [2] for conditioned Galton–Watson trees with finite offspring variance; this was extended to general simply generated trees in [21, Theorem 7.12]. In those cases, the degree statistic is random, but Condition 1.1 holds in probability, with a non-random limiting probability distribution $\mathbf{p}$. We return to simply generated trees in Section 7. 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 1.3 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

$$\eta_{\mathbf{p}}(T,T') := (|T|-1)(|T'|-1) - \sum_{i\geq 0} \frac{n_T(i) n_{T'}(i)}{p_i}, \tag{1.8}$$

where we interpret $0/0 := 0$, and

$$\gamma_{\mathbf{p}}(T,T) := \pi_{\mathbf{p}}(T) + \eta_{\mathbf{p}}(T,T)(\pi_{\mathbf{p}}(T))^2, \tag{1.9}$$

$$\gamma_{\mathbf{p}}(T,T') := N_{T'}(T)\pi_{\mathbf{p}}(T) + N_T(T')\pi_{\mathbf{p}}(T') + \eta_{\mathbf{p}}(T,T')\pi_{\mathbf{p}}(T)\pi_{\mathbf{p}}(T'), \qquad T \neq T'. \tag{1.10}$$

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 (1.9)–(1.10); thus $\gamma_{\mathbf{p}}(T,T')$ is always finite.

Theorem 1.5. *Let $\mathbf{n}_\kappa$, $\kappa \geq 1$, be some degree statistics that satisfy Condition 1.1 and let $\mathcal{T}_{\mathbf{n}_\kappa} \sim \mathrm{Unif}(\mathbb{T}_{\mathbf{n}_\kappa})$. For a fixed $m \geq 1$, let $T_1, \dots, T_m \in \mathbb{T}$ be a fixed sequence of rooted plane trees. Then, as $\kappa \to \infty$,*

$$\mathbb{E} N_{T_i}(\mathcal{T}_{\mathbf{n}_\kappa}) = \pi_{\mathbf{p}}(T_i)|\mathbf{n}_\kappa| + o(|\mathbf{n}_\kappa|), \tag{1.11}$$

$$\mathrm{Var}(N_{T_i}(\mathcal{T}_{\mathbf{n}_\kappa})) = \gamma_{\mathbf{p}}(T_i,T_i)|\mathbf{n}_\kappa| + o(|\mathbf{n}_\kappa|), \tag{1.12}$$

$$\mathrm{Cov}\Big(N_{T_i}(\mathcal{T}_{\mathbf{n}_\kappa}), N_{T_j}(\mathcal{T}_{\mathbf{n}_\kappa})\Big) = \gamma_{\mathbf{p}}(T_i,T_j)|\mathbf{n}_\kappa| + o(|\mathbf{n}_\kappa|), \tag{1.13}$$

for $1 \leq i,j \leq m$, and

$$\left(\frac{N_{T_1}(\mathcal{T}_{\mathbf{n}_\kappa}) - \mathbb{E}[N_{T_1}(\mathcal{T}_{\mathbf{n}_\kappa})]}{\sqrt{|\mathbf{n}_\kappa|}}, \dots, \frac{N_{T_m}(\mathcal{T}_{\mathbf{n}_\kappa}) - \mathbb{E}[N_{T_m}(\mathcal{T}_{\mathbf{n}_\kappa})]}{\sqrt{|\mathbf{n}_\kappa|}} \right) \xrightarrow{\mathrm{d}} \mathrm{N}(0, \Gamma_{\mathbf{p}}), \tag{1.14}$$

where the covariance matrix $\Gamma_{\mathbf{p}} := (\gamma_{\mathbf{p}}(T_i,T_j))_{i,j=1}^m$. Furthermore, in (1.14), 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 (1.12) and (1.14) (with $m = 1$) show that $N_T(\mathcal{T}_{\mathbf{n}_\kappa})$ is asymptotically normal, with

$$\frac{N_T(\mathcal{T}_{\mathbf{n}_\kappa}) - \mathbb{E}[N_T(\mathcal{T}_{\mathbf{n}_\kappa})]}{\sqrt{\mathrm{Var}(N_T(\mathcal{T}_{\mathbf{n}_\kappa}))}} \xrightarrow{\mathrm{d}} \mathrm{N}(0,1), \qquad \kappa \to \infty. \tag{1.15}$$

The case $|T| = 1$ is trivial, with $N_T(\mathcal{T}_{\mathbf{n}_\kappa}) = n_\kappa(0)$ non-random. Ignoring this case, Theorem 1.5 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, the following theorems are more precise.

Theorem 1.6. *Let $T \in \mathbb{T}$ be a fixed tree. Then, uniformly for all degree statistics $\mathbf{n} = (n(i))_{i\geq 0}$,*

$$\mathbb{E} N_T(\mathcal{T}_{\mathbf{n}}) = |\mathbf{n}|\pi_{\mathbf{p}(\mathbf{n})}(T) + O(1), \tag{1.16}$$

$$\mathrm{Var}\, N_T(\mathcal{T}_{\mathbf{n}}) = |\mathbf{n}|\gamma_{\mathbf{p}(\mathbf{n})}(T,T) + O(1). \tag{1.17}$$

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

$$\operatorname{Cov}\Big(N_T(\mathcal{T}_{\mathbf{n}}), N_{T'}(\mathcal{T}_{\mathbf{n}})\Big) = |\mathbf{n}|\gamma_{\mathbf{p}(\mathbf{n})}(T, T') + O(1). \tag{1.18}$$

In view of (1.16), we define, for any degree statistic $\mathbf{n}$ and tree $T \in \mathbb{T}$,

$$\mu_{\mathbf{n}}(T) := |\mathbf{n}|\pi_{\mathbf{p}(\mathbf{n})}(T) = |\mathbf{n}|\prod_{i\geq 0} p_i(\mathbf{n})^{n_T(i)} = |\mathbf{n}| \prod_{i \in \mathcal{D}(T)} p_i(\mathbf{n})^{n_T(i)}. \tag{1.19}$$

This is thus a convenient approximation of $\mathbb{E}\, N_T(\mathcal{T}_{\mathbf{n}})$. We define also

$$\widehat{\eta}_{\mathbf{p}}(T, T') := (\pi_{\mathbf{p}}(T)\pi_{\mathbf{p}}(T'))^{1/2}\eta_{\mathbf{p}}(T, T'), \qquad \text{if } \pi_{\mathbf{p}}(T), \pi_{\mathbf{p}}(T') > 0, \tag{1.20}$$

and extend this by continuity to the case $\pi_{\mathbf{p}}(T)\pi_{\mathbf{p}}(T') = 0$; this yields by (1.8) the general formula

$$\widehat{\eta}_{\mathbf{p}}(T, T') = (\pi_{\mathbf{p}}(T)\pi_{\mathbf{p}}(T'))^{1/2}(|T|-1)(|T'|-1) - \sum_{i\geq 0} n_T(i) n_{T'}(i) \prod_{j \in \mathcal{D}(T)\cup\mathcal{D}(T')} p_j^{(n_T(j)+n_{T'}(j))/2-\delta_{ij}}. \tag{1.21}$$

We interpret again $0\cdot\infty := 0$; thus the sum in (1.21) is finite also if $p_i = 0$ for some $i \in \mathcal{D}(T)\cup\mathcal{D}(T')$. In fact, (1.21) is a polynomial in $p_0^{1/2}, p_1^{1/2}, \dots$, and is thus continuous in $\mathbf{p}$ as asserted.

Similarly, we define

$$\widehat{\gamma}_{\mathbf{p}}(T, T') := (\pi_{\mathbf{p}}(T)\pi_{\mathbf{p}}(T'))^{-1/2}\gamma_{\mathbf{p}}(T, T'), \qquad \text{if } \pi_{\mathbf{p}}(T), \pi_{\mathbf{p}}(T') > 0, \tag{1.22}$$

and extend this by continuity, which by (1.9)–(1.10) and (1.20) yields

$$\widehat{\gamma}_{\mathbf{p}}(T, T) := 1 + \widehat{\eta}_{\mathbf{p}}(T, T), \tag{1.23}$$

$$\widehat{\gamma}_{\mathbf{p}}(T, T') := N_{T'}(T)\prod_{i\in\mathcal{D}(T)} p_i^{(n_i(T)-n_i(T'))/2} + N_T(T')\prod_{i\in\mathcal{D}(T')} p_i^{(n_i(T')-n_i(T))/2} + \widehat{\eta}_{\mathbf{p}}(T, T'), \qquad T \neq T'. \tag{1.24}$$

Note that $N_{T'}(T) > 0$ implies $\mathcal{D}(T') \subseteq \mathcal{D}(T)$ and $n_T(i) \geq n_{T'}(i)$, for $i \geq 0$; thus (1.24) always yields a finite value (again interpreting $0\cdot\infty := 0$); again, this is a polynomial in $p_0^{1/2}, p_1^{1/2}, \dots$, and thus continuous in $\mathbf{p}$.

Theorem 1.7. *Let $\mathbf{n}_\kappa$, $\kappa \geq 1$, be some degree statistics that satisfy Condition 1.1 and let $\mathcal{T}_{\mathbf{n}_\kappa} \sim \operatorname{Unif}(\mathbb{T}_{\mathbf{n}_\kappa})$. For fixed $m \geq 1$, let $T_1, \dots, T_m \in \mathbb{T}$ be a fixed sequence of rooted plane trees such that, as $\kappa \to \infty$, $\mu_{\mathbf{n}_\kappa}(T_i) := |\mathbf{n}_\kappa|\pi_{\mathbf{p}(\mathbf{n}_\kappa)}(T_i) \to \infty$ for each $1 \leq i \leq m$. Then,*

$$\left(\frac{N_{T_1}(\mathcal{T}_{\mathbf{n}_\kappa}) - \mathbb{E}[N_{T_1}(\mathcal{T}_{\mathbf{n}_\kappa})]}{\sqrt{\mu_{\mathbf{n}_\kappa}(T_1)}}, \dots, \frac{N_{T_m}(\mathcal{T}_{\mathbf{n}_\kappa}) - \mathbb{E}[N_{T_m}(\mathcal{T}_{\mathbf{n}_\kappa})]}{\sqrt{\mu_{\mathbf{n}_\kappa}(T_m)}}\right) \xrightarrow{\mathrm{d}} \mathrm{N}(0, \widehat{\Gamma}_{\mathbf{p}}), \quad \text{as } \kappa \to \infty, \tag{1.25}$$

where the covariance matrix $\widehat{\Gamma}_{\mathbf{p}} := (\widehat{\gamma}_{\mathbf{p}}(T_i, T_j))_{i,j=1}^m$. Furthermore, in (1.25), we can replace $\mathbb{E}[N_{T_i}(\mathcal{T}_{\mathbf{n}_\kappa})]$ by $\mu_{\mathbf{n}_\kappa}(T_i)$.

Moreover, $\widehat{\gamma}_{\mathbf{p}}(T, T) > 0$, and thus the asymptotic normality (1.15) holds, except in the following three *exceptional cases:*

(i) $|T| = 1$,

(ii) *T is a path and* $p_1 = 1$,

(iii) *T is a star with a root of degree d joined to d leaves, and* $p_0 = 1$.

The exceptional cases (ii) and (iii) are discussed further in Example 5.3.

Remark 1.8. Theorem 1.7 shows that excluding the exceptional cases (i)–(iii), the condition $\mu_{\mathbf{n}_\kappa}(T) \to \infty$, as $\kappa \to \infty$, is sufficient for asymptotic normality of $N_T(\mathcal{T}_{\mathbf{n}_\kappa})$. This condition is also necessary, since otherwise (at least for a subsequence) $\mathbb{E} N_T(\mathcal{T}_{\mathbf{n}_\kappa}) = O(1)$ by (1.16), and since $N_T(\mathcal{T}_{\mathbf{n}_\kappa})$ is integer-valued, it is easy to see that then it cannot converge to a non-degenerate normal distribution for any normalization.

Problem 1.9. *In Theorem 1.5, suppose that $T_1, \dots, T_m$ are distinct with $|T_i| > 1$ and $\pi_{\mathbf{p}}(T_i) > 0$ for every $1 \le i \le m$. Theorem 1.5 says that $\gamma_{\mathbf{p}}(T_i, T_i) > 0$, for every $1 \le i \le m$. Is the covariance matrix $\Gamma_{\mathbf{p}}$ non-singular?*

In the case of critical conditioned Galton–Watson trees with finite offspring variance, (joint) normal convergence of the subtree counts in analogy to (1.14) was proved in [22, Corollary 1.8] (together with convergence of mean and variance). Indeed, [22, Theorem 1.5] proved, more generally, asymptotic normality of additive functionals that are defined via toll functions (under some conditions); see Section 8 for further discussion on additive functionals.

Remark 1.10. 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 [6], [7], [5], [11], [16]; for random recursive trees see [10], [16]; for increasing trees see [12]; for m-ary search trees and preferential attachment trees see [18]; for random tries see [23].

Our approach relies on a multivariate version of the Gao–Wormald theorem [13, Theorem 1]; see Theorem A.1 in Appendix A. The original Gao–Wormald theorem [13] 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.) Our multivariate version 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 (1.2) 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]; see further Appendix A.)

The (one dimensional) Gao–Wormald theorem has been used before by Cai and Devroye [4] 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

Some standard facts on coding trees by walks are recalled in Section 2; these facts are used in Section 3 to give exact formulas for factorial moments of $N_T(\mathcal{T}_{\mathbf{n}})$. These formulas are then used in Sections 4–5 to prove our main results.

Applications to labelled trees with given vertex degrees, simply generated trees and additive functionals are given in Sections 6–8.

Appendix A contains a general statement and proof of the multivariate version of the Gao–Wormald theorem that we use in our proof of the main theorems. Appendix B uses that theorem to give a new simple proof (in two cases) of a known result on asymptotic normality of degree statistics in conditioned Galton–Watson trees (Theorem 7.6) that we use in the proof in Section 7.

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. Recall that $a_\kappa = \Theta(b_\kappa)$ means $a_\kappa = O(b_\kappa)$ and $b_\kappa = O(a_\kappa)$.

For two sequence of positive real numbers $(a_n)_{n\ge1}$ and $(b_n)_{n\ge1}$, we write $a_n \ll b_n$ or $b_n \gg a_n$ if and only if $a_n/b_n \to 0$ as $n \to \infty$.

$\mathbf{1}_{\mathcal{E}}$ 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 Trees and walks

For $k \in \mathbb{N}$, we view a sequence $x = (x(0),x(1),\dots,x(k)) \in \mathbb{Z}^{k+1}$ as a walk with steps (or increments) given by $\Delta x(i) := x(i) - x(i-1)$, for $1 \le i \le k$. Define the set of all (discrete) bridges finishing in position -1 at time k, as

$$\mathbb{B}^k := \Big\{(x(0),x(1),\dots,x(k)) \in \mathbb{Z}^{k+1} : x(0) = 0,\, \Delta x(i) \ge -1, \text{ for } 1 \le i \le k, \text{ and } x(k) = -1\Big\}. \tag{2.1}$$

For $1 \le i \le k$ and $x = (x(0),x(1),\dots,x(k)) \in \mathbb{B}^k$, define $\omega_i(x)$ as the cyclic shift of x by i, that is, the sequence of length k starting at 0 whose j-th increment is $\Delta x(i+j)$ with $i+j$ interpreted mod k. For any $x \in \mathbb{B}^k$, let τ_x be the first time that x hits its overall minimum, i.e., $\min_{1\le i\le k} x(i)$. The (discrete form of) Vervaat's transformation of x is defined by $V(x) = \omega_{\tau_x}(x)$; see [28, Exercise 6.1.1] or [30]. This transformation maps the set of bridges $\mathbb{B}^k$ to the set of excursions of size k finishing at -1:

$$\mathbb{E}^k := \Big\{(w(0), w(1), \dots, w(k)) \in \mathbb{Z}^{k+1} : w(0) = 0,\ \Delta w(i) \ge -1 \text{ for } 1 \le i \le k, \text{ and } w \text{ first hits } -1 \text{ at time } k\Big\}. \tag{2.2}$$

Let $\mathbf{n} = (n(i))_{i\ge 0}$ be a degree statistic with associated degree sequence $c(\mathbf{n}) = (c(i))_{i\ge 1}$, that is, the sequence obtained by writing $n(0)$ zeros, $n(1)$ ones, and so on. Let σ be a uniformly random permutation of $\{1,\dots,|\mathbf{n}|\}$. Define the bridge $W^{\mathrm{b}}_{\mathbf{n}} \in \mathbb{B}^{|\mathbf{n}|}$ by letting

$$W^{\mathrm{b}}_{\mathbf{n}}(0) := 0 \ \text{ and } \ W^{\mathrm{b}}_{\mathbf{n}}(j) := \sum_{i=1}^{j} (c(\sigma(i)) - 1), \ \text{ for } \ 1 \le j \le |\mathbf{n}|. \tag{2.3}$$

Note that $W^{\mathrm{b}}_{\mathbf{n}}(|\mathbf{n}|) = -1$. $W^{\mathrm{b}}_{\mathbf{n}}$ is a discrete random process with exchangeable increments. The set of paths taken by $W^{\mathrm{b}}_{\mathbf{n}}$ is

$$\mathbb{B}_{\mathbf{n}} := \Big\{(x(0), x(1), \dots, x(|\mathbf{n}|)) \in \mathbb{B}^{|\mathbf{n}|} : |\{1 \le j \le |\mathbf{n}| : \Delta x(j) = i-1\}| = n(i), \text{ for every } i \ge 0\Big\}. \tag{2.4}$$

From the excursions in $\mathbb{E}^{|\mathbf{n}|}$, we consider those with fixed number of increments of given size, i.e.,

$$\mathbb{E}_{\mathbf{n}} := \mathbb{E}^{|\mathbf{n}|} \cap \mathbb{B}_{\mathbf{n}} = \Big\{(w(0), w(1), \dots, w(|\mathbf{n}|)) \in \mathbb{E}^{|\mathbf{n}|} : |\{1 \le j \le |\mathbf{n}| : \Delta w(j) = i-1\}| = n(i), \text{ for every } i \ge 0\Big\}. \tag{2.5}$$

It is well known that there exists a bijection between $\mathbb{E}_{\mathbf{n}}$ and $\mathbb{T}_{\mathbf{n}}$ (see [28, Lemma 6.3]), and it is also well known that (see [28, Exercise 6.2.1])

$$|\mathbb{T}_{\mathbf{n}}| = \frac{1}{|\mathbf{n}|}\binom{|\mathbf{n}|}{\mathbf{n}} = \frac{1}{|\mathbf{n}|}\frac{|\mathbf{n}|!}{\prod_{i\ge 0} n(i)!}. \tag{2.6}$$

It should be clear that bridges in $\mathbb{B}_{\mathbf{n}}$ are sent to excursions in $\mathbb{E}_{\mathbf{n}}$ by the Vervaat transformation. Moreover, for $w \in \mathbb{E}_{\mathbf{n}}$, the number of $x \in \mathbb{B}_{\mathbf{n}}$ such that $V(x) = w$ is exactly $|\mathbf{n}|$; see [21, Corollary 15.4] or [28, Lemma 6.1].

Let $u(1) \prec \cdots \prec u(|T|)$ be the sequence of vertices of $T \in \mathbb{T}$ in depth-first order (also called pre-order); thus $u(1)$ is the root of T. Set $d_T(i) = d_T(u(i))$, for $i = 1,\dots,|T|$, and call $(d_T(1),\dots,d_T(|T|))$ the pre-order degree sequence of T. For $k \in \mathbb{N}$, it is well-known that a sequence $(d(1),\dots,d(k)) \in \mathbb{N}_0^k$ is the pre-order degree sequence of a tree $T \in \mathbb{T}$ if and only if

$$\sum_{i=1}^{j} d(i) \ge j, \ \text{ for } \ 1 \le j \le k-1, \ \text{ and } \ \sum_{i=1}^{k} d(i) = k-1; \tag{2.7}$$

see [21, Lemma 15.2]. Indeed, T is uniquely determined by its pre-order degree sequence. The depth first queue process (DFQP, or Łukasiewicz path) $W_T = (W_T(i), 0 \le i \le |T|)$ of a tree $T \in \mathbb{T}$, associated to the depth-first ordering $u(1) \prec \cdots \prec u(|T|)$ of its vertices, is defined by $W_T(0) := 0$ and $W_T(i) := W_T(i-$

$1) + d_T(i) - 1$, for $1 \le i \le |T|$. Note that $\Delta W_T(i) = d_T(i) - 1 \ge -1$ for every $1 \le i \le |T|$, with equality if and only if $u(i)$ is a leaf of T. In addition, $W_T(i) \ge 0$ for every $0 \le i \le |T| - 1$, but $W_T(|T|) = -1$, i.e., $W_T \in \mathbb{E}^{|T|}$.

The next (well known) proposition summaries some of the previous definitions and remarks.

Proposition 2.1. *Let* $\mathbf{n}$ *be a degree statistic and let* $\mathcal{T}_{\mathbf{n}} \sim \mathrm{Unif}(\mathbb{T}_{\mathbf{n}})$. *If* $W_{\mathcal{T}_{\mathbf{n}}}$ *is the DFQP of* $\mathcal{T}_{\mathbf{n}}$, *then* $\mathbb{P}(W_{\mathcal{T}_{\mathbf{n}}} = w) = \frac{1}{|\mathbb{T}_{\mathbf{n}}|}$, *for* $w \in \mathbb{E}_{\mathbf{n}}$. *Moreover, if* U *is a uniform random variable on* $\{1,\dots,|\mathbf{n}|\}$ *independent of* $W_{\mathcal{T}_{\mathbf{n}}}$, *then* $(W_{\mathcal{T}_{\mathbf{n}}}, U) \stackrel{\mathrm{d}}{=} (V(W^{\mathrm{b}}_{\mathbf{n}}), \tau_{W^{\mathrm{b}}_{\mathbf{n}}})$. □

Note, in particular, that $\tau_{W^{\mathrm{b}}_{\mathbf{n}}}$ is uniform on $\{1,\dots,|\mathbf{n}|\}$ and independent of $V(W^{\mathrm{b}}_{\mathbf{n}})$.

3 Moment computations

In this section, we compute the moments of the number of fringe subtrees of a uniformly random tree $\mathcal{T}_{\mathbf{n}}$ of $\mathbb{T}_{\mathbf{n}}$, for a degree statistic $\mathbf{n}$. As a warm-up, we use some of the ideas used in [22] to compute the first moment.

Recall that $T \in \mathbb{T}$ is uniquely described by its pre-order degree sequence $(d_T(1),\dots,d_T(|T|))$. Then, for $i = 1,\dots,|T|$, the fringe subtree $T_{u(i)}$ has pre-order degree sequence $(d_T(i),\dots,d_T(i+k-1))$, where $1 \le k \le |T| - i + 1$ is the unique index such that $(d_T(i),\dots,d_T(i+k-1))$ is a pre-order degree sequence of a tree, i.e., it satisfies (2.7). Thus, for $T, T' \in \mathbb{T}$, we can write (1.2) as

$$N_{T'}(T) = \sum_{i=1}^{|T|-|T'|+1} \mathbf{1}_{\{d_T(i)=d_{T'}(1),\dots,d_T(i+|T'|-1))=d_{T'}(|T'|)\}}, \tag{3.1}$$

where the sum is interpreted as 0 when $|T'| > |T|$.

Lemma 3.1. *Let* $\mathbf{n}$ *be a degree statistic and let* $\mathcal{T}_{\mathbf{n}} \sim \mathrm{Unif}(\mathbb{T}_{\mathbf{n}})$. *For* $T \in \mathbb{T}$ *such that* $|\mathbf{n}| \ge |T|$,

$$\mathbb{E}[N_T(\mathcal{T}_{\mathbf{n}})] = \frac{|\mathbf{n}|}{(|\mathbf{n}|)_{|T|}} \prod_{i \ge 0} (n(i))_{n_T(i)}. \tag{3.2}$$

Proof. If $n_T(i) > n(i)$ for some $i \ge 0$, then both sides of (3.2) are 0. Assume therefore that $n_T(i) \le n(i)$ for all $i \ge 0$. Then, $|T| \le |\mathcal{T}_{\mathbf{n}}| = |\mathbf{n}|$. Let $W_{\mathcal{T}_{\mathbf{n}}}$ be the DFQP of $\mathcal{T}_{\mathbf{n}}$ and $(d_{\mathcal{T}_{\mathbf{n}}}(1),\dots,d_{\mathcal{T}_{\mathbf{n}}}(|\mathbf{n}|))$ its pre-order degree sequence. Note that $d_{\mathcal{T}_{\mathbf{n}}}(i) = \Delta W_{\mathcal{T}_{\mathbf{n}}}(i) + 1$, for $i = 1,\dots,|\mathbf{n}|$. Let $(d_T(1),\dots,d_T(|T|))$ be the pre-order degree sequence of T. Hence, by (3.1),

$$N_T(\mathcal{T}_{\mathbf{n}}) = \sum_{i=1}^{|\mathbf{n}|-|T|+1} \mathbf{1}_{\{\Delta W_{\mathcal{T}_{\mathbf{n}}}(i)=d_T(1)-1,\dots,\Delta W_{\mathcal{T}_{\mathbf{n}}}(i+|T|-1)=d_T(|T|)-1\}}. \tag{3.3}$$

For $1 \le i \le |\mathbf{n}| - |T| + 1$ and $(y(1)\dots,y(i-1),y(i+|T|),\dots,y(|\mathbf{n}|)) \in \mathbb{Z}^{|\mathbf{n}|-|T|}$, define the walk

$$w^{y}_{i,T}(0) = 0 \text{ and } w^{y}_{i,T}(j) = \sum_{r=1}^{j} (y(r) - 1), \text{ for } 1 \le j \le |\mathbf{n}|, \tag{3.4}$$

where $y(i) = d_T(1), \ldots, y(i+|T|-1) = d_T(|T|)$. In particular, $\Delta w^y_{i,T}(i+j-1) = d_T(j)-1$, for $j = 1, \ldots, |T|$. We then consider the set of admissible excursions obtained in this way:

$$\mathbb{A}_{i,T} = \{w^y_{i,T} : (y(1)\ldots, y(i-1), y(i+|T|), \ldots, y(|\mathbf{n}|)) \in \mathbb{Z}^{|\mathbf{n}|-|T|} \text{ and } w^y_{i,T} \in \mathbb{E}_{\mathbf{n}}\}, \tag{3.5}$$

i.e., $\mathbb{A}_{i,T}$ is the set of excursions in $\mathbb{E}_{\mathbf{n}}$ that code trees in $\mathbb{T}_{\mathbf{n}}$ with a fringe subtree with pre-order degree sequence $(d_T(1), \ldots, d_T(|T|))$ that is rooted at its i-th vertex in depth-first order. Let $\tilde{\mathbf{n}} = (\tilde{n}(i))_{i\geq 0}$ be given by $\tilde{n}(0) = n(0) - n_T(0) + 1$ and $\tilde{n}(i) = n(i) - n_T(i)$ for $i \geq 1$. If we instead of inserting the degree sequence of T as above, insert only $y(i) = 0$ (corresponding to a leaf), and then relabel $y(j+|T|)$ as $y(j+1)$ for $i \leq j \leq |\mathbf{n}| - |T|$, we obtain a bijection between $\mathbb{A}_{i,T}$ and the excursions in $\mathbb{E}_{\tilde{\mathbf{n}}}$ that correspond to a tree with a leaf as its i-th vertex. Thus, due to the bijection between $\mathbb{E}_{\tilde{\mathbf{n}}}$ and $\mathbb{T}_{\tilde{\mathbf{n}}}$, we see that

$$\sum_{i=1}^{|\mathbf{n}|-|T|+1} |\mathbb{A}_{i,T}| = |\mathbb{T}_{\tilde{\mathbf{n}}}| \cdot \tilde{n}(0). \tag{3.6}$$

By Proposition 2.1 and (3.3)–(3.6), this yields

$$\mathbb{E}[N_T(\mathcal{T}_{\mathbf{n}})] = \sum_{i=1}^{|\mathbf{n}|-|T|+1} \frac{|\mathbb{A}_{i,T}|}{|\mathbb{T}_{\mathbf{n}}|} = \frac{|\mathbb{T}_{\tilde{\mathbf{n}}}| \cdot \tilde{n}(0)}{|\mathbb{T}_{\mathbf{n}}|}, \tag{3.7}$$

and the result (3.2) follows by (2.6). □

Lemma 3.1 can be generalized to joint factorial moments of the random variables $N_{T_1}(\mathcal{T}_{\mathbf{n}}), \ldots, N_{T_m}(\mathcal{T}_{\mathbf{n}})$, for $m \geq 1$ and a sequence of distinct rooted plane trees $T_1, \ldots, T_m \in \mathbb{T}$. Before that, we need to introduce some notation. For $1 \leq i, j \leq m$, let

$$\tau_{ij} := N_{T_i}(T_j)\mathbf{1}_{\{i\neq j\}} \tag{3.8}$$

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 $T_1, \ldots, T_m$ according to their sizes, the matrix $(\tau_{ij})^m_{i,j=1}$ is strictly triangular.)

For $q_1, \ldots, q_m \in \mathbb{N}_0$, 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 $q := q_1 + \cdots + q_m$ 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 $b = (b_1, \ldots, b_m) \in \mathbb{N}_0^m$, 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 $1 \leq i \leq m$. We thus have

$$\mathbb{E}\Big[(N_{T_1}(\mathcal{T}_{\mathbf{n}}))_{q_1} \cdots (N_{T_m}(\mathcal{T}_{\mathbf{n}}))_{q_m}\Big] = \sum_{b\in\mathbb{N}_0^m} \mathbb{E}[S_b(\mathcal{T}_{\mathbf{n}})]. \tag{3.9}$$

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 3.2. *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 $|\mathbf{n}| \ge \sum_{i=1}^m (q_i|T_i| - q_i) + 1$. Then*

$$\mathbb{E}[S_b(\mathcal{T}_{\mathbf{n}})] = \frac{|\mathbf{n}|}{(|\mathbf{n}|)_{1+\sum_{j=1}^m (q_j - b_j)(|T_j|-1)}} \prod_{i \ge 0} (n(i))_{\sum_{j=1}^m (q_j - b_j) n_{T_j}(i)} \prod_{j=1}^m \frac{(q_j)_{b_j} \left(\sum_{k=1}^m (q_k - b_k)\tau_{jk}\right)_{b_j}}{b_j!}, \tag{3.10}$$

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 sides of (3.10) 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

$$\tilde{n}(i) := \begin{cases} n(i) - \sum_{j=1}^m q_j n_{T_j}(i), & i \ge 1, \\ n(0) - \sum_{j=1}^m q_j n_{T_j}(0) + \sum_{j=1}^m q_j, & i = 0, \end{cases} \tag{3.11}$$

and has size

$$|\tilde{\mathbf{n}}| := \sum_{i \ge 0} \tilde{n}(i) = |\mathbf{n}| - \sum_{j=1}^m q_j(|T_j| - 1). \tag{3.12}$$

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 (3.11) and a sequence of q marked leaves. If we ignore the marks, the number of possible reduced trees is given by (2.6) with the degree statistic $\tilde{\mathbf{n}}$ in (3.11). 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.,

$$\frac{(|\tilde{\mathbf{n}}| - 1)!}{\prod_{i \ge 0} \tilde{n}(i)!} (\tilde{n}(0))_{\sum_{j=1}^m q_j} = \frac{(|\tilde{\mathbf{n}}| - 1)!}{\prod_{i \ge 0} (n(i) - \sum_{j=1}^m q_j n_{T_j}(i))!}. \tag{3.13}$$

By dividing with $|\mathbb{T}_{\mathbf{n}}|$, which is given by (2.6), and using (3.12), we find

$$\mathbb{E}[S_0(\mathcal{T}_{\mathbf{n}})] = \frac{1}{(|\mathbf{n}| - 1)_{\sum_{j=1}^m q_j(|T_j|-1)}} \prod_{i \ge 0} (n(i))_{\sum_{j=1}^m q_j n_{T_j}(i)}. \tag{3.14}$$

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

$$\prod_{j=1}^m \frac{(q_j)_{b_j}\left(\sum_{k=1}^m (q_k - b_k)\tau_{jk}\right)_{b_j}}{b_j!}. \tag{3.15}$$

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 (3.11)–(3.13). Hence, we obtain (3.10), extending (3.14). $\square$

We record two important special cases of the computation above.

Lemma 3.3. *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* $|\mathbf{n}| \ge q|T| - q + 1$,

$$\mathbb{E}[(N_T(\mathcal{T}_{\mathbf{n}}))_q] = \frac{|\mathbf{n}|}{(|\mathbf{n}|)_{q|T|-q+1}} \prod_{i \ge 0} (n(i))_{q n_T(i)}. \tag{3.16}$$

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

$$\mathbb{E}[N_T(\mathcal{T}_{\mathbf{n}}) N_{T'}(\mathcal{T}_{\mathbf{n}})] = N_T(T')\,\mathbb{E}[N_{T'}(\mathcal{T}_{\mathbf{n}})] + N_{T'}(T)\,\mathbb{E}[N_T(\mathcal{T}_{\mathbf{n}})] + \frac{|\mathbf{n}|}{(|\mathbf{n}|)_{|T|+|T'|-1}} \prod_{i \ge 0} (n(i))_{n_T(i)+n_{T'}(i)}. \tag{3.17}$$

Proof. (i): This is the case $m = 1$ of (3.9) and (3.10), when we consider only one tree T_1. In this case, there are no bound fringe trees, and thus we only have to consider $b = 0$ in (3.9). Taking $b_1 = 0$ (and $q_1 = q$) in (3.10) yields (3.16).

(ii): This is the case $m = 2$ and $q_1 = q_2 = 1$ of (3.9). The possible vectors $b = (b_1, b_2)$ are $(1,0)$, $(0,1)$, and $(0,0)$, and it is easily verified that taking these three vectors in (3.10), and using (3.16) with $q = 1$ in two cases, yields the three terms on the right-hand side of (3.17). $\square$

4 Proof of Theorems 1.3 and 1.6

In this section we prove Theorems 1.3 and 1.6 (in opposite order). In what follows we will frequently use the following well-known estimate.

Lemma 4.1. *If* $x \ge 1$ *is a real number and* $0 \le k \le x/2$ *is an integer, then*

$$(x)_k = x^k \exp\left(-\frac{k(k-1)}{2x} + O\left(\frac{k^3}{x^2}\right)\right). \tag{4.1}$$

Proof. Since $\ln(1-y) = -y + O(y^2)$ for $0 \le y \le 1/2$, the result follows from the identity

$$\frac{(x)_k}{x^k} = \prod_{i=0}^{k-1} \frac{x-i}{x} = \exp\left(\sum_{i=0}^{k-1} \ln\left(1-\frac{i}{x}\right)\right). \tag{4.2}$$

□

Proof of Theorem 1.6. Note first the trivial bound

$$N_T(\mathcal{T}_{\mathbf{n}}) \le \frac{n(i)}{n_T(i)} \le n(i), \qquad i \in \mathcal{D}(T), \tag{4.3}$$

since the copies of T in $\mathcal{T}_{\mathbf{n}}$ are distinct. Furthermore, by (1.6) and (1.4),

$$|\mathbf{n}|\pi_{\mathbf{p}(\mathbf{n})}(T) \le |\mathbf{n}|p_i(\mathbf{n}) = n(i), \qquad i \in \mathcal{D}(T). \tag{4.4}$$

Hence, (1.16) is trivial if $n(i) = O(1)$ for some $i \in \mathcal{D}(T)$. In particular, we may in the sequel assume $n(i) \ge 2n_T(i)$ for every $i \ge 0$, and thus $|\mathbf{n}| \ge 2|T|$. Then, by (3.2) and Lemma 4.1,

$$\begin{aligned}
\mathbb{E}\, N_T(\mathcal{T}_{\mathbf{n}}) &= |\mathbf{n}|^{1-|T|} \prod_{i\in\mathcal{D}(T)} n(i)^{n_T(i)} \cdot \exp\left(\frac{|T|(|T|-1)}{2|\mathbf{n}|} - \sum_{i\in\mathcal{D}(T)} \frac{n_T(i)(n_T(i)-1)}{2n(i)} + O\Big(\sum_{i\in\mathcal{D}(T)} \frac{1}{n(i)^2}\Big)\right) \\
&= |\mathbf{n}|\pi_{\mathbf{p}(\mathbf{n})}(T) \cdot \exp\left(\frac{|T|(|T|-1)}{2|\mathbf{n}|} - \sum_{i\in\mathcal{D}(T)} \frac{n_T(i)(n_T(i)-1)}{2n(i)} + O\Big(\sum_{i\in\mathcal{D}(T)} \frac{1}{n(i)^2}\Big)\right),
\end{aligned} \tag{4.5}$$

which implies (1.16) by (4.4).

Similarly, taking $q = 2$ in (3.16), and now assuming as we may $n(i) \ge 4n_T(i)$ for every $i \ge 0$,

$$\begin{aligned}
\mathbb{E}(N_T(\mathcal{T}_{\mathbf{n}}))_2 &= \frac{|\mathbf{n}|}{(|\mathbf{n}|)_{2|T|-1}} \prod_{i\in\mathcal{D}(T)} (n(i))_{2n_T(i)} \\
&= |\mathbf{n}|^{2-2|T|} \prod_{i\in\mathcal{D}(T)} n(i)^{2n_T(i)} \cdot \exp\left(\frac{(2|T|-1)(2|T|-2)}{2|\mathbf{n}|} - \sum_{i\in\mathcal{D}(T)} \frac{2n_T(i)(2n_T(i)-1)}{2n(i)} + O\Big(\sum_{i\in\mathcal{D}(T)} \frac{1}{n(i)^2}\Big)\right) \\
&= \Big(|\mathbf{n}|\pi_{\mathbf{p}(\mathbf{n})}(T)\Big)^2 \cdot \exp\left(\frac{(2|T|-1)(|T|-1)}{|\mathbf{n}|} - \sum_{i\in\mathcal{D}(T)} \frac{n_T(i)(2n_T(i)-1)}{n(i)} + O\Big(\sum_{i\in\mathcal{D}(T)} \frac{1}{n(i)^2}\Big)\right),
\end{aligned} \tag{4.6}$$

Hence, using also (4.5),

$$\mathbb{E}(N_T(\mathcal{T}_{\mathbf{n}}))_2 = \Big(\mathbb{E}\, N_T(\mathcal{T}_{\mathbf{n}})\Big)^2 \cdot \exp\left(\frac{(|T|-1)^2}{|\mathbf{n}|} - \sum_{i\in\mathcal{D}(T)} \frac{n_T(i)^2}{n(i)} + O\Big(\sum_{i\in\mathcal{D}(T)} \frac{1}{n(i)^2}\Big)\right). \tag{4.7}$$

Consequently, using (1.16) and noting that $\mathbb{E}\, N_T(\mathcal{T}_{\mathbf{n}}) = O(n(i))$ for $i \in \mathcal{D}(T)$ by (1.16) and (4.4),

$$\mathrm{Var}[N_T(\mathcal{T}_{\mathbf{n}})] = \mathbb{E}(N_T(\mathcal{T}_{\mathbf{n}}))_2 + \mathbb{E}\, N_T(\mathcal{T}_{\mathbf{n}}) - \Big(\mathbb{E}\, N_T(\mathcal{T}_{\mathbf{n}})\Big)^2$$

$$= \Big(\mathbb{E} N_T(\mathcal{T}_{\mathbf{n}})\Big)^2 \cdot \left(\frac{(|T|-1)^2}{|\mathbf{n}|} - \sum_{i\in\mathcal{D}(T)} \frac{n_T(i)^2}{n(i)} \right) + \mathbb{E} N_T(\mathcal{T}_{\mathbf{n}}) + O(1)$$

$$= \Big(|\mathbf{n}|\pi_{\mathbf{p}(\mathbf{n})}(T)\Big)^2 \cdot \left(\frac{(|T|-1)^2}{|\mathbf{n}|} - \sum_{i\in\mathcal{D}(T)} \frac{n_T(i)^2}{n(i)} \right) + |\mathbf{n}|\pi_{\mathbf{p}(\mathbf{n})}(T) + O(1), \tag{4.8}$$

which yields (1.17) by the definitions (1.9), (1.8) and (1.4).

For the proof of (1.18) we use (3.17). The first two terms are handled by (1.16), and the final term is treated as in (4.6)–(4.8) with mainly notational differences; we omit the details. ☐

Proof of Theorem 1.3. By Condition 1.1, 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 (1.16).

Moreover, it follows from (1.8)–(1.9) that $\gamma_{\mathbf{p}(\mathbf{n}_\kappa)}(T,T) = O(1)$ (for a fixed T), and thus (1.17) yields $\operatorname{Var} N_T(\mathcal{T}_{\mathbf{n}_\kappa}) = O(|\mathbf{n}_\kappa|)$. Therefore, (ii) follows from (i) and Chebyshev's inequality. ☐

5 Proof of Theorems 1.5 and 1.7

We have now all the ingredients to prove Theorems 1.5 and Theorem 1.7. Theorem 1.5 is essentially a special case Theorem 1.7, combined with the already proved Theorem 1.6. Nevertheless, in order to focus on the main ideas, we give first a separate proof of Theorem 1.5, and then the rather small modifications required for the more technical general version in Theorem 1.7.

Proof of Theorem 1.5. First note that Condition 1.1 implies

$$\pi_{\mathbf{p}(\mathbf{n}_\kappa)}(T_i) \to \pi_{\mathbf{p}}(T_i) \quad \text{and} \quad \gamma_{\mathbf{p}(\mathbf{n}_\kappa)}(T_i,T_j) \to \gamma_{\mathbf{p}}(T_i,T_j), \qquad \text{for } 1 \leq i,j \leq m. \tag{5.1}$$

Hence, (1.11)–(1.13) follow from (1.16)–(1.18) in Theorem 1.6.

We next prove the asymptotic normality result in (1.14). Note first that (1.16) implies that it does not matter whether we use $\mathbb{E}[N_{T_i}(\mathcal{T}_{\mathbf{n}_\kappa})]$ or $\mu_{\mathbf{n}_\kappa}(T_i) = |\mathbf{n}_\kappa|\pi_{\mathbf{p}(\mathbf{n}_\kappa)}(T_i)$ in (1.14).

If $\pi_{\mathbf{p}}(T_i) = 0$, for some $1 \leq i \leq m$, then it follows from (1.9) that $\gamma_{\mathbf{p}}(T_i,T_i) = 0$, and thus (1.12) yields $\operatorname{Var}[N_{T_i}(\mathcal{T}_{\mathbf{n}_\kappa})] = o(|\mathbf{n}_\kappa|)$; consequently, (1.16) and Chebyshev's inequality yield, as $\kappa \to \infty$,

$$\frac{N_{T_i}(\mathcal{T}_{\mathbf{n}_\kappa}) - \mathbb{E}[N_{T_i}(\mathcal{T}_{\mathbf{n}_\kappa})]}{\sqrt{|\mathbf{n}_\kappa|}} \xrightarrow{\text{p}} 0. \tag{5.2}$$

Hence, convergence of the i-th component in (1.14) 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 (1.10), noting that if $N_{T_i}(T_j) > 0$ then also $\pi_{\mathbf{p}}(T_j) = 0$. Thus, we may ignore all i in (1.14) with $\pi_{\mathbf{p}}(T_i) = 0$ and show (joint) convergence for the remaining ones, because then (1.14) in general will follow from [3, 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 $T_1,\dots,T_m$ are distinct.

To see the main idea of the proof, consider first the univariate case $m = 1$. (The impatient reader may skip this and proceed directly to the general case.) We omit the index 1 and write T instead of T_1. In this case, we can use the Gao–Wormald theorem [13, Theorem 1] and the following estimate. For any $q_\kappa = O(|\mathbf{n}_\kappa|^{1/2})$, (3.16) and Lemma 4.1 yield, recalling the definitions (1.4), (1.6), (1.8), (1.9), and (1.19) of $p_i(\mathbf{n})$, $\pi_{\mathbf{p}}(T)$, $\eta_{\mathbf{p}}(T,T)$, $\gamma_{\mathbf{p}}(T,T)$, and $\mu_{\mathbf{n}_\kappa}(T)$,

$$\begin{aligned}
\mathbb{E}[(N_T(\mathcal{T}_{\mathbf{n}_\kappa}))_{q_\kappa}] &= \frac{\prod_{i\geq 0} n_\kappa(i)^{q_\kappa n_T(i)}}{|\mathbf{n}_\kappa|^{q_\kappa(|T|-1)}} \exp\left(\frac{\left(q_\kappa(|T|-1)\right)^2}{2|\mathbf{n}_\kappa|} - \sum_{i\geq 0} \frac{\left(q_\kappa n_T(i)\right)^2}{2n_\kappa(i)} + o(1)\right) \\
&= |\mathbf{n}_\kappa|^{q_\kappa} \prod_{i\geq 0} p_i(\mathbf{n}_\kappa)^{q_\kappa n_T(i)} \exp\left(\frac{\left(q_\kappa(|T|-1)\right)^2}{2|\mathbf{n}_\kappa|} - \sum_{i\geq 0} \frac{\left(q_\kappa n_T(i)\right)^2}{2n_\kappa(i)} + o(1)\right) \\
&= \left(|\mathbf{n}_\kappa| \pi_{\mathbf{p}(\mathbf{n}_\kappa)}(T)\right)^{q_\kappa} \exp\left(\frac{q_\kappa^2}{2|\mathbf{n}_\kappa|} \eta_{\mathbf{p}(\mathbf{n}_\kappa)}(T,T) + o(1)\right) \\
&= \mu_{\mathbf{n}_\kappa}(T)^{q_\kappa} \exp\left(\frac{(\gamma_{\mathbf{p}(\mathbf{n}_\kappa)}(T,T) - \pi_{\mathbf{p}(\mathbf{n}_\kappa)}(T))|\mathbf{n}_\kappa|}{2\mu_{\mathbf{n}_\kappa}(T)^2} q_\kappa^2 + o(1)\right) \\
&= \mu_{\mathbf{n}_\kappa}(T)^{q_\kappa} \exp\left(\frac{\gamma_{\mathbf{p}}(T,T)|\mathbf{n}_\kappa| - \mu_{\mathbf{n}_\kappa}(T)}{2\mu_{\mathbf{n}_\kappa}(T)^2} q_\kappa^2 + o(1)\right).
\end{aligned} \tag{5.3}$$

If $\gamma_{\mathbf{p}}(T,T) > 0$, we may now apply the Gao–Wormald theorem [13, Theorem 1] with $\mu_\kappa := \mu_{\mathbf{n}_\kappa}(T)$ and $\sigma_\kappa^2 := \gamma_{\mathbf{p}}(T,T)|\mathbf{n}_\kappa|$ and conclude (1.15), which by (1.12) is equivalent to (1.14) (with $m = 1$). The case $\gamma_{\mathbf{p}}(T,T) = 0$ is trivial, since then (1.12) implies (5.2). 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 Theorem A.1 of the Gao–Wormald theorem.

The general case follows similarly by a multidimensional version of the Gao–Wormald theorem, which we state and prove as Theorem A.1 in Appendix A, together with the following estimates of (joint) factorial moments. 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 (3.9) and estimate the terms separately.

Write for convenience $\mu_{i\kappa} := \mu_{\mathbf{n}_\kappa}(T_i) = |\mathbf{n}_\kappa| \pi_{\mathbf{p}(\mathbf{n}_\kappa)}(T_i)$, for $1 \leq i \leq m$. Note that our assumption $\pi_{\mathbf{p}}(T_i) > 0$ and (5.1) imply that (possibly ignoring some small κ)

$$\mu_{i\kappa} = \Theta(|\mathbf{n}_\kappa|), \qquad 1 \leq i \leq m. \tag{5.4}$$

Furthermore, for every $i \in \bigcup_{j=1}^m \mathcal{D}(T_j)$, we have by assumption $p_i > 0$, and thus Condition 1.1 implies

$$n_\kappa(i) = \Theta(|\mathbf{n}_\kappa|). \tag{5.5}$$

Let $q_{1\kappa}, \dots, q_{m\kappa} \in \mathbb{N}_0$ be such that

$$q_{i\kappa} = O(|\mathbf{n}_\kappa|^{1/2}), \qquad 1 \leq i \leq m. \tag{5.6}$$

Let $b = (b_1, \dots, b_m) \in \mathbb{N}_0^m$ be a fixed sequence. Assume first that $q_{j\kappa} \geq b_j$ for $1 \leq j \leq m$. Then we deduce from (3.10) in Lemma 3.2 and Lemma 4.1, using (5.4)–(5.6) and recalling (1.4), (1.6) and (1.19),

$$\begin{aligned}
\mathbb{E}\,S_b(\mathcal{T}_{\mathbf{n}_\kappa}) &= \frac{\prod_{i\geq 0} n_\kappa(i)^{\sum_{j=1}^m (q_{j\kappa}-b_j)n_{T_j}(i)}}{|\mathbf{n}_\kappa|^{\sum_{j=1}^m (q_{j\kappa}-b_j)(|T_j|-1)}} \prod_{j=1}^m \frac{(q_{j\kappa}+O(1))^{b_j}\big(\sum_{k=1}^m (q_{k\kappa}-b_k)\tau_{jk}+O(1)\big)^{b_j}}{b_j!} \\
&\qquad \times \exp\left(\frac{\big(\sum_{j=1}^m (q_{j\kappa}-b_j)(|T_j|-1)\big)^2}{2|\mathbf{n}_\kappa|} - \sum_{i\geq 0}\frac{\big(\sum_{j=1}^m (q_{j\kappa}-b_j)n_{T_j}(i)\big)^2}{2n_\kappa(i)} + o(1)\right) \\
&= |\mathbf{n}_\kappa|^{\sum_{j=1}^m (q_{j\kappa}-b_j)} \prod_{i\geq 0}\prod_{j=1}^m p_i(\mathbf{n}_\kappa)^{(q_{j\kappa}-b_j)n_{T_j}(i)} \prod_{j=1}^m \frac{\big(q_{j\kappa}\sum_{k=1}^m q_{k\kappa}\tau_{jk} + O(q_{j\kappa}+\sum_{k=1}^m \tau_{jk}q_{k\kappa})\big)^{b_j}}{b_j!} \\
&\qquad \times \exp\left(\frac{\big(\sum_{j=1}^m q_{j\kappa}(|T_j|-1)\big)^2}{2|\mathbf{n}_\kappa|} - \sum_{i\geq 0}\frac{\big(\sum_{j=1}^m q_{j\kappa}n_{T_j}(i)\big)^2}{2n_\kappa(i)} + o(1)\right) \\
&= |\mathbf{n}_\kappa|^{\sum_{j=1}^m (q_{j\kappa}-b_j)} \prod_{j=1}^m \pi_{\mathbf{p}(\mathbf{n}_\kappa)}(T_j)^{q_{j\kappa}-b_j} \prod_{j=1}^m \frac{\big(q_{j\kappa}\sum_{k=1}^m q_{k\kappa}\tau_{jk} + O(\mu_{j\kappa}^{1/2})\big)^{b_j}}{b_j!} \\
&\qquad \times \exp\left(\frac{\big(\sum_{j=1}^m q_{j\kappa}(|T_j|-1)\big)^2}{2|\mathbf{n}_\kappa|} - \sum_{i\geq 0}\frac{\big(\sum_{j=1}^m q_{j\kappa}n_{T_j}(i)\big)^2}{2n_\kappa(i)} + o(1)\right) \\
&= \prod_{j=1}^m \mu_{j\kappa}^{q_{j\kappa}} \prod_{j=1}^m \frac{\big(q_{j\kappa}\sum_{k=1}^m q_{k\kappa}\tau_{jk}/\mu_{j\kappa} + o(1)\big)^{b_j}}{b_j!} \\
&\qquad \times \exp\left(\frac{\big(\sum_{j=1}^m q_{j\kappa}(|T_j|-1)\big)^2}{2|\mathbf{n}_\kappa|} - \sum_{i\geq 0}\frac{\big(\sum_{j=1}^m q_{j\kappa}n_{T_j}(i)\big)^2}{2n_\kappa(i)} + o(1)\right).
\end{aligned} \tag{5.7}$$

On the right-hand side, by (5.4)–(5.6), each factor in the second product is $O(1)$, and so is the exponential factor. Consequently, (5.7) yields

$$\frac{\mathbb{E}\,S_b(\mathcal{T}_{\mathbf{n}_\kappa})}{\prod_{j=1}^m \mu_{j\kappa}^{q_{j\kappa}}} = \prod_{j=1}^m \frac{\big(q_{j\kappa}\sum_{k=1}^m q_{k\kappa}\tau_{jk}/\mu_{j\kappa}\big)^{b_j}}{b_j!}\cdot\exp\left(\frac{\big(\sum_{j=1}^m q_{j\kappa}(|T_j|-1)\big)^2}{2|\mathbf{n}_\kappa|} - \sum_{i\geq 0}\frac{\big(\sum_{j=1}^m q_{j\kappa}n_{T_j}(i)\big)^2}{2n_\kappa(i)}\right) + o(1). \tag{5.8}$$

This is trivially true also if $q_{j\kappa} < b_j$ for some $1 \leq j \leq m$, since then $S_b(\mathcal{T}_{\mathbf{n}_\kappa}) = 0$ and the first term on the right-hand side of (5.8) then is easily seen to be $o(1)$. Hence, (5.8) holds for every fixed $b \in \mathbb{N}_0^m$, uniformly for all $(q_{j\kappa})_{j=1}^m$ that satisfy (5.6).

Furthermore, a simple variant of this calculation shows that, for each constant $C > 0$, uniformly for all $\kappa \geq 1$, $b = (b_1, \dots, b_m) \in \mathbb{N}_0^m$ and $q_{1\kappa}, \dots, q_{m\kappa} \in \mathbb{N}_0$ such that (5.6) holds,

$$\mathbb{E}\,S_b(\mathcal{T}_{\mathbf{n}_\kappa}) \leq \frac{\prod_{i\geq 0} n_\kappa(i)^{\sum_{j=1}^m (q_{j\kappa}-b_j)n_{T_j}(i)}}{|\mathbf{n}_\kappa|^{\sum_{j=1}^m (q_{j\kappa}-b_j)(|T_j|-1)}} \prod_{j=1}^m \frac{q_{j\kappa}^{b_j}\big(\sum_{k=1}^m q_{k\kappa}\tau_{jk}\big)^{b_j}}{b_j!}\exp(O(1))$$

$$\leq C_1 \prod_{j=1}^{m} \mu_{j\kappa}^{q_{j\kappa}} \prod_{j=1}^{m} \frac{C_2^{b_j}}{b_j!}, \tag{5.9}$$

for some constants $C_1 > 0$ and $C_2 > 0$ (depending on C). Equivalently,

$$\frac{\mathbb{E}[S_b(\mathcal{T}_{\mathbf{n}_\kappa})]}{\prod_{j=1}^{m} \mu_{j\kappa}^{q_{j\kappa}}} \leq C_1 \prod_{j=1}^{m} \frac{C_2^{b_j}}{b_j!}. \tag{5.10}$$

By (5.4)–(5.6), the same estimate holds for the first term on the right-hand side of (5.8), and thus it holds also for the $o(1)$ term there.

Now sum (5.8) over all $b = (b_1, \dots, b_m) \in \mathbb{N}_0^m$. The sum of the error terms $o(1)$ is $o(1)$ by dominated convergence justified by (5.10) and the comments after it. Hence, (3.9) and (5.8) yield, uniformly for all $q_{1\kappa}, \dots, q_{m\kappa}$ satisfying (5.6), recalling $n_\kappa(i) = |\mathbf{n}_\kappa| p_i(\mathbf{n}_\kappa)$ for $i \geq 0$,

$$\begin{aligned}
\frac{\mathbb{E}[(N_{T_1}(\mathcal{T}_{\mathbf{n}}))_{q_{1\kappa}} \cdots (N_{T_m}(\mathcal{T}_{\mathbf{n}}))_{q_{m\kappa}}]}{\prod_{j=1}^{m} \mu_{j\kappa}^{q_{j\kappa}}} &= \prod_{j=1}^{m} \exp\left(\frac{q_{j\kappa} \sum_{k=1}^{m} q_{k\kappa} \tau_{jk}}{\mu_{j\kappa}} \right) \\
&\quad \times \exp\left(\frac{\left(\sum_{j=1}^{m} q_{j\kappa}(|T_j| - 1)\right)^2}{2|\mathbf{n}_\kappa|} - \sum_{i \geq 0} \frac{\left(\sum_{j=1}^{m} q_{j\kappa} n_{T_j}(i)\right)^2}{2|\mathbf{n}_\kappa| p_i(\mathbf{n}_\kappa)} \right) + o(1) \\
&= \exp\left(\frac{1}{2} \sum_{j,k=1}^{m} \frac{r_\kappa(j,k)|\mathbf{n}_\kappa|}{\mu_{j\kappa}\mu_{k\kappa}} q_{j\kappa} q_{k\kappa} \right) + o(1),
\end{aligned} \tag{5.11}$$

where, by a simple calculation recalling (1.19) and (1.8),

$$r_\kappa(j,k) := 2\tau_{jk} \pi_{\mathbf{p}(\mathbf{n}_\kappa)}(T_k) + \eta_{\mathbf{p}(\mathbf{n}_\kappa)}(T_j, T_k) \pi_{\mathbf{p}(\mathbf{n}_\kappa)}(T_j) \pi_{\mathbf{p}(\mathbf{n}_\kappa)}(T_k), \qquad 1 \leq j,k \leq m. \tag{5.12}$$

In (5.11), we may replace $r_\kappa(j,k)$ by the symmetrization $\widetilde{r}_\kappa(j,k) := \frac{1}{2}(r_\kappa(j,k) + r_\kappa(k,j))$. Comparing (5.12) and (1.9)–(1.10), using (3.8) and treating the cases $j = k$ and $j \neq k$ separately, we obtain

$$\widetilde{r}_\kappa(j,k) = \gamma_{\mathbf{p}(\mathbf{n}_\kappa)}(T_j, T_k) - \delta_{jk} \pi_{\mathbf{p}(\mathbf{n}_\kappa)}(T_j), \qquad 1 \leq j,k \leq m. \tag{5.13}$$

Define $\sigma_{j\kappa} := |\mathbf{n}_\kappa|^{1/2}$ for every $1 \leq j \leq m$; then (5.11) yields, using (5.13) and (5.1),

$$\mathbb{E}\Big[(N_{T_1}(\mathcal{T}_{\mathbf{n}}))_{q_{1\kappa}} \cdots (N_{T_m}(\mathcal{T}_{\mathbf{n}}))_{q_{m\kappa}}\Big] = \prod_{j=1}^{m} \mu_{j\kappa}^{q_{j\kappa}} \exp\left(\frac{1}{2} \sum_{j,k=1}^{m} \frac{\gamma_{\mathbf{p}}(T_j, T_k)\sigma_{j\kappa}\sigma_{k\kappa} - \delta_{jk}\mu_{j\kappa}}{\mu_{j\kappa}\mu_{k\kappa}} q_{j\kappa} q_{k\kappa} + o(1) \right), \tag{5.14}$$

uniformly in all $q_{1\kappa}, \dots, q_{m\kappa}$ that satisfy (5.6). We apply Theorem A.1, and note that (5.14) is the condition (A.2) (with obvious changes of notation); furthermore, by (5.1), our choices $\mu_{i\kappa} := |\mathbf{n}_\kappa| \pi_{\mathbf{p}(\mathbf{n}_\kappa)}(T_i)$ and $\sigma_{i\kappa} := |\mathbf{n}_\kappa|^{1/2}$ satisfy (A.1). Hence, Theorem A.1 yields (1.14).

Finally, the last assertion of Theorem 1.5 is proved in Lemma 5.1 below. $\square$

Lemma 5.1. *Suppose that* $\mathbf{p} = (p_i)_{i\geq 0} \in \mathcal{P}_1(\mathbb{N}_0)$, *and let* $T \in \mathbb{T}$ *with* $|T| > 1$. *If* $\pi_{\mathbf{p}}(T) > 0$, *then* $\gamma_{\mathbf{p}}(T,T) > 0$.

Recall that the assumption $\pi_{\mathbf{p}}(T) > 0$ is equivalent to $p_i > 0$ for all $i \in \mathcal{D}(T)$.

Proof. Suppose that $\pi_{\mathbf{p}}(T) > 0$ but $\gamma_{\mathbf{p}}(T,T) = 0$. By (1.9) and (1.8), this means

$$0 = \frac{\gamma_{\mathbf{p}}(T,T)}{\pi_{\mathbf{p}}(T)^2} = \frac{1}{\pi_{\mathbf{p}}(T)} + \eta_{\mathbf{p}}(T,T) = \frac{1}{\pi_{\mathbf{p}}(T)} + (|T|-1)^2 - \sum_{i\in\mathcal{D}(T)} \frac{n_T(i)^2}{p_i}. \tag{5.15}$$

Furthermore, since (1.14) in Theorem 1.5 applies, for any $m \geq 1$ and trees $T_1,\dots,T_m \in \mathbb{T}$, the matrix $(\gamma_{\mathbf{p}}(T_i,T_j))_{i,j=1}^m$ is a covariance matrix, and thus positive semidefinite. Hence, the Cauchy–Schwarz inequality holds for $\gamma_{\mathbf{p}}(T_i,T_j)$. In particular, for any tree $T' \in \mathbb{T}$,

$$|\gamma_{\mathbf{p}}(T,T')| \leq \gamma_{\mathbf{p}}(T,T)^{1/2}\gamma_{\mathbf{p}}(T',T')^{1/2} = 0. \tag{5.16}$$

Fix $d \in \mathcal{D}(T)$ with $d \geq 1$. Let $1 \leq k \leq d$ and define $T_{d,k}$ to be the tree that has a root of degree d, the first k children of the root are copies of T, and the remaining $d-k$ children of the root are leaves. Thus $|T_{d,k}| = 1 + k|T| + d - k$ and

$$n_{T_{d,k}}(i) = kn_T(i) + (d-k)\mathbf{1}_{\{i=0\}} + \mathbf{1}_{\{i=d\}}, \quad i \geq 0. \tag{5.17}$$

Moreover, there are exactly k fringe trees in $T_{d,k}$ that are equal to T, so $N_T(T_{d,k}) = k$, while $N_{T_{d,k}}(T) = 0$. Hence, (5.16), (1.10), and (5.17) yield

$$\begin{aligned}
0 &= \frac{\gamma_{\mathbf{p}}(T,T_{d,k})}{\pi_{\mathbf{p}}(T)\pi_{\mathbf{p}}(T_{d,k})} = \frac{k}{\pi_{\mathbf{p}}(T)} + \eta_{\mathbf{p}}(T,T_{d,k}) \\
&= \frac{k}{\pi_{\mathbf{p}}(T)} + (|T|-1)(k|T|+d-k) - \sum_{i\in\mathcal{D}(T)} \frac{n_T(i)\big(kn_T(i) + (d-k)\mathbf{1}_{\{i=0\}} + \mathbf{1}_{\{i=d\}}\big)}{p_i} \\
&= \frac{k}{\pi_{\mathbf{p}}(T)} + (|T|-1)(k|T|+d-k) - k\sum_{i\in\mathcal{D}(T)} \frac{n_T(i)^2}{p_i} - (d-k)\frac{n_T(0)}{p_0} - \frac{n_T(d)}{p_d}.
\end{aligned} \tag{5.18}$$

Subtracting k times (5.15) from (5.18) yields

$$0 = d(|T|-1) - (d-k)\frac{n_T(0)}{p_0} - \frac{n_T(d)}{p_d}. \tag{5.19}$$

This has to hold for every $k = 1,\dots,d$. If $d \geq 2$, this is a contradiction, since $n_T(0) > 0$ for every tree T.

It remains only to consider the case when $\mathcal{D}(T)$ has no element d with $d \geq 2$, i.e., when $\mathcal{D}(T) \subseteq \{0,1\}$. In this case, $|T|$ is a path, so $n_T(0) = 1$, $n_T(1) = |T|-1$ and $n_T(i) = 0$ for $i \geq 2$. We still have (5.19) with $k = d = 1$, which gives

$$|T| - 1 = \frac{n_T(1)}{p_1} = \frac{|T|-1}{p_1}. \tag{5.20}$$

Hence, $p_1 = 1$. Since $\sum_{i\geq 0} p_i = 1$, this implies $p_0 = 0$, which contradicts $\pi_{\mathbf{p}}(T) > 0$ because $n_T(0) > 0$.

These contradictions complete the proof. □

Proof of Theorem 1.7. The proof is very similar to the proof of Theorem 1.5, and we focus on the differences. We may again assume that $T_1,\dots,T_m$ are distinct, and we define again $\mu_{i\kappa} := \mu_{\mathbf{n}_\kappa}(T_i) = |\mathbf{n}_\kappa|\pi_{\mathbf{p}(\mathbf{n}_\kappa)}(T_i)$ for $1 \leq i \leq m$. Furthermore, by (1.16), $\mathbb{E} N_{T_j}(\mathcal{T}_{\mathbf{n}_\kappa}) = \mu_{\mathbf{n}_\kappa}(T_j) + O(1)$ for $1 \leq j \leq m$, and thus it does not matter whether we use $\mathbb{E} N_{T_j}(\mathcal{T}_{\mathbf{n}_\kappa})$ or $\mu_{i\kappa} = \mu_{\mathbf{n}_\kappa}(T_j)$ in (1.25).

We now assume that $q_{1\kappa},\dots,q_{m\kappa} \in \mathbb{N}_0$ are such that for some fixed constant $C > 0$,

$$q_{i\kappa} \leq C\mu_{i\kappa}^{1/2}, \qquad \text{for } \ 1 \leq i \leq m. \tag{5.21}$$

By (4.4), we then have

$$q_{j\kappa} = O\Big(\mu_{j\kappa}^{1/2}\Big) = O\Big(n_\kappa(i)^{1/2}\Big), \qquad \text{for } \ i \in \mathcal{D}(T_j), \tag{5.22}$$

and in particular, $q_{j\kappa} = O\Big(|\mathbf{n}_\kappa|^{1/2}\Big)$. Furthermore, by assumption $\mu_{j\kappa} \to \infty$, and thus (4.4) yields $n_\kappa(i) \to \infty$ for every $i \in \bigcup_{j=1}^m \mathcal{D}(T_j)$. Recall also the definition of τ_{ij} in (3.8), for $1 \leq i,j \leq m$. For $1 \leq j,k \leq m$, note that if $\tau_{jk} > 0$, then T_j is a fringe subtree of T_k; hence $n_{T_j}(i) \leq n_{T_k}(i)$ for every $i \geq 0$, and thus, by (1.19), $\mu_{k\kappa} \leq \mu_{j\kappa}$. Hence, by (5.21), if $\tau_{jk} > 0$, then $q_{k\kappa} = O\Big(\mu_{k\kappa}^{1/2}\Big) = O\Big(\mu_{j\kappa}^{1/2}\Big)$. Consequently, for every $1 \leq j \leq m$,

$$\sum_{k=1}^m \tau_{jk} q_{k\kappa} = O\Big(\mu_{j\kappa}^{1/2}\Big). \tag{5.23}$$

We now argue as in the proof of Theorem 1.5. It is easily checked that all calculations in (5.7)–(5.13) are valid in the present situation too, using (5.21)–(5.23) instead of (5.4)–(5.6).

We then use (5.13) together with (1.22) and (1.19) and obtain

$$\widetilde{r}_\kappa(j,k)|\mathbf{n}_\kappa| = \widehat{\gamma}_{\mathbf{p}(\mathbf{n}_\kappa)}(T_j,T_k)\Big(\mu_{\mathbf{n}_\kappa}(T_j)\mu_{\mathbf{n}_\kappa}(T_k)\Big)^{1/2} - \delta_{jk}\mu_{\mathbf{n}_\kappa}(T_j) = \widehat{\gamma}_{\mathbf{p}(\mathbf{n}_\kappa)}(T_j,T_k)(\mu_{j\kappa}\mu_{k\kappa})^{1/2} - \delta_{jk}\mu_{j\kappa}. \tag{5.24}$$

As $\kappa \to \infty$, we have $\widehat{\gamma}_{\mathbf{p}(\mathbf{n}_\kappa)}(T_j,T_k) \to \widehat{\gamma}_{\mathbf{p}}(T_j,T_k)$ by Condition 1.1 and the continuity of $\widehat{\gamma}_{\mathbf{p}}(T_j,T_k)$ in $\mathbf{p}$. Hence, if we now define $\sigma_{j\kappa} := \mu_{j\kappa}^{1/2}$, for $1 \leq j \leq m$, then (5.11) and (5.24) yield

$$\mathbb{E}\Big[(N_{T_1}(\mathcal{T}_{\mathbf{n}}))_{q_{1\kappa}} \cdots (N_{T_m}(\mathcal{T}_{\mathbf{n}}))_{q_{m\kappa}}\Big] = \prod_{j=1}^m \mu_{j\kappa}^{q_{j\kappa}} \exp\left(\frac{1}{2}\sum_{j,k=1}^m \frac{(\widehat{\gamma}_{\mathbf{p}}(T_j,T_k)\sigma_{j\kappa}\sigma_{k\kappa} - \delta_{jk}\mu_{j\kappa})}{\mu_{j\kappa}\mu_{k\kappa}} q_{j\kappa}q_{k\kappa} + o(1)\right), \tag{5.25}$$

uniformly in all $q_{1\kappa},\dots,q_{m\kappa}$ that satisfy (5.21). Since $\mu_{j\kappa}/\sigma_{j\kappa} = \mu_{j\kappa}^{1/2}$ for $1 \leq j \leq m$, this is precisely the condition (A.2) in Theorem A.1. Moreover, $\mu_{j\kappa} = \sigma_{j\kappa}^2$ and $\mu_{j\kappa} \to \infty$ by assumption; thus (A.1) holds too. Hence, Theorem A.1 applies and yields (1.25).

The final claim follows by Lemma 5.2 below. □

Lemma 5.2. *Suppose that $\mathbf{p} = (p_i)_{i\geq 0} \in \mathcal{P}_1(\mathbb{N}_0)$, and let $T \in \mathbb{T}$. Then $\widehat{\gamma}_{\mathbf{p}}(T,T) > 0$, except in the three exceptional cases (i), (ii) and (iii) of Theorem 1.7.*

Proof. If $\pi_{\mathbf{p}}(T) > 0$, then the result follows by Lemma 5.1. Thus suppose $\pi_{\mathbf{p}}(T) = 0$. Then, by (1.21),

$$\widehat{\eta}_{\mathbf{p}}(T,T) = -\sum_{i\in\mathcal{D}(T)} n_T(i)^2 \prod_{j\in\mathcal{D}(T)} p_j^{n_T(j)-\delta_{ij}}. \tag{5.26}$$

Since $\pi_{\mathbf{p}}(T) = 0$, there exists at least one $i_0 \in \mathcal{D}(T)$ with $p_{i_0} = 0$. Fix one such i_0. Then each product in (5.26) with $i \neq i_0$ vanishes because it contains the factor $p_{i_0}^{n_T(i_0)} = 0$. Hence,

$$\widehat{\eta}_{\mathbf{p}}(T,T) = -n_T(i_0)^2 \prod_{j\in\mathcal{D}(T)} p_j^{n_T(j)-\delta_{i_0 j}}. \tag{5.27}$$

If $n_T(i_0) \geq 2$, then (5.27) yields $\widehat{\eta}_{\mathbf{p}}(T,T) = 0$, and thus $\widehat{\gamma}_{\mathbf{p}}(T,T) = 1$ by (1.23). In the remaining case, $n_T(i_0) = 1$; thus (1.23) and (5.27) yield

$$\widehat{\gamma}_{\mathbf{p}}(T,T) = 1 + \widehat{\eta}_{\mathbf{p}}(T,T) = 1 - \prod_{j\in\mathcal{D}(T), j\neq i_0} p_j^{n_T(j)}. \tag{5.28}$$

Consequently, if $\widehat{\gamma}_{\mathbf{p}}(T,T) = 0$, then $p_j = 1$ for every $j \in \mathcal{D}(T)$ with $j \neq i_0$. Obviously, there is at most one such j, and thus $|\mathcal{D}(T)| \leq 2$.

We always have $0 \in \mathcal{D}(T)$, and thus either $|T| = 1$ (case (i)), or $\mathcal{D}(T) = \{0, d\}$ for some $d \geq 1$. In the latter case, we have either $i_0 = 0$ or $i_0 = d$.

If $i_0 = 0$, then, as shown above, $n_T(0) = 1$, so T has only one leaf and thus T is a path. Then $\mathcal{D}(T) = \{0, 1\}$ and we need $p_1 = 1$; this is case (ii).

If $i_0 = d$, then $n_T(d) = 1$. Thus, T has only one non-leaf, so T is a star where the root has degree d; furthermore, $p_0 = 1$. This is (iii). □

While the exceptional case (i) in Theorem 1.7 (and Lemma 5.2) is completely trivial, with $N_T(\mathcal{T}_{\mathbf{n}_\kappa})$ deterministic, the cases (ii) and (iii) are not. We illustrate this with a simple example, which shows that in some such cases $N_T(\mathcal{T}_{\mathbf{n}_\kappa})$ is still asymptotically normal, but not in all cases.

Example 5.3. Let T be the tree with $|T| = 2$; thus T consists of a root and a leaf, and $n_T(0) = n_T(1) = 1$. Note that T is an example of both exceptional cases (ii) (if $p_1 = 1$) and (iii) (if $p_0 = 1$).

We consider for simplicity only degree statistics $\mathbf{n}_\kappa$ such that $\mathcal{T}_{\mathbf{n}_\kappa}$ has exactly one vertex of degree ≥ 2; it then follows from (1.1) that this degree equals $n_\kappa(0)$, and thus the degree statistic $\mathbf{n}_\kappa$ has $n_\kappa(0) \geq 2$, $n_\kappa(1) \geq 0$, and $n_\kappa(i) = \delta_{i,n_\kappa(0)}$, for $i \geq 2$. For such $\mathbf{n}_\kappa$, the tree $\mathcal{T}_{\mathbf{n}_\kappa}$ consists of a vertex, v say, of degree $n_\kappa(0)$, $n_\kappa(0)$ paths from v to the leaves, and a path (which might be empty) from the root to v. Let $X_0 \geq 0$ be the number of vertices on the path from the root to v (thus $X_0 = 0$ if v is the root), and let $X_i \geq 0$ be the number of vertices of degree 1 on the i-th path from v to a leaf. Then,

$$\sum_{i=0}^{n_\kappa(0)} X_i = n_\kappa(1), \tag{5.29}$$

and there is a bijection between such vectors $(X_i)_{i=0}^{n_\kappa(0)} \in \mathbb{N}_0^{n_\kappa(0)+1}$ and possible trees $\mathcal{T}_{\mathbf{n}_\kappa}$. Vectors $(X_i)_{i=0}^{n_\kappa(0)} \in$

$\mathbb{N}_0^{n_\kappa(0)+1}$ satisfying (5.29) are called compositions of $n_\kappa(1)$ in $n_\kappa(0)+1$ parts. Hence, the random tree $\mathcal{T}_{\mathbf{n}_\kappa}$ corresponds to a uniformly random composition of $n_\kappa(1)$ in $n_\kappa(0)+1$ parts. As is well known, see [25, Example 1.3.3], [19], and e.g. [21, Example 12.2], such a random composition can be obtained by choosing any $p \in (0,1)$ and then letting $X_i \sim \mathrm{Ge}(p)$ be independent, and condition on the event that (5.29) holds.

There is one fringe subtree T in each path to a leaf for which $X_i \geq 1$; thus we obtain, with $x_+ := \max(x,0)$,

$$N_T(\mathcal{T}_{\mathbf{n}_\kappa}) \overset{\mathrm{d}}{=} \left(\sum_{i=1}^{n_\kappa(0)} \mathbf{1}_{X_i \geq 1} \;\middle|\; \sum_{i=0}^{n_\kappa(0)} X_i = n_\kappa(1)\right) = n_\kappa(1) - \left(\sum_{i=0}^{n_\kappa(0)} (X_i-1)_+ + \mathbf{1}_{\{X_0 \geq 1\}} \;\middle|\; \sum_{i=0}^{n_\kappa(0)} X_i = n_\kappa(1)\right). \tag{5.30}$$

Consider now the case when $n_\kappa(0) \to \infty$ and $n_\kappa(1) = o(n_\kappa(0))$, as $\kappa \to \infty$. This implies that Condition 1.1 holds, with $p_0 = 1$. By symmetry, $\mathbb{E}\Big[X_0 \mid \sum_{i=0}^{n_\kappa(0)} X_i = n_\kappa(1)\Big] = n_\kappa(1)/(n_\kappa(0)+1) \to 0$, and thus we may ignore the term $\mathbf{1}_{\{X_0 \geq 1\}}$ in (5.30).

For example, suppose that $n_\kappa(1)/\sqrt{n_\kappa(0)} \to \lambda \in (0,\infty)$. It is then easy to see that $n_\kappa(1) - N_T(\mathcal{T}_{\mathbf{n}_\kappa}) \overset{\mathrm{d}}{\longrightarrow} \mathrm{Po}(\lambda^2)$ with a Poisson limit distribution; see [19, Example 5]. Moreover, by the methods in [19], see also [20, Theorem 2.1], it follows easily that if $n_\kappa(0)^{1/2} \ll n_\kappa(1) \ll n_\kappa(0)$, then $N_T(\mathcal{T}_{\mathbf{n}_\kappa})$ is asymptotically normal; the variance is $\sim n_\kappa(1)^2/n_\kappa(0)$. Conversely, it is easy to see that if $n_\kappa(1) \ll n_\kappa(0)^{1/2}$, then $\mathbb{P}(N_T(\mathcal{T}_{\mathbf{n}_\kappa}) = n_\kappa(1)) \to 1$, so the distribution is asymptotically degenerate.

It is interesting to note that if $n_\kappa(0)^{3/4} \ll n_\kappa(1) \ll n_\kappa(0)$, then the asymptotic normality of $N_T(\mathcal{T}_{\mathbf{n}_\kappa})$ can easily be proved by the Gao–Wormald theorem, similarly to Theorem 1.7, using Lemmas 3.3(i) and 4.1. (With $\mu_\kappa = |n_\kappa| p_0(n_\kappa) p_1(n_\kappa)$ and $\sigma_\kappa = p_1(n_\kappa)|n_\kappa|^{1/2} = n_\kappa(1)/|n_\kappa|^{1/2}$.) However, we do not see how to use this method to prove the full range of asymptotic normality in this example.

We have here concentrated on case (iii), i.e., $p_0 = 1$. By similar arguments, one can also study the case $n_\kappa(1)/n_\kappa(0) \to \infty$, when Condition 1.1 holds with $p_1 = 1$, and we are in the exceptional case (ii). Again, normal, Poisson and degenerate limits occur for various ranges; we leave the details to the reader.

Example 5.3 treats only a simple example, and we leave the general case as an open problem.

Problem 5.4. *Find criteria for asymptotic normality of $N_T(\mathcal{T}_{\mathbf{n}_\kappa})$ in the exceptional cases (ii) and (iii) in Theorem 1.7. When is there a Poisson limit? Are there any other possible non-degenerate limit distributions?*

6 Application to labelled trees with given vertex degrees

For $n \in \mathbb{N}$, let $\mathbb{T}_n^{\mathrm{lab}}$ be the set of unordered rooted trees with n vertices labelled by $\{1,\dots,n\}$. (I.e., the labelled rooted trees of size n.) We use the notations above for such trees too, *mutatis mutandis*. In particular, for a tree $T \in \mathbb{T}_n^{\mathrm{lab}}$ and a vertex $i \in T$, $d_T(i)$ is the (out)degree of $i \in T$. We define the degree sequence of T as the sequence $\mathbf{d}_T = (d_T(i))_{i=1}^n$.

Let $\mathbb{D}_n := \{\mathbf{d}_T : T \in \mathbb{T}_n^{\mathrm{lab}}\}$ be the set of degree sequences of labelled trees of size n; if $\mathbf{d} \in \mathbb{D}_n$, we say that $\mathbf{d}$ is a *degree sequence of length n*. Note that $\mathbb{D}_n := \Big\{\mathbf{d} = (d_i)_{i=1}^n \in \mathbb{N}_0^n : \sum_{i=1}^n d_i = n-1\Big\}$. We further let $\mathbb{D} := \bigcup_{n \geq 1} \mathbb{D}_n$, the set of all degree sequences. If $\mathbf{d}$ is a degree sequence, we write $|\mathbf{d}| = n$ if

$\mathbf{d} \in \mathbb{D}_n$; we then say that n is the *length* of $\mathbf{d}$. We also define the degree statistic $\mathbf{n_d} = (n_{\mathbf{d}}(i))_{i \geq 0}$, where $n_{\mathbf{d}}(i) := |\{v \in \{1, \dots, n\} : d_v = i\}|$. Note that $|\mathbf{n_d}| = |\mathbf{d}|$.

For a degree sequence $\mathbf{d}$, let $\mathbb{T}^{\mathrm{lab}}_{\mathbf{d}}$ be the set of labelled trees $T \in \mathbb{T}^{\mathrm{lab}}_{|\mathbf{d}|}$ that have degree sequence $\mathbf{d}$. We let $\mathcal{T}^{\mathrm{lab}}_{\mathbf{d}}$ be a uniformly random element of $\mathbb{T}^{\mathrm{lab}}_{\mathbf{d}}$, i.e., a uniformly random labelled tree with degree sequence $\mathbf{d}$; we denote this by $\mathcal{T}^{\mathrm{lab}}_{\mathbf{d}} \sim \mathrm{Unif}(\mathbb{T}^{\mathrm{lab}}_{\mathbf{d}})$.

Although the random trees $\mathcal{T}_{\mathbf{n}}$ (for a degree statistics $\mathbf{n}$) and $\mathcal{T}^{\mathrm{lab}}_{\mathbf{d}}$ (for a degree sequence $\mathbf{d}$) are different types of trees, it is well known that they are closely related and for many purposes equivalent. We state one version of this as a lemma.

Lemma 6.1. *Let $\mathbf{d}$ be a degree sequence, let $\mathbf{n_d}$ be the corresponding degree statistic, and let $\mathcal{T}_{\mathbf{n_d}} \sim \mathrm{Unif}(\mathbb{T}_{\mathbf{n_d}})$. We may construct $\mathcal{T}^{\mathrm{lab}}_{\mathbf{d}} \sim \mathrm{Unif}(\mathbb{T}^{\mathrm{lab}}_{\mathbf{d}})$ as follows: randomly label the vertices of $\mathcal{T}_{\mathbf{n_d}}$ such that the tree has degree sequence $\mathbf{d}$, and then ignore the ordering.*

Proof. Let $\overline{\mathcal{T}_{\mathbf{d}}}^{\mathrm{lab}}$ be the intermediary labelled ordered random tree. The tree $\mathcal{T}_{\mathbf{n_d}}$ may be labelled in exactly $\prod_{i \geq 0} n_{\mathbf{d}}(i)!$ ways to have the specified degree sequence $\mathbf{d}$. Since this number is constant (for a given $\mathbf{d}$), $\overline{\mathcal{T}_{\mathbf{d}}}^{\mathrm{lab}}$ is uniformly distributed over all labelled ordered trees with degree sequence $\mathbf{d}$. Similarly, each labelled unordered tree with degree sequence $\mathbf{d} = (d_i)_{i=1}^n$ can be ordered in $\prod_{i=1}^n d_i!$ ways; again this number is constant, and thus the tree obtained from $\overline{\mathcal{T}_{\mathbf{d}}}^{\mathrm{lab}}$ by forgetting the ordering is uniformly distributed on $\mathbb{T}^{\mathrm{lab}}_{\mathbf{d}}$. $\square$

For a tree $T \in \mathbb{T}^{\mathrm{lab}}_n$ and a vertex $v \in T$, we define the fringe tree T_v as before. We ignore the labels on T_v; thus, T_v is regarded as an unordered unlabelled rooted tree. Let $\mathbb{T}^{\mathrm{un}}$ be the set of unordered unlabelled rooted trees. If $T \in \mathbb{T}^{\mathrm{lab}}_n$ and $T' \in \mathbb{T}^{\mathrm{un}}$, let as before $N_{T'}(T)$ be the number of fringe trees of T that are equal (i.e., isomorphic to) T'; this is again given by (1.2).

For a tree $T \in \mathbb{T}^{\mathrm{un}}$, let $\mathrm{Ord}(T)$ be the set of ordered trees $\overline{T} \in \mathbb{T}$ that reduce to T if we ignore the ordering. It follows from the construction in Lemma 6.1 that for any degree sequence $\mathbf{d}$ and tree $T \in \mathbb{T}^{\mathrm{un}}$,

$$N_T(\mathcal{T}^{\mathrm{lab}}_{\mathbf{d}}) = \sum_{\overline{T} \in \mathrm{Ord}(T)} N_{\overline{T}}(\mathcal{T}_{\mathbf{n_d}}). \tag{6.1}$$

Versions for random labelled trees $\mathbb{T}^{\mathrm{lab}}_{\mathbf{d}}$ of Theorems 1.3, 1.5, 1.6 and 1.7 now follow as a consequence of (6.1). We state only the two first of these in detail, and leave the others to the reader. We first need some notation.

Note that the definitions in Section 1 of $\pi_{\mathbf{p}}(T)$, $\eta_{\mathbf{p}}(T,T')$, $\gamma_{\mathbf{p}}(T,T')$, $\widehat{\eta}_{\mathbf{p}}(T,T')$, and $\widehat{\gamma}_{\mathbf{p}}(T,T')$ use only the degree statistics and not the orderings; these quantities are thus well defined also for unordered trees $T, T' \in \mathbb{T}^{\mathrm{un}}$; moreover, they have the same value as if we give the trees any orderings. Recall also that $\mathcal{T}_{\mathbf{p}}$ is a Galton–Watson tree with offspring distribution $\mathbf{p}$; we let $\mathcal{T}^{\mathrm{un}}_{\mathbf{p}}$ denote this Galton–Watson tree regarded as an unordered tree in $\mathbb{T}^{\mathrm{un}}$. In analogy with (1.6) we define, for $T \in \mathbb{T}^{\mathrm{un}}$,

$$\pi^{\mathrm{un}}_{\mathbf{p}}(T) := \mathbb{P}\Big(\mathcal{T}^{\mathrm{un}}_{\mathbf{p}} = T\Big) = |\mathrm{Ord}(T)| \pi_{\mathbf{p}}(T). \tag{6.2}$$

Furthermore, in analogy with (1.9)–(1.10), for $T, T' \in \mathbb{T}^{\mathrm{un}}$,

$$\gamma_{\mathbf{p}}^{\mathrm{un}}(T,T) := \pi_{\mathbf{p}}^{\mathrm{un}}(T) + \eta_{\mathbf{p}}(T,T)(\pi_{\mathbf{p}}^{\mathrm{un}}(T))^2, \tag{6.3}$$

$$\gamma_{\mathbf{p}}^{\mathrm{un}}(T,T') := N_{T'}(T)\pi_{\mathbf{p}}^{\mathrm{un}}(T) + N_T(T')\pi_{\mathbf{p}}^{\mathrm{un}}(T') + \eta_{\mathbf{p}}(T,T')\pi_{\mathbf{p}}^{\mathrm{un}}(T)\pi_{\mathbf{p}}^{\mathrm{un}}(T'), \qquad T \neq T'. \tag{6.4}$$

Theorem 6.2. *Let $\mathbf{d}_\kappa$, $\kappa \geq 1$, be some degree sequences such that the corresponding degree statistics $\mathbf{n}_{\mathbf{d}_\kappa}$ satisfy Condition 1.1, and let $\mathcal{T}_{\mathbf{d}_\kappa}^{\mathrm{lab}} \sim \mathrm{Unif}(\mathbb{T}_{\mathbf{d}_\kappa}^{\mathrm{lab}})$. For every fixed $T \in \mathbb{T}^{\mathrm{un}}$, as $\kappa \to \infty$:*

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

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

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

Proof. This follows by Theorem 1.3 together with (6.1) since, for any $T \in \mathbb{T}^{\mathrm{un}}$, using (6.2),

$$\sum_{\overline{T} \in \mathrm{Ord}(T)} \pi_{\mathbf{p}}(\overline{T}) = |\mathrm{Ord}(T)| \pi_{\mathbf{p}}(T) = \pi^{\mathrm{un}}(T). \tag{6.5}$$

□

Theorem 6.3. *Let $\mathbf{d}_\kappa$, $\kappa \geq 1$, be some degree sequences such that the corresponding degree statistics $\mathbf{n}_{\mathbf{d}_\kappa}$ satisfy Condition 1.1, and let $\mathcal{T}_{\mathbf{d}_\kappa}^{\mathrm{lab}} \sim \mathrm{Unif}(\mathbb{T}_{\mathbf{d}_\kappa}^{\mathrm{lab}})$. For a fixed $m \geq 1$, let $T_1, \ldots, T_m \in \mathbb{T}^{\mathrm{un}}$ be a fixed sequence of rooted unordered unlabelled trees. Then, as $\kappa \to \infty$,*

$$\mathbb{E}\, N_{T_i}(\mathcal{T}_{\mathbf{d}_\kappa}^{\mathrm{lab}}) = \pi_{\mathbf{p}}^{\mathrm{un}}(T_i)|\mathbf{d}_\kappa| + o(|\mathbf{d}_\kappa|), \tag{6.6}$$

$$\mathrm{Var}\, N_{T_i}(\mathcal{T}_{\mathbf{d}_\kappa}^{\mathrm{lab}}) = \gamma_{\mathbf{p}}^{\mathrm{un}}(T_i, T_i)|\mathbf{d}_\kappa| + o(|\mathbf{d}_\kappa|), \tag{6.7}$$

$$\mathrm{Cov}\Big(N_{T_i}(\mathcal{T}_{\mathbf{d}_\kappa}^{\mathrm{lab}}), N_{T_j}(\mathcal{T}_{\mathbf{d}_\kappa}^{\mathrm{lab}})\Big) = \gamma_{\mathbf{p}}^{\mathrm{un}}(T_i, T_j)|\mathbf{d}_\kappa| + o(|\mathbf{d}_\kappa|), \tag{6.8}$$

for $1 \leq i, j \leq m$, and

$$\left(\frac{N_{T_1}(\mathcal{T}_{\mathbf{d}_\kappa}^{\mathrm{lab}}) - \mathbb{E}[N_{T_1}(\mathcal{T}_{\mathbf{d}_\kappa}^{\mathrm{lab}})]}{\sqrt{|\mathbf{d}_\kappa|}}, \ldots, \frac{N_{T_m}(\mathcal{T}_{\mathbf{d}_\kappa}^{\mathrm{lab}}) - \mathbb{E}[N_{T_m}(\mathcal{T}_{\mathbf{d}_\kappa}^{\mathrm{lab}})]}{\sqrt{|\mathbf{d}_\kappa|}}\right) \xrightarrow{\mathrm{d}} \mathrm{N}(0, \Gamma_{\mathbf{p}}^{\mathrm{un}}), \tag{6.9}$$

where the covariance matrix $\Gamma_{\mathbf{p}}^{\mathrm{un}} := (\gamma_{\mathbf{p}}^{\mathrm{un}}(T_i, T_j))_{i,j=1}^m$. Furthermore, in (6.9), we can replace $\mathbb{E}[N_{T_i}(\mathcal{T}_{\mathbf{d}_\kappa}^{\mathrm{lab}})]$ by $|\mathbf{d}_\kappa| \pi_{\mathbf{p}(\mathbf{n}_{\mathbf{d}_\kappa})}^{\mathrm{un}}(T_i)$.

Proof. This follows by Theorem 1.5 together with (6.1), using (6.5) and the following calculations. First, for any $T \in \mathbb{T}^{\mathrm{un}}$, by (1.9)–(1.10), (6.2), and the fact that if $\overline{T}, \overline{T}' \in \mathbb{T}$ with $|\overline{T}| = |\overline{T}'|$ but $\overline{T} \neq \overline{T}'$, then

$N_{\overline{T}}(\overline{T}') = 0$,

$$\sum_{\overline{T}\in\mathrm{Ord}(T),\overline{T}'\in\mathrm{Ord}(T)} \gamma_{\mathbf{p}}(\overline{T},\overline{T}') = \sum_{\overline{T}\in\mathrm{Ord}(T),\overline{T}'\in\mathrm{Ord}(T)} \Big(\pi_{\mathbf{p}}(T)\mathbf{1}_{\{\overline{T}=\overline{T}'\}} + \eta_{\mathbf{p}}(\overline{T},\overline{T}')\pi_{\mathbf{p}}(T)^2\Big)$$
$$= |\mathrm{Ord}(T)|\pi_{\mathbf{p}}(T) + |\mathrm{Ord}(T)|^2\pi_{\mathbf{p}}(T)^2\eta_{\mathbf{p}}(T,T) = \gamma^{\mathrm{un}}_{\mathbf{p}}(T,T). \tag{6.10}$$

Secondly, for $T,T'\in\mathbb{T}^{\mathrm{un}}$ with $T\neq T'$, we have, cf. (6.1),

$$\sum_{\overline{T}\in\mathrm{Ord}(T),\overline{T}'\in\mathrm{Ord}(T')} N_{\overline{T}'}(\overline{T}) = \sum_{\overline{T}\in\mathrm{Ord}(T)} N_{T'}(\overline{T}) = |\mathrm{Ord}(T)|N_{T'}(T), \tag{6.11}$$

and thus, similarly, by (1.10), (6.2) and (6.4),

$$\sum_{\overline{T}\in\mathrm{Ord}(T),\overline{T}'\in\mathrm{Ord}(T')} \gamma_{\mathbf{p}}(\overline{T},\overline{T}') = \sum_{\overline{T}\in\mathrm{Ord}(T),\overline{T}'\in\mathrm{Ord}(T')} \Big(N_{\overline{T}'}(\overline{T})\pi_{\mathbf{p}}(T) + N_{\overline{T}}(\overline{T}')\pi_{\mathbf{p}}(T') + \eta_{\mathbf{p}}(T,T')\pi_{\mathbf{p}}(T)\pi_{\mathbf{p}}(T')\Big)$$
$$= \gamma^{\mathrm{un}}_{\mathbf{p}}(T,T'). \tag{6.12}$$

□

Problem 6.4. *Suppose that $T\in\mathbb{T}^{\mathrm{un}}$ and $\mathbf{p}\in\mathcal{P}_1(\mathbb{N}_0)$ with $|T|>1$ and $\pi^{\mathrm{un}}_{\mathbf{p}}(T)>0$. Is $\gamma^{\mathrm{un}}_{\mathbf{p}}(T,T)>0$?*

Note that an affirmative answer to Problem 1.9 would imply a positive answer to this too.

7 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

$$w(T) := \prod_{v\in T} w_{d_T(v)} = \prod_{i\geq 0} w_i^{n_T(i)}. \tag{7.1}$$

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.,

$$\mathbb{P}(\mathcal{T}_{\mathbf{w},n} = T) = \frac{w(T)}{Z_n(\mathbf{w})}, \quad \text{for } T\in\mathbb{T}_n. \tag{7.2}$$

The random tree $\mathcal{T}_{\mathbf{w},n}$ is called simply generated tree of size n and weight sequence $\mathbf{w}$; see e.g. [8] and [21]. 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 $0 \le s < \rho_{\mathbf{w}}$, we let

$$\Psi_{\mathbf{w}}(s) := \frac{s\Phi'_{\mathbf{w}}(s)}{\Phi_{\mathbf{w}}(s)} = \frac{\sum_{i\ge 0} i w_i s^i}{\sum_{i\ge 0} w_i s^i}. \tag{7.3}$$

Furthermore, if $\Phi_{\mathbf{w}}(\rho_{\mathbf{w}}) < \infty$, we define also $\Psi_{\mathbf{w}}(\rho_{\mathbf{w}})$ by (7.3); 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 [21, Lemma 3.1 (i)]. Let $\nu_{\mathbf{w}} := \Psi_{\mathbf{w}}(\rho_{\mathbf{w}}) \in [0,\infty]$, and define

$$\tau_{\mathbf{w}} = \begin{cases} \rho_{\mathbf{w}} & \text{if } \nu_{\mathbf{w}} < 1, \\ \Psi_{\mathbf{w}}^{-1}(1) & \text{if } \nu_{\mathbf{w}} \ge 1. \end{cases} \tag{7.4}$$

It follows from [21, Lemma 3.1] that

$$\rho_{\mathbf{w}} > 0 \iff \nu_{\mathbf{w}} > 0 \iff \tau_{\mathbf{w}} > 0. \tag{7.5}$$

The following result from [21] shows that simply generated trees satisfy Condition 1.1 in probability.

Theorem 7.1 ([21, 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* $i \ge 2$*. Define*

$$\theta_i(\mathbf{w}) = \frac{w_i \tau_{\mathbf{w}}^i}{\Phi_{\mathbf{w}}(\tau_{\mathbf{w}})}, \quad \textit{for} \ \ i \ge 0. \tag{7.6}$$

Then, $\theta(\mathbf{w}) = (\theta_i(\mathbf{w}))_{i\ge 0}$ *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* $i \ge 0$*,* $p_i(\mathbf{n}_{\mathcal{T}_{\mathbf{w},n}}) \xrightarrow{\mathrm{p}} \theta_i(\mathbf{w})$*, 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 $i \ge 1$; otherwise, $\tau_{\mathbf{w}} > 0$ and (7.6) shows that $\theta_i(\mathbf{w}) > 0 \iff w_i > 0$ for $i \ge 0$.

Using Theorem 7.1, we will show that Theorem 1.5 implies the following version for conditioned Galton–Watson trees. The asymptotic normality (7.9) was proved in case (i) by different methods in [22, Corollary 1.8]; (ii) and (iii) are new.

Theorem 7.2 (partly [22])**.** *Let* $\mathbf{w}$ *be a sequence of non-negative real weights with* $w_0 > 0$ *and* $w_i > 0$ *for at least one* $i \ge 2$*. 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* $m \ge 1$*, 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* $Z_n(\mathbf{w}) > 0$*),*

$$\left(\frac{N_{T_1}(\mathcal{T}_{\mathbf{w},n}) - \mathbb{E}[N_{T_1}(\mathcal{T}_{\mathbf{w},n}) \mid \mathbf{n}_{\mathcal{T}_{\mathbf{w},n}}]}{\sqrt{n}}, \dots, \frac{N_{T_m}(\mathcal{T}_{\mathbf{w},n}) - \mathbb{E}[N_{T_m}(\mathcal{T}_{\mathbf{w},n}) \mid \mathbf{n}_{\mathcal{T}_{\mathbf{w},n}}]}{\sqrt{n}} \right) \xrightarrow{\mathrm{d}} \mathrm{N}(0, \Gamma_{\theta(\mathbf{w})}), \tag{7.7}$$

where the covariance matrix $\Gamma_{\theta(\mathbf{w})}$ is defined by (1.9)–(1.10)*, and for* $1 \le j \le m$,

$$\mathbb{E}[N_{T_j}(\mathcal{T}_{\mathbf{w},n}) \mid \mathbf{n}_{\mathcal{T}_{\mathbf{w},n}}] = \frac{n}{(n)_{|T_j|}} \prod_{i \ge 0} (n_{\mathcal{T}_{\mathbf{w},n}}(i))_{n_{T_j}(i)}. \tag{7.8}$$

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

(i) $\nu_{\mathbf{w}} \ge 1$ *and* $\sigma^2_{\mathbf{w}} \in (0, \infty)$.

(ii) $\nu_{\mathbf{w}} \ge 1$, $\sigma^2_{\mathbf{w}} = \infty$ *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* $k \to \infty$ *[9, Theorem XVII.5.2].)*

(iii) $0 < \nu_{\mathbf{w}} < 1$ *and* $\theta_i(\mathbf{w}) = ci^{-\beta} + o(i^{-\beta})$*, as* $i \to \infty$*, with fixed* $c > 0$ *and* $\beta > 2$.

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

$$\left(\frac{N_{T_1}(\mathcal{T}_{\mathbf{w},n}) - n\pi_{\theta(\mathbf{w})}(T_1)}{\sqrt{n}}, \dots, \frac{N_{T_m}(\mathcal{T}_{\mathbf{w},n}) - n\pi_{\theta(\mathbf{w})}(T_m)}{\sqrt{n}} \right) \xrightarrow{\mathrm{d}} \mathrm{N}(0, \widetilde{\Gamma}_{\theta(\mathbf{w})}), \tag{7.9}$$

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* $T \neq T'$,

$$\widetilde{\gamma}_{\theta(\mathbf{w})}(T, T) = \pi_{\theta(\mathbf{w})}(T) - \left(2|T| - 1 + \varsigma_{\mathbf{w}}^{-2}\right)(\pi_{\theta(\mathbf{w})}(T))^2, \tag{7.10}$$

$$\widetilde{\gamma}_{\theta(\mathbf{w})}(T, T') = N_{T'}(T)\pi_{\theta(\mathbf{w})}(T) + N_T(T')\pi_{\theta(\mathbf{w})}(T') - \left(|T| + |T'| - 1 + \varsigma_{\mathbf{w}}^{-2}\right)\pi_{\theta(\mathbf{w})}(T)\pi_{\theta(\mathbf{w})}(T'), \tag{7.11}$$

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

Remark 7.3. Recall that for any weight sequence $\mathbf{w}$ and any constants $a, b > 0$, the weight sequence $\widehat{\mathbf{w}} = (\widehat{w}_i)_{i \ge 0}$ with $\widehat{w}_i := ab^i w_i$ is equivalent to $\mathbf{w}$, i.e., it satisfies that $\mathcal{T}_{\mathbf{w},n} \overset{\mathrm{d}}{=} \mathcal{T}_{\widehat{\mathbf{w}},n}$, for all n for which either (and thus both) of the random trees are defined; this is a consequence of (7.2). In the setting of Theorem 7.1, if $\rho_{\mathbf{w}} > 0$, then the weight sequence $\mathbf{w}$ is equivalent to the weight sequence $\theta(\mathbf{w}) = (\theta_i(\mathbf{w}), i \ge 0)$, which is a probability distribution with mean $\mu_{\mathbf{w}} = \min(1, \nu_{\mathbf{w}})$; see further [21, 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 7.2, 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 $\nu_{\mathbf{w}} \ge 1$, and subcritical if $0 < \nu_{\mathbf{w}} < 1$.

We postpone the proof of Theorem 7.2. A central idea is to obtain the unconditional limit (7.9) by combining the conditional limit (7.7) with a limit result for the conditional expectations in (7.8). For this, we will use the following theorem on asymptotic normality of the degree statistics, which is proved in [22] and [31]. To be more precise, case (i) is shown in [22, Example 2.2], while cases (ii) and (iii) are shown in [31, Theorems 6.2 and 6.7] (although the asymptotic (co)variances are not explicitly given in [31, Theorem 6.2]). Moreover, the approach used in the present paper allows us to give a different (and simpler) proof of Theorem 7.6 in cases (i) and (ii), using the multidimensional version of the Gao–Wormald theorem (Theorem A.1); we give this proof in Appendix B.

Remark 7.4. In case (i), i.e. $\nu_{\mathbf{w}} \ge 1$ and $\sigma_{\mathbf{w}}^2 \in (0,\infty)$, the univariate version of Theorem 7.6 was first proved by Kolchin [25, Theorem 2.3.1]; the multivariate result (7.9) was proved in general in [22] as said above, and earlier under extra moment assumptions on $\theta(\mathbf{w})$ by Janson [20] (assuming a third moment), Minami [26] and Drmota [8, Section 3.2.1] (both assuming an exponential moment), using different proofs.

Remark 7.5. If $\nu_{\mathbf{w}} \ge 1$ and $\sigma_{\mathbf{w}}^2 \in (0,\infty)$ as in Theorem 7.2(i), then the offspring distribution $\theta(\mathbf{w})$ is critical (i.e., has mean 1) with finite variance. This is the framework assumed in [22], and as said above, in this case, (7.9) is proved in [22, Corollary 1.8]. Our proof uses Theorem 7.6, which in this case also is a result from [22]. However, note that our proof is quite different; we use Theorem 7.6 to prove (7.9), while [22] essentially does the opposite.

Theorem 7.6 ([22] and [31])**.** *Let* $\mathbf{w}$ *and* $\mathcal{T}_{\mathbf{w},n}$ *be as in Theorem 7.2, and assume that one of the conditions (i)–(iii) there holds. Then, for any fixed* $k \in \mathbb{N}_0$*, as* $n \to \infty$ *(along integers* n *such that* $Z_n(\mathbf{w}) > 0$*),*

$$\left(\frac{n_{\mathcal{T}_{\mathbf{w},n}}(0) - n\theta_0(\mathbf{w})}{\sqrt{n}}, \dots, \frac{n_{\mathcal{T}_{\mathbf{w},n}}(k) - n\theta_k(\mathbf{w})}{\sqrt{n}}\right) \xrightarrow{\mathrm{d}} \mathrm{N}(0, \Gamma_k^*), \tag{7.12}$$

where the covariance matrix $\Gamma_k^* := (\gamma^*(i,j))_{i,j=0}^k$ *is given by*

$$\gamma^*(i,i) = \theta_i(\mathbf{w})(1-\theta_i(\mathbf{w})) - (i-1)^2\theta_i(\mathbf{w})^2/\varsigma_{\mathbf{w}}^2, \tag{7.13}$$

$$\gamma^*(i,j) = -\theta_i(\mathbf{w})\theta_j(\mathbf{w}) - (i-1)(j-1)\theta_i(\mathbf{w})\theta_j(\mathbf{w})/\varsigma_{\mathbf{w}}^2, \qquad i \ne j, \tag{7.14}$$

where $\varsigma_{\mathbf{w}}^2$ *is as in Theorem 7.2. (In particular,* $\varsigma_{\mathbf{w}}^2 = \infty$ *in cases (ii) and (iii); we then interpret the final terms in (7.13) and (7.14) as* 0*.)*

Proof of Theorem 7.2. For any fixed degree statistic $\mathbf{n}$ with $\mathbb{P}(\mathbf{n}_{\mathcal{T}_{\mathbf{w},n}} = \mathbf{n}) > 0$, (7.2) 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 [24, Theorem 4.30], we can and will assume that the convergence in Theorem 7.1 holds a.s.; in other words, Condition 1.1 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 $\mathcal{T}_{\mathbf{w},n}$, $n \ge 1$, have the (conditional) distributions $\mathrm{Unif}(\mathbb{T}_{\mathbf{n}_{\mathcal{T}_{\mathbf{w},n}}})$. It follows that we may apply Theorem 1.5 conditioned on the sequence of degree statistics $(\mathbf{n}_{\mathcal{T}_{\mathbf{w},n}})_{n=1}^\infty$; this shows that (7.7) holds conditioned on $(\mathbf{n}_{\mathcal{T}_{\mathbf{w},n}})_{n=1}^\infty$. Then, (7.7) also holds unconditionally by the dominated convergence theorem. Furthermore, (7.8) follows from Lemma 3.1.

In the rest of the proof we assume that either (i), (ii) or (iii) holds, and thus Theorem 7.6 applies. We fix k so large that $k \ge i$ for every $i \in \bigcup_{j=1}^m \mathcal{D}(T_j)$. (Recall (1.7).) For convenience, we use again the Skorohod coupling theorem, and may thus assume that the limits $\mathbf{p}(\mathbf{n}_{\mathcal{T}_{\mathbf{w},n}}) \to \theta(\mathbf{w})$ in Theorem 7.1 and (7.12) in Theorem 7.6 hold almost surely.

Let $1 \le j \le m$. First, suppose that $\pi_{\theta(\mathbf{w})}(T_j) > 0$, i.e., $\theta_i(\mathbf{w}) > 0$ for $i \in \mathcal{D}(T_j)$. Then, a.s., by (7.8) and

Lemma 4.1 and the assumed a.s. convergence in (7.12),

$$\begin{aligned}
\mathbb{E}[N_{T_j}(\mathcal{T}_{\mathbf{w},n}) \mid \mathbf{n}_{\mathcal{T}_{\mathbf{w},n}}] &= n \prod_{i\in\mathcal{D}(T_j)} \left(\frac{n_{\mathcal{T}_{\mathbf{w},n}}(i)}{n}\right)^{n_{T_j}(i)} + O(1) \\
&= n \exp\left(\sum_{i\in\mathcal{D}(T_j)} n_{T_j}(i) \ln\left(\frac{n_{\mathcal{T}_{\mathbf{w},n}}(i) - n\theta_i(\mathbf{w})}{n} + \theta_i(\mathbf{w})\right)\right) + O(1) \\
&= \pi_{\theta(\mathbf{w})}(T_j)n + \pi_{\theta(\mathbf{w})}(T_j) \sum_{i\geq 0} \frac{n_{T_j}(i)}{\theta_i(\mathbf{w})} \left(n_{\mathcal{T}_{\mathbf{w},n}}(i) - n\theta_i(\mathbf{w})\right) + O(1). \qquad (7.15)
\end{aligned}$$

On the other hand, if $\pi_{\theta(\mathbf{w})}(T_j) = 0$, then $\theta_{i'}(\mathbf{w}) = 0$ for some $i' \geq 0$ such that $n_{T_j}(i') > 0$. Since we assume $\nu_{\mathbf{w}} > 0$, this implies by (7.5) and (7.6) that $w_{i'} = 0$; hence, $\mathcal{T}_{\mathbf{w},n}$ a.s. contains no vertex of degree i', and thus no fringe subtree T_j, so $N_{T_j}(\mathcal{T}_{\mathbf{w},n}) = 0$. Consequently, (7.15) holds trivially in this case too. Thus (7.15) holds in both cases.

Since the sum in (7.15) really only contains a finite number of terms, it and (7.12) imply that, as $n \to \infty$, a.s., for some random vector $\mathbf{W}$,

$$\left(\frac{\mathbb{E}[N_{T_1}(\mathcal{T}_{\mathbf{w},n}) \mid \mathbf{n}_{\mathcal{T}_{\mathbf{w},n}}] - n\pi_{\theta(\mathbf{w})}(T_1)}{\sqrt{n}}, \dots, \frac{\mathbb{E}[N_{T_m}(\mathcal{T}_{\mathbf{w},n}) \mid \mathbf{n}_{\mathcal{T}_{\mathbf{w},n}}] - n\pi_{\theta(\mathbf{w})}(T_m)}{\sqrt{n}}\right) \to \mathbf{W} \sim \mathrm{N}(0, \Gamma'_{\theta(\mathbf{w})}), \qquad (7.16)$$

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

$$\begin{aligned}
\gamma'_{\theta(\mathbf{w})}(T, T') &= \pi_{\theta(\mathbf{w})}(T)\pi_{\theta(\mathbf{w})}(T') \sum_{r,r'\geq 0} \frac{n_T(r)n_{T'}(r')}{\theta_r(\mathbf{w})\theta_{r'}(\mathbf{w})} \gamma^*(r, r') \\
&= \pi_{\theta(\mathbf{w})}(T)\pi_{\theta(\mathbf{w})}(T') \left(-|T||T'| - \varsigma_{\mathbf{w}}^{-2} + \sum_{r\geq 0} \frac{n_T(r)n_{T'}(r)}{\theta_r(\mathbf{w})}\right). \qquad (7.17)
\end{aligned}$$

To obtain the second equality, we have used (7.13)–(7.14) and (1.1), i.e., $|T| = \sum_{r\geq 0} n_T(r) = 1 + \sum_{r\geq 0} r n_T(r)$, noting that it suffices to consider the case $\pi_{\theta(\mathbf{w})}(T)\pi_{\theta(\mathbf{w})}(T') > 0$, which implies that $\theta_r(\mathbf{w})\theta_{r'}(\mathbf{w}) > 0$ when $n_T(r)n_{T'}(r') > 0$.

Recall that (7.7) holds conditioned on the sequence $(\mathbf{n}_{\mathcal{T}_{\mathbf{w},n}})_{n=1}^\infty$. Therefore, the limits (7.7) and (7.16) hold jointly, with independent limits. It follows that (7.9) holds with $\widetilde{\Gamma}_{\theta(\mathbf{w})} = \Gamma_{\theta(\mathbf{w})} + \Gamma'_{\theta(\mathbf{w})}$, and a simple calculation yields (7.10)–(7.11). ☐

Theorem 7.2 gives a partial solution to [21, Problem 21.4], but the general case remains open.

Problem 7.7. *Does (7.9) in Theorem 7.2 hold for any weight sequence* $\mathbf{w}$*, with some covariance matrix* $\widetilde{\Gamma}_{\theta(\mathbf{w})} = (\widetilde{\gamma}_{\theta(\mathbf{w})}(T_i, T_j))_{i,j=1}^m$*? If so, is* $\widetilde{\Gamma}_{\theta(\mathbf{w})}$ *given by (7.10)–(7.11), for a suitable* $\varsigma_{\mathbf{w}}^2$*?*

The argument used to answer the previous question in the cases (i), (ii) or (iii) (second part of Theorem 7.2) works as soon as one has a general version of Theorem 7.6.

Problem 7.8. *Does (7.12) in Theorem 7.6 hold for any weight sequence* $\mathbf{w}$*?*

This was conjectured in [31] (at least for $\rho_{\mathbf{w}} > 0$), but remains open as far as we know. See also Remark B.1.

8 Application to additive functionals

Let $f : \mathbb{T} \to \mathbb{R}$ be a functional of rooted trees (in this context often called *toll function*) and for $T \in \mathbb{T}$, consider the functional F (often called an *additive functional*) that is defined as the sum over all fringe subtrees

$$F(T) = F(T, f) := \sum_{v \in T} f(T_v). \tag{8.1}$$

In particular, by choosing $f(T) = \mathbf{1}_{\{T=T'\}}$ for some $T' \in \mathbb{T}$, we obtain $F(T) = N_{T'}(T)$. Moreover, for any f,

$$F(T) = \sum_{T' \in \mathbb{T}} f(T') N_{T'}(T), \tag{8.2}$$

i.e., $F(T)$ can be written as a linear combination of subtree counts $N_{T'}(T)$.

For a probability distribution $\mathbf{p} = (p_i)_{i\geq 0} \in \mathcal{P}_1(\mathbb{N}_0)$ and a functional $f : \mathbb{T} \to \mathbb{R}$ such that

$$\sum_{T \in \mathbb{T}} |f(T)| \pi_{\mathbf{p}}(T) < \infty, \tag{8.3}$$

we let

$$\mathbb{E}_{\pi_{\mathbf{p}}}[f(\mathcal{T})] := \mathbb{E}[f(\mathcal{T}_{\mathbf{p}})] = \sum_{T \in \mathbb{T}} f(T) \pi_{\mathbf{p}}(T). \tag{8.4}$$

We say that the functional $f : \mathbb{T} \to \mathbb{R}$ has *finite support* if $f(T) \neq 0$ only for finitely many trees; equivalently, there exists a constant $K > 0$ such that $f(T) = 0$ unless $|T| \leq K$. It then follows from (8.2) that the additive functionals F associated to f with finite support are exactly the (finite) linear combinations of subtree counts. Theorem 1.5 implies the following corollary. Note that a functional f with finite support always satisfies (8.3) for any distribution $\mathbf{p} \in \mathcal{P}_1(\mathbb{N}_0)$; indeed, in this case, the left-hand side of (8.3) is a sum with finitely many non-zero summands.

Theorem 8.1. *Let $\mathbf{n}_\kappa$, $\kappa \geq 1$, be some degree statistics that satisfy Condition 1.1 and let $\mathcal{T}_{\mathbf{n}_\kappa} \sim \mathrm{Unif}(\mathbb{T}_{\mathbf{n}_\kappa})$. Suppose that $f : \mathbb{T} \to \mathbb{R}$ is a functional of rooted trees with finite support, and let F be the corresponding additive functional. Then, as $\kappa \to \infty$,*

$$\mathbb{E}\, F(\mathcal{T}_{\mathbf{n}_\kappa}) = |\mathbf{n}_\kappa|\, \mathbb{E}_{\pi_{\mathbf{p}}}[f(\mathcal{T})] + o(|\mathbf{n}_\kappa|), \tag{8.5}$$

$$\mathrm{Var}\, F(\mathcal{T}_{\mathbf{n}_\kappa}) = |\mathbf{n}_\kappa| \gamma_{\mathbf{p}}(f) + o(|\mathbf{n}_\kappa|), \tag{8.6}$$

where

$$\gamma_{\mathbf{p}}(f) := 2\,\mathbb{E}_{\pi_{\mathbf{p}}}[f(\mathcal{T})F(\mathcal{T})] - \mathbb{E}_{\pi_{\mathbf{p}}}[f(\mathcal{T})^2] + \Big(\mathbb{E}_{\pi_{\mathbf{p}}}[f(\mathcal{T})(|\mathcal{T}|-1)]\Big)^2 - \sum_{i\geq 0}\frac{1}{p_i}\Big(\mathbb{E}_{\pi_{\mathbf{p}}}[f(\mathcal{T})n_{\mathcal{T}}(i)]\Big)^2 \mathbf{1}_{\{p_i>0\}} \qquad (8.7)$$

is finite, with $0 \leq \gamma_{\mathbf{p}}(f) < \infty$. *Furthermore,*

$$\frac{F(\mathcal{T}_{\mathbf{n}_\kappa}) - \mathbb{E}[F(\mathcal{T}_{\mathbf{n}_\kappa})]}{\sqrt{|\mathbf{n}_\kappa|}} \xrightarrow{\mathrm{d}} \mathrm{N}(0, \gamma_{\mathbf{p}}(f)), \quad \textit{as} \ \ \kappa \to \infty. \qquad (8.8)$$

Remark 8.2. We have excluded terms with $p_i = 0$ in the sum in (8.7); these terms make no difference, since if $p_i = 0$, then for every T either $n_T(i) = 0$ or $\pi_{\mathbf{p}}(T) = 0$, and thus $\mathbb{E}_{\pi_{\mathbf{p}}}[f(\mathcal{T})n_{\mathcal{T}}(i)] = 0$. Furthermore, this sum has only finitely many non-zero terms, since we only need to consider $i \in \bigcup_{T: f(T)\neq 0}\mathcal{D}(T)$, which is a finite union of finite sets.

Proof. Since f has finite support, there exists a constant $K > 0$ such that $f(T) = 0$ unless $|T| \leq K$, for $T \in \mathbb{T}$. Let $\mathbb{T}^{(K)} := \{T \in \mathbb{T} : |T| \leq K\}$, and note that $\mathbb{T}^{(K)}$ is a finite set of trees. It follows from (8.1) and (8.2) that the corresponding additive functional F is given by

$$F(T) = \sum_{T' \in \mathbb{T}^{(K)}} f(T')N_{T'}(T), \quad T \in \mathbb{T}; \qquad (8.9)$$

note that (8.9) has finitely many summands. We label the elements of $\mathbb{T}^{(K)}$ as $T_1, \dots, T_m$, for some $m \in \mathbb{N}$, and apply Theorem 1.5.

Since f has finite support, it satisfies (8.3), and thus $\mathbb{E}_{\pi_{\mathbf{p}}}[f(\mathcal{T})]$ is defined (and finite). Then, as $\kappa \to \infty$, (8.5) follows from (8.9), (1.11) in Theorem 1.5, and (8.4), which yield

$$\mathbb{E}\,F(\mathcal{T}_{\mathbf{n}_\kappa}) = \sum_{T \in \mathbb{T}^{(K)}} f(T)\,\mathbb{E}\,N_T(\mathcal{T}_{\mathbf{n}_\kappa}) = |\mathbf{n}_\kappa| \sum_{T \in \mathbb{T}^{(K)}} f(T)\pi_{\mathbf{p}}(T) + o(|\mathbf{n}_\kappa|) = |\mathbf{n}_\kappa|\,\mathbb{E}_{\pi_{\mathbf{p}}}[f(\mathcal{T})] + o(|\mathbf{n}_\kappa|). \qquad (8.10)$$

Similarly, (8.9) and (1.12)–(1.13) in Theorem 1.5 imply that, as $\kappa \to \infty$,

$$\begin{aligned}\operatorname{Var} F(\mathcal{T}_{\mathbf{n}_\kappa}) &= \sum_{T \in \mathbb{T}^{(K)}} \sum_{T' \in \mathbb{T}^{(K)}} f(T)f(T')\operatorname{Cov}\Big(N_T(\mathcal{T}_{\mathbf{n}_\kappa}), N_{T'}(\mathcal{T}_{\mathbf{n}_\kappa})\Big) \\ &= |\mathbf{n}_\kappa| \sum_{T \in \mathbb{T}^{(K)}} \sum_{T' \in \mathbb{T}^{(K)}} f(T)f(T')\gamma_{\mathbf{p}}(T, T') + o(|\mathbf{n}_\kappa|), \end{aligned} \qquad (8.11)$$

where $\gamma_{\mathbf{p}}(T, T')$ is defined in (1.9)–(1.10). In other words, (8.6) holds with

$$\gamma_{\mathbf{p}}(f) := \sum_{T \in \mathbb{T}^{(K)}} \sum_{T' \in \mathbb{T}^{(K)}} f(T)f(T')\gamma_{\mathbf{p}}(T, T'). \qquad (8.12)$$

Moreover, (8.9) and (8.10) imply that

$$F(\mathcal{T}_{\mathbf{n}_\kappa}) - \mathbb{E}\,F(\mathcal{T}_{\mathbf{n}_\kappa}) = \sum_{T \in \mathbb{T}^{(K)}} f(T)\Big(N_T(\mathcal{T}_{\mathbf{n}_\kappa}) - \mathbb{E}\,N_T(\mathcal{T}_{\mathbf{n}_\kappa})\Big), \qquad (8.13)$$

which together with (1.14) implies that (8.8) holds with the same $\gamma_{\mathbf{p}}(f)$ given by (8.12).

It remains to evaluate this $\gamma_{\mathbf{p}}(f)$ and show that (8.12) agrees with (8.7). First, (8.4) yields

$$\sum_{T\in\mathbb{T}^{(K)}} f(T)^2\pi_{\mathbf{p}}(T) = \mathbb{E}_{\pi_{\mathbf{p}}}[f(\mathcal{T})^2]. \tag{8.14}$$

Next, observe that (8.4) and (8.9) yield

$$\begin{aligned}\sum_{T\in\mathbb{T}^{(K)}}\sum_{T'\in\mathbb{T}^{(K)}} f(T)f(T')N_{T'}(T)\pi_{\mathbf{p}}(T) &= \sum_{T'\in\mathbb{T}^{(K)}} f(T')\sum_{T\in\mathbb{T}^{(K)}} f(T)N_{T'}(T)\pi_{\mathbf{p}}(T)\\ &= \sum_{T'\in\mathbb{T}^{(K)}} f(T')\,\mathbb{E}_{\pi_{\mathbf{p}}}[f(\mathcal{T})N_{T'}(\mathcal{T})]\\ &= \mathbb{E}_{\pi_{\mathbf{p}}}[f(\mathcal{T})F(\mathcal{T})]. \end{aligned}\tag{8.15}$$

Hence, recalling (8.14), since $N_T(T)=1$ for every tree T,

$$\begin{aligned}\sum_{T\neq T'} f(T)f(T')N_{T'}(T)\pi_{\mathbf{p}}(T) &= \mathbb{E}_{\pi_{\mathbf{p}}}[f(\mathcal{T})F(\mathcal{T})] - \sum_{T\in\mathbb{T}^{(K)}} f(T)^2\pi_{\mathbf{p}}(T)\\ &= \mathbb{E}_{\pi_{\mathbf{p}}}[f(\mathcal{T})F(\mathcal{T})] - \mathbb{E}_{\pi_{\mathbf{p}}}[f(\mathcal{T})^2]. \end{aligned}\tag{8.16}$$

Furthermore, by the definition (1.8) of $\eta_{\mathbf{p}}(T,T')$ and (8.4),

$$\begin{aligned}&\sum_{T\in\mathbb{T}^{(K)}}\sum_{T'\in\mathbb{T}^{(K)}} f(T)f(T')\eta_{\mathbf{p}}(T,T')\pi_{\mathbf{p}}(T)\pi_{\mathbf{p}}(T')\\ &\quad= \Bigg(\sum_{T\in\mathbb{T}^{(K)}} f(T)(|T|-1)\pi_{\mathbf{p}}(T)\Bigg)^2 - \sum_{i\geq 0}\frac{1}{p_i}\Bigg(\sum_{T\in\mathbb{T}^{(K)}} f(T)n_T(i)\pi_{\mathbf{p}}(T)\Bigg)^2\\ &\quad= \Big(\mathbb{E}_{\pi_{\mathbf{p}}}[f(\mathcal{T})(|\mathcal{T}|-1)]\Big)^2 - \sum_{i\geq 0}\frac{1}{p_i}\Big(\mathbb{E}_{\pi_{\mathbf{p}}}[f(\mathcal{T})n_{\mathcal{T}}(i)]\Big)^2\mathbf{1}_{\{p_i>0\}}, \end{aligned}\tag{8.17}$$

where we recall Remark 8.2. Note also that all sums and expectations in (8.14)–(8.17) are finite, since f has finite support.

Finally, by combining (8.12) and (1.9)–(1.10), with (8.14), (8.16), and (8.17), we obtain (8.7). We have already remarked that all terms in (8.7) are finite. □

Note that we have not ruled out the possibility that $\gamma_{\mathbf{p}}(f)=0$. This may happen in trivial cases; whether it may happen in non-trivial cases (properly defined) is equivalent to Problem 1.9. We may restate that problem in a slightly stronger form:

Problem 8.3. *Let $\mathbf{p}\in\mathcal{P}_1(\mathbb{N}_0)$ be given. For which functionals f with finite support is $\gamma_{\mathbf{p}}(f)=0$?*

In analogy to [22] (which treats conditioned Galton–Watson trees), we may, more generally, also study additive functionals where the associated toll function does not necessarily have a finite support. We use a standard truncation argument. Let $f:\mathbb{T}\to\mathbb{R}$ be an arbitrary toll function. For a constant

$K > 0$, define the truncation and tail functionals associated to f by letting

$$f^{(K)}(T) := f(T)\mathbf{1}_{\{|T|\le K\}} \quad \text{and} \quad \widehat{f}^{(K)}(T) := f(T) - f^{(K)}(T), \quad T \in \mathbb{T}. \tag{8.18}$$

For $T \in \mathbb{T}$, we also let $F^{(K)}(T) := F(T, f^{(K)})$ and $\widehat{F}^{(K)}(T) := F(T, \widehat{f}^{(K)})$ be the additive functionals associated to $f^{(K)}$ and $\widehat{f}^{(K)}$, respectively; see (8.1).

Theorem 8.4. *Let $\mathbf{n}_\kappa$, $\kappa \ge 1$, be some degree statistics that satisfy Condition 1.1 and let $\mathcal{T}_{\mathbf{n}_\kappa} \sim \mathrm{Unif}(\mathbb{T}_{\mathbf{n}_\kappa})$. Suppose that $f : \mathbb{T} \to \mathbb{R}$ is a functional of rooted trees such that the following conditions are satisfied:*

(i) *There exists a real constant $0 \le \gamma < \infty$ such that $\lim_{K\to\infty} \gamma_{\mathbf{p}}(f^{(K)}) = \gamma$;*

(ii) $\lim_{K\to\infty} \limsup_{\kappa\to\infty} \frac{\mathrm{Var}(\widehat{F}^{(K)}(\mathcal{T}_{\mathbf{n}_\kappa}))}{|\mathbf{n}_\kappa|} = 0.$

Then, for the corresponding additive functional F,

$$\frac{F(\mathcal{T}_{\mathbf{n}_\kappa}) - \mathbb{E}[F(\mathcal{T}_{\mathbf{n}_\kappa})]}{\sqrt{|\mathbf{n}_\kappa|}} \to \mathrm{N}(0,\gamma), \quad \text{as } \kappa \to \infty. \tag{8.19}$$

Proof. Note that (8.8) in Theorem 8.1 applies to each $f^{(K)}$, as $\kappa \to \infty$. Then, (i), (ii) and [3, Theorem 3.2] (or [24, Theorem 4.28]) imply that we can let $K \to \infty$ and conclude (8.19). $\square$

The truncation argument used in the proof of Theorem 8.4 is the same as in [22], where critical conditioned Galton–Watson trees with finite offspring variance were studied (under some conditions on the toll functions). Indeed, the conditions in [22, Theorem 1.5] say roughly that the functional $f(T)$ is small when $|T|$ is large, and the proof there consists of verifying (under these conditions) results analogous to (i) and (ii) in Theorem 8.4. Furthermore, in [22, Theorem 1.13], asymptotic normality was also proved (in the same way) for additive functionals under the assumption that the toll function $f(T)$ is bounded and local (i.e., only depends on a fixed neighbourhood of the root of T). This was extended further by Ralaivaosaona et al. [29], who extended the result to "almost local" functionals (but assuming higher moments of the offspring distribution). For applications of our results above for random trees with given degree statistics, it would be useful to have some explicit sufficient conditions on f, similar to the conditions in the references just mentioned for conditioned Galton–Watson trees.

Problem 8.5. *Find suitable conditions on the functional f, and perhaps also on the degree statistics $\mathbf{n}_\kappa$, such that (i) and (ii) in Theorem 8.4 are satisfied, and thus (8.19) holds.*

Appendix A A multidimensional Gao–Wormald theorem

In this section, we generalize the Gao–Wormald theorem [13, Theorem 1] to the multidimensional setting. Note that a different but closely related multidimensional generalization of the Gao–Wormald theorem has been shown recently (and independently) by Hitczenko and Wormald [15]; the two multidimensional versions use essentially the same condition on (high) factorial moments, but the conditions and the result are stated in different ways. It is not clear exactly how the two versions are related, but it seems that the version in [15] is more flexible, and gives more precise results in cases when our asymptotic covariance matrix Γ is singular. On the other hand, it seems that our version is easier to apply in the present paper, and perhaps also in some other applications. (We thank the authors of [15] for interesting discussions on multidimensional versions of the Gao–Wormald theorem.)

For a sequence of real-valued random variables $(Z_n)_{n\geq 1}$, and a sequence of real numbers $(a_n)_{n\geq 1}$ such that $a_n > 0$, we write $Z_n = o_{\mathrm{p}}(a_n)$ when $Z_n/a_n \to 0$, as $n \to \infty$, in probability. For a complex number x we denote by $\mathrm{Re}(x)$ its real part.

Theorem A.1. *For $m, n \in \mathbb{N}$, let $(X_{1n}, \dots, X_{mn})$ be vectors of non-negative random variables. Suppose that μ_{in} and σ_{in} are positive real numbers such that for each $1 \leq i \leq m$, as $n \to \infty$,*

$$\sigma_{in} \ll \mu_{in} \ll \sigma_{in}^3. \tag{A.1}$$

Let $\Gamma := (\gamma_{ij})_{i,j=1}^m$ be a fixed matrix. Let $c > 0$ be a constant, and suppose further that, as $n \to \infty$, uniformly for all integer sequences $(k_{in})_{i=1}^m$ with $0 \leq k_{in} \leq c\mu_{in}/\sigma_{in}$,

$$\mathbb{E}\prod_{i=1}^m (X_{in})_{k_{in}} = \prod_{i=1}^m \mu_{in}^{k_{in}} \cdot \exp\left(\frac{1}{2}\sum_{i,j=1}^m \frac{\gamma_{ij}\sigma_{in}\sigma_{jn} - \delta_{ij}\mu_{in}}{\mu_{in}\mu_{jn}} k_{in}k_{jn} + o(1)\right), \tag{A.2}$$

Then,

$$\left(\frac{X_{1n} - \mu_{1n}}{\sigma_{1n}}, \dots, \frac{X_{mn} - \mu_{mn}}{\sigma_{mn}}\right) \xrightarrow{\mathrm{d}} \mathrm{N}(0, \Gamma), \quad \text{as } n \to \infty. \tag{A.3}$$

Proof. We follow closely the proof of the one-dimensional version in [13]. In the following, unspecified limits are as $n \to \infty$. Note first that (A.1) implies $\sigma_{in} \to \infty$ and $\mu_{in} \to \infty$. For $1 \leq i \leq m$, we let

$$\zeta_{in} := \frac{\mu_{in}}{\sigma_{in}} \ln\left(\frac{X_{in}}{\mu_{in}}\right), \tag{A.4}$$

$$Q_{in}(k) := \frac{(X_{in})_k}{(\mu_{in})_k}, \tag{A.5}$$

$$t_{in} := k_{in}\frac{\sigma_{in}}{\mu_{in}}. \tag{A.6}$$

By Lemma 4.1 and (A.1), the assumption (A.2) implies, uniformly for allowed sequences $(k_{in})_{i=1}^m$,

$$\mathbb{E}\prod_{i=1}^m Q_{in}(k_{in}) = \exp\left(\frac{1}{2}\sum_{i,j=1}^m \frac{\gamma_{ij}\sigma_{in}\sigma_{jn}k_{in}k_{jn}}{\mu_{in}\mu_{jn}} + o(1)\right) = \exp\left(\frac{1}{2}\sum_{i,j=1}^m \gamma_{ij}t_{in}t_{jn}\right) + o(1). \tag{A.7}$$

Note that (A.6) implies that $0 \leq t_{in} \leq c$, and thus $t_{in} = O(1)$. In particular, the expectation in (A.7) is $O(1)$.

We recall the inequality [13, (2.8)], which says that if $a \geq b > k$, then

$$k \ln(a/b) \leq \ln\left(\frac{(a)_k}{(b)_k}\right) \leq k \log\left(\frac{a-k}{b-k}\right) \leq \frac{k \ln(a/b)}{1-k/b}. \tag{A.8}$$

Note that $k_{in}/\mu_{in} \leq c/\sigma_{in} \to 0$, for $1 \leq i \leq m$. In the sequel, we will consider only n that are large enough so that $\sigma_{in} > 2c$, for every $1 \leq i \leq m$. Hence, for allowed sequences $(k_{in})_{i=1}^m$, we have $k_{in}/\mu_{in} < 1/2$. Then, when $X_{in} \geq \mu_{in}$, (A.8) yields $2 \ln Q_{in}(k_{in}) \leq 4k_{in} \ln(X_{in}/\mu_{in}) \leq \ln Q_{in}(4k_{in})$. Thus, when $X_{in} \geq \mu_{in}$,

$$Q_{in}(k_{in})^2 \leq Q_{in}(4k_{in}). \tag{A.9}$$

Furthermore, if $k_{in} \leq X_{in} \leq \mu_{in}$, we trivially have $Q_{in}(k_{in}) \leq 1$. Similarly, it is easy to see that if $0 \leq X_{in} < k_{in}$, then $|(X_{in})_{k_{in}}| \leq (k_{in})_{k_{in}} \leq (\mu_{in})_{k_{in}}$ and thus $|Q_{in}(k_{in})| \leq 1$ in this case too. (This is trivial if X_{in} is integer-valued.) By combining these observations with (A.9) for the case $X_{in} \geq \mu_{in}$, we see that for any allowed sequences $(k_{in})_{i=1}^m$, and any X_{in},

$$Q_{in}(k_{in})^2 \leq Q_{in}(4k_{in}) + 1 = Q_{in}(4k_{in}) + Q_{in}(0). \tag{A.10}$$

We set $c' := c/4$, and we henceforth consider only sequences $(k_{in})_{i=1}^m$ such that

$$k_{in} \leq c' \mu_{in}/\sigma_{in}. \tag{A.11}$$

Then (A.2) and, as a consequence, (A.7) hold also with $4k_{in}$ or 0 instead of k_{in}. Hence, (A.10) implies that

$$\mathbb{E} \prod_{i=1}^m \left(1 + Q_{in}(k_{in})^2\right) \leq C, \tag{A.12}$$

for some constant C. (We similarly use C below to denote unknown constants, possibly with different values on each occasion.)

For $1 \leq i \leq m$, we set $\varepsilon_{in} := 2c'/\sigma_{in}$. Then, (A.11) implies $k_{in}/(\mu_{in}/2) \leq \varepsilon_{in}$. Note that $\varepsilon_{in} < 2c'/(2c) = 1/4$ and $\varepsilon_{in} \to 0$, as $n \to \infty$. It is shown in [13] that

$$\left|e^{t_{in}\zeta_{in}} - Q_{in}(k_{in})\right| \leq \begin{cases} \min(Q_{in}(k_{in}), \varepsilon_{in} Q_{in}(k_{in}) \ln(Q_{in}(k_{in}))) & \text{if} \quad X_{in} \geq \mu_{in}, \\ \varepsilon_{in} Q_{in}(k_{in})^{1-\varepsilon_{in}} \ln(1/Q_{in}(k_{in})) & \text{if} \quad \mu_{in}/2 \leq X_{in} < \mu_{in}, \\ 2^{-k_{in}} & \text{if} \quad 0 \leq X_{in} < \mu_{in}/2. \end{cases} \tag{A.13}$$

Note also that it follows from (A.5) that $X_{in} \geq \mu_{in}$ if and only if $Q_{in} \geq 1$. By considering the four cases $X_{in} < \mu_{in}/2$, $\mu_{in}/2 \leq X_{in} < \mu_{in}$, $1 \leq Q_{in} \leq \varepsilon_{in}^{-1/2}$ and $Q_{in} > \varepsilon_{in}^{-1/2}$, it follows from (A.14) that

$$\begin{aligned} \left|e^{t_{in}\zeta_{in}} - Q_{in}(k_{in})\right| &\leq 2^{-k_{in}} + C\varepsilon_{in} + C\varepsilon_{in}^{1/2} \ln(\varepsilon_{in}^{-1/2}) + Q_{in}(k_{in}) \mathbf{1}_{\{Q_{in}(k_{in}) > \varepsilon_{in}^{-1/2}\}} \\ &\leq 2^{-k_{in}} + C\varepsilon_{in}^{1/3} + \varepsilon_{in}^{1/2} Q_{in}(k_{in})^2 \\ &\leq C(2^{-k_{in}} + C\varepsilon_{in}^{1/3})(1 + Q_{in}(k_{in})^2). \end{aligned} \tag{A.14}$$

Furthermore, (A.8) shows that if $X_{in} \geq \mu_{in}$, then $e^{t_{in}\zeta_{in}} = (X_{in}/\mu_{in})^{k_{in}} \leq Q_{in}(k_{in})$. Hence, for any X_{in},

$$0 < e^{t_{in}\zeta_{in}} \leq 1 + Q_{in}(k_{in}). \tag{A.15}$$

It follows from (A.14) and (A.15) that

$$\begin{aligned}\left|\prod_{i=1}^{m} e^{t_{in}\zeta_{in}} - \prod_{i=1}^{m} Q_{in}(k_{in})\right| &\leq \sum_{i=1}^{m} \left|e^{t_{in}\zeta_{in}} - Q_{in}(k_{in})\right| \prod_{j\neq i}(1 + Q_{jn}(k_{jn})) \\ &\leq C\sum_{i=1}^{m}(2^{-k_{in}} + C\varepsilon_{in}^{1/3}) \cdot \prod_{i=1}^{m}(1 + Q_{in}(k_{in})^2).\end{aligned} \tag{A.16}$$

Now, we further restrict to sequences $(k_{in})_{i=1}^{m}$ such that $c''\mu_{in}/\sigma_{in} \leq k_{in} \leq c'\mu_{in}/\sigma_{in}$, where $c'' := c'/2$. Then $2^{-k_{in}} \to 0$, uniformly for all allowed sequences $(k_{in})_{i=1}^{m}$. Thus, (A.16) implies, uniformly,

$$\left|\prod_{i=1}^{m} e^{t_{in}\zeta_{in}} - \prod_{i=1}^{m} Q_{in}(k_{in})\right| = o(1) \cdot \prod_{i=1}^{m}(1 + Q_{in}(k_{in})^2). \tag{A.17}$$

It then follows from (A.12) and (A.7) that, uniformly for all allowed sequences $(k_{in})_{i=1}^{m}$,

$$\mathbb{E}\prod_{i=1}^{m} e^{t_{in}\zeta_{in}} = \exp\left(\frac{1}{2}\sum_{i,j=1}^{m}\gamma_{ij}t_{in}t_{jn}\right) + o(1). \tag{A.18}$$

We claim that (A.18) implies that

$$(\zeta_{1n},\ldots,\zeta_{mn}) \xrightarrow{\mathrm{d}} \mathrm{N}(0,\Gamma), \quad \text{as } n \to \infty. \tag{A.19}$$

Then, by the definition (A.4), since ζ_{in} is tight by (A.19) and $\sigma_{in}/\mu_{in} \to 0$,

$$\frac{X_{in} - \mu_{in}}{\sigma_{in}} = \frac{\mu_{in}}{\sigma_{in}}\left(\exp\left(\frac{\sigma_{in}}{\mu_{in}}\zeta_{in}\right) - 1\right) = \zeta_{in} + o_{\mathrm{p}}(1), \tag{A.20}$$

for $1 \leq i \leq m$. Thus, the result (A.3) follows from (A.19).

The rest of the proof is dedicated to prove the claim in (A.19). The proof is routine, although we do not know any reference. So, we include the argument for completeness.

Note that (A.18) holds for all $t_{in} \in [c'', c']$ such that $t_{in}\mu_{in}/\sigma_{in}$ is an integer. By assumption, $\mu_{in}/\sigma_{in} \to \infty$, and thus the gaps between the allowed t_{in} tend to 0 as $n \to \infty$; in particular, there exist such t_{in} for large enough n. Define $Y_{in} := e^{\zeta_{in}}$, for $1 \leq i \leq m$. It follows easily from (A.18) (or, rather, the one-dimensional version, which follows as above) that the sequences $(Y_{in})_{n\geq 1}$ are tight, for $1 \leq i \leq m$, and thus so is the sequence of random vectors $(Y_{1n},\ldots,Y_{mn})_{n\geq 1}$. Consider a subsequence such that $(Y_{1n},\ldots,Y_{mn}) \xrightarrow{\mathrm{d}} (Y_1,\ldots,Y_m)$, for some random variables $Y_1,\ldots,Y_m$. Set $c''' := 0.8c'$, say. Fix any real numbers $t_i \in [c'', c''']$, for $1 \leq i \leq m$. For sufficiently large n, we may find $t_{in} \in [c'', c''']$ such that $t_{in} \to t_i$ and $t_{in}\mu_{in}/\sigma_{in}$ are integers, and also $t'_{in} \in [c'', 1.1c'']$ and $t''_{in} \in [0.9c', c']$ such that $t'_{in}\mu_{in}/\sigma_{in}$ and $t''_{in}\mu_{in}/\sigma_{in}$ are integers. Then,

$t'_{in} \le 1.1 t_{in} \le t''_{in}$, and thus, using (A.18),

$$\mathbb{E}\left(\prod_{i=1}^m Y_{in}^{t_{in}}\right)^{1.1} = \mathbb{E}\left[\prod_{i=1}^m Y_{in}^{1.1 t_{in}}\right] \le \mathbb{E}\left[\prod_{i=1}^m \left(Y_{in}^{t'_{in}} + Y_{in}^{t''_{in}}\right)\right] \le C. \tag{A.21}$$

Hence, the sequence $\prod_{i=1}^m Y_{in}^{t_{in}}$ is uniformly integrable. Furthermore $\prod_{i=1}^m Y_{in}^{t_{in}} \xrightarrow{\mathrm{d}} \prod_{i=1}^m Y_i^{t_i}$ (along the subsequence), and thus it follows from (A.18) that

$$\mathbb{E}\prod_{i=1}^m Y_i^{t_i} = \lim_{n\to\infty} \mathbb{E}\prod_{i=1}^m Y_{in}^{t_{in}} = \exp\left(\frac{1}{2}\sum_{i,j=1}^m \gamma_{ij} t_i t_j\right). \tag{A.22}$$

This holds for any $t_1,\dots,t_m \in [c'',c''']$. Since the expectation on the left-hand side of (A.22) thus is finite for these $t_1,\dots,t_m$, it follows that it is finite for all complex $t_1,\dots,t_m$ with real parts in (c'',c'''), and that it is a bounded analytic function in this domain. By analytic continuation, we thus have

$$\mathbb{E}\prod_{i=1}^m Y_i^{t_i} = \exp\left(\frac{1}{2}\sum_{i,j=1}^m \gamma_{ij} t_i t_j\right), \qquad \mathrm{Re}(t_1),\dots,\mathrm{Re}(t_m) \in (c'',c'''). \tag{A.23}$$

Furthermore, by taking $t_i = \frac{1}{2}(c''+c''') + \mathrm{i}u_i$ with real numbers u_i, it follows from the boundedness of (A.23) that the matrix $\Gamma := (\gamma_{ij})_{i,j=1}^m$ is positive semi-definite. Hence the multivariate normal distribution $\mathrm{N}(0,\Gamma)$ exists. For $1 \le i \le m$, let $\zeta_i := \ln(Y_i) \in [-\infty,\infty)$, and let $(\hat\zeta_1,\dots,\hat\zeta_m) \sim \mathrm{N}(0,\Gamma)$. Then, (A.23) can be written as

$$\mathbb{E}\prod_{i=1}^m e^{t_i\zeta_i} = \mathbb{E}\prod_{i=1}^m Y_i^{t_i} = \mathbb{E}\prod_{i=1}^m e^{t_i\hat\zeta_i}, \qquad \mathrm{Re}(t_1),\dots,\mathrm{Re}(t_m) \in (c'',c'''). \tag{A.24}$$

Let ν be the distribution of $(\zeta_1,\dots,\zeta_m)$; note that this is a probability measure on $[-\infty,\infty)^m$. Also, let $\hat\nu$ be the distribution $\mathrm{N}(0,\Gamma)$. Fix some $\tau \in (c'',c''')$, and define the conjugated measures ν^* and $\hat\nu^*$ on $\mathbb{R}^m$ by

$$\frac{\mathrm{d}\nu^*}{\mathrm{d}\nu}(x_1,\dots,x_m) = \frac{\mathrm{d}\hat\nu^*}{\mathrm{d}\hat\nu}(x_1,\dots,x_m) = e^{\sum_{i=1}^m \tau x_i}, \qquad x_1,\dots,x_m \in \mathbb{R}. \tag{A.25}$$

Then (A.24) (with $\mathrm{Re}(t_i) = \tau$ for each $1 \le i \le m$) implies that the finite measures ν^* and $\hat\nu^*$ have the same Fourier transform; consequently, $\nu^* = \hat\nu^*$. This implies that the measures ν and $\hat\nu$ coincide on $\mathbb{R}^m$. Since $\hat\nu$ is a probability measure, this shows that $\nu(\mathbb{R}^m) = 1$, and thus ν is concentrated on $\mathbb{R}^m$. In other words, $\zeta_i > -\infty$ a.s., and $(\zeta_1,\dots,\zeta_m) \sim \mathrm{N}(0,\Gamma)$.

Consequently, we have shown that

$$(e^{\zeta_{1n}},\dots,e^{\zeta_{mn}}) = (Y_{1n},\dots,Y_{mn}) \xrightarrow{\mathrm{d}} (Y_1,\dots,Y_m) \overset{\mathrm{d}}{=} (e^{\hat\zeta_1},\dots,e^{\hat\zeta_m}), \quad \text{as } n\to\infty, \tag{A.26}$$

along every convergent subsequence, and thus for the full sequence. Then, the claim (A.19) follows by taking logarithms. $\square$

Remark A.2. The original theorem by Gao and Wormald [13, Theorem 1] is essentially the case $m = 1$ of Theorem A.1. (The condition $\sigma_{in} \ll \mu_{in}$ is omitted in [13] by mistake, but it is clearly needed.) However, there is a technical difference in that they assume (A.2) (for $m = 1$) only for $c_1\mu_n/\sigma_n \le k_n \le c\mu_n/\sigma_n$, for some arbitrary $0 < c_1 < c$, while we assume it for all $k_n \le c\mu_n/\sigma_n$.

In fact, the proof above, with minor modifications, shows that in the multivariate version in Theorem A.1, it suffices to assume (A.2) for $k_{in} \in [c_1\mu_{in}/\sigma_{in}, c\mu_{in}/\sigma_{in}] \cup \{0\}$, i.e., it suffices to assume, for each i, that either $k_i = 0$ or k_i is in a range as in [13]. In other words, it suffices to restrict k_{in} as in [13] if we also assume that (A.2) holds for the vector $(X_{in})_{i\in J}$, for every subset J of $\{1,\dots,m\}$.

Appendix B Proof of Theorem 7.6

In this appendix, we give a new, simple, proof of Theorem 7.6 under condition (i) or (ii) there (stated in Theorem 7.2), by using the multivariate Gao–Wormald theorem in Appendix A.

Proof of Theorem 7.6 assuming (i) or (ii). We consider first a general weight sequence $\mathbf{w}$, assuming only $\rho_{\mathbf{w}} > 0$. By Remark 7.3, we may replace $\mathbf{w}$ by the probability distribution $\theta(\mathbf{w})$. In other words, we may and will assume that $\mathbf{w}$ is a probability distribution on $\mathbb{N}_0$; thus $\mathcal{T}_{\mathbf{w},n}$ is a conditioned Galton–Watson tree with this offspring distribution.

Let $\xi_1, \xi_2, \dots$ be a sequence of independent random variables with distribution $\mathbf{w}$, and define the partial sums

$$S_n := \sum_{j=1}^{n} \xi_j, \qquad n \ge 0, \tag{B.1}$$

and

$$\widetilde{S}_n := \sum_{j=1}^{n} (\xi_j - 1) = S_n - n, \qquad n \ge 0. \tag{B.2}$$

There is a well-known bijection between $\mathbb{T}_n$ and the set $\mathbb{E}^n$ of excursions of length n, which combines the bijections between $\mathbb{T}_{\mathbf{n}}$ and $\mathbb{E}_{\mathbf{n}}$ mentioned in Section 2 for different $\mathbf{n}$ with $|\mathbf{n}| = n$. Let us denote this bijection by Υ. By (2.5), for any tree $T \in \mathbb{T}_n$, the corresponding excursion $\Upsilon(T) \in \mathbb{E}^n$ has increments that are the vertex degrees minus 1, i.e., $d_T(i) - 1$, $i = 1\dots,n$, in some order. It follows that for the conditioned Galton–Watson tree $\mathcal{T}_{\mathbf{w},n}$, the random excursion $\Upsilon(\mathcal{T}_{\mathbf{w},n}) \in \mathbb{E}^n$ has a distribution that equals the distribution of $\widetilde{\mathbb{S}}_n := (\widetilde{S}_0, \dots, \widetilde{S}_n)$ conditioned on $\widetilde{\mathbb{S}}_n \in \mathbb{E}^n$. This means that the degree statistic $\mathbf{n}_{\mathcal{T}_{\mathbf{w},n}}$ has the same distribution as the sequence $\bigl(|\{1 \le j \le n : \xi_j = i\}| : i \ge 0\bigr)$ conditioned on $\widetilde{\mathbb{S}}_n \in \mathbb{E}^n$. Furthermore, since the Vervaat transformation (see Section 2) is an n-to-1 map of the set of bridges $\mathbb{B}^n$ to $\mathbb{E}^n$, which only permutes the set of increments, we also have the equality in distribution

$$\begin{aligned} \mathbf{n}_{\mathcal{T}_{\mathbf{w},n}} = (n_{\mathcal{T}_{\mathbf{w},n}}(i) : i \ge 0) &\overset{\mathrm{d}}{=} \Bigl((|\{1 \le j \le n : \xi_j = i\}| : i \ge 0) \mid \widetilde{\mathbb{S}}_n \in \mathbb{B}^n\Bigr) \\ &= \Bigl((|\{1 \le j \le n : \xi_j = i\}| : i \ge 0) \mid S_n = n - 1\Bigr), \end{aligned} \tag{B.3}$$

where the final equality follows from (2.1) and (B.2). We use this to compute factorial moments of $n_{\mathcal{T}_{\mathbf{w},n}}(i)$. Fix $k \in \mathbb{N}_0$, and let $q_{0n}, \dots, q_{kn}$ be non-negative integers. Then (B.3) shows that $\prod_{i=0}^{k}(n_{\mathcal{T}_{\mathbf{w},n}}(i))_{q_{in}}$ has the same distribution as the number of arrays of distinct indices $(j_{i\ell} : 0 \le i \le k, 1 \le \ell \le q_{in})$ such that $\xi_{j_{i\ell}} = i$ for all i and ℓ, conditioned on $S_n = n-1$. For each choice of indices $(j_{i\ell})$, the probability that $\xi_{j_{i\ell}} = i$ for all i and ℓ is $\prod_{i=0}^{k} w_i^{q_{in}}$, and $S_n = n-1$ if and only if the sum of the remaining $n - \sum_{i=0}^{k} q_{in}$ variables ξ_j equals $n - 1 - \sum_{i=0}^{k} i q_{in}$. Hence, (B.3) yields the factorial moments

$$\mathbb{E}\prod_{i=0}^{k}(n_{\mathcal{T}_{\mathbf{w},n}}(i))_{q_{in}} = (n)_{\sum_{i=0}^{k} q_{in}} \prod_{i=0}^{k} w_i^{q_{in}} \cdot \frac{\mathbb{P}\Big(S_{n-\sum_{i=0}^{k} q_{in}} = n-1-\sum_{i=0}^{k} i q_{in}\Big)}{\mathbb{P}(S_n = n-1)}. \tag{B.4}$$

Note that we may ignore indices i with $w_i = 0$ in (7.12), since then $n_{\mathcal{T}_{\mathbf{w},n}}(i) = 0 = \theta_i(\mathbf{w})$, while (7.13)–(7.14) yield $\gamma^*(i,j) = 0$ for every $0 \le j \le k$. Hence, in the sequel we assume $q_{in} = 0$ when $w_i = 0$, which means that we really consider only $0 \le i \le k$ with $w_i > 0$. Also, assume that $q_{in} \le C\sqrt{n}$ for all $0 \le i \le k$ and some fixed constant $C > 0$. We estimate below (B.4) as $n \to \infty$; the estimates will be uniform for all such $(q_{in})_{i=0}^{k}$.

Let $\mu_{in} := n w_i$ and $\sigma_{in} := \sqrt{n}$, for $0 \le i \le k$. Then (B.4) and Lemma 4.1 yield

$$\mathbb{E}\prod_{i=0}^{k}(n_{\mathcal{T}_{\mathbf{w},n}}(i))_{q_{in}} = \prod_{i=0}^{k} \mu_{in}^{q_{in}} \cdot \exp\left(-\frac{1}{2n}\Big(\sum_{i=0}^{k} q_{in}\Big)^2 + O(n^{-1/2})\right) \frac{\mathbb{P}\Big(S_{n-\sum_{i=0}^{k} q_{in}} = n-1-\sum_{i=0}^{k} i q_{in}\Big)}{\mathbb{P}(S_n = n-1)}. \tag{B.5}$$

Hence we see that what remains to apply Theorem A.1 is a suitable local limit theorem for the sums $(S_n, n \ge 0)$. We consider the cases (i) and (ii) separately.

(i): In this case, the distribution $\mathbf{w} = \theta(\mathbf{w})$ has mean 1 and finite variance $\sigma_{\mathbf{w}}^2 > 0$. Let $h \ge 1$ be the span of this distribution, i.e., the largest integer such that ξ_1 a.s. is a multiple of h. Then $Z_n(\mathbf{w}) > 0$ only if $n \equiv 1 \pmod{h}$, so we consider only such n; moreover, by assumption $q_{in} > 0$ only if $w_i > 0$ and thus $i \equiv 0 \pmod{h}$. Hence, the standard local central limit theorem, see e.g. [27, Theorem 7.1], yields, recalling $q_{in} = O(\sqrt{n})$,

$$\mathbb{P}\Big(S_{n-\sum_{i=0}^{k} q_{in}} = n-1-\sum_{i=0}^{k} i q_{in}\Big) = \frac{h}{\sqrt{2\pi\sigma_{\mathbf{w}}^2 n}} \exp\left(-\frac{\Big(\sum_{i=0}^{k} i q_{in} - \sum_{i=0}^{k} q_{in}\Big)^2}{2\sigma_{\mathbf{w}}^2 n} + o(1)\right), \tag{B.6}$$

$$\mathbb{P}\Big(S_n = n-1\Big) = \frac{h}{\sqrt{2\pi\sigma_{\mathbf{w}}^2 n}} \exp\left(o(1)\right). \tag{B.7}$$

Hence, (B.5) yields, recalling (7.13)–(7.14) and our assumption $\theta(\mathbf{w}) = \mathbf{w}$,

$$\begin{aligned}\mathbb{E}\prod_{i=0}^{k}(n_{\mathcal{T}_{\mathbf{w},n}}(i))_{q_{in}} &= \prod_{i=0}^{k} \mu_{in}^{q_{in}} \cdot \exp\left(-\frac{\Big(\sum_{i=0}^{k} q_{in}\Big)^2 + \Big(\sum_{i=0}^{k}(i-1) q_{in}\Big)^2/\sigma_{\mathbf{w}}^2}{2n} + o(1)\right) \\ &= \prod_{i=0}^{k} \mu_{in}^{q_{in}} \cdot \exp\left(\frac{1}{2}\sum_{i,j=0}^{k} \frac{n\gamma^*(i,j) - \delta_{ij}\mu_{in}}{\mu_{in}\mu_{jn}} q_{in} q_{jn} + o(1)\right).\end{aligned} \tag{B.8}$$

Hence, Theorem 7.6 follows by Theorem A.1 (with $\sigma_{in} := \sqrt{n}$) in case (i).

(ii): This is similar. By assumption, there exist sequences of constants $a_n > 0$ and b_n such that $(S_n - b_n)/a_n$ converges in distribution to some stable random variable Y of index $\alpha \in (1,2]$. Let again h be the span of the distribution $\mathbf{w}$, and let $g_Y(x)$ be the density function of the stable limit Y. Then, a local limit theorem holds, see e.g. [14, § 50]. Since we assume that the variance of ξ_1 is infinite, we have $a_n \gg n^{1/2}$, see [9, XVII.(5.23)], and thus $q_{in} = o(a_n)$, and then the local limit theorem yields, simply,

$$\mathbb{P}\Big(S_{n-\sum_{i=0}^k q_{in}} = n-1-\sum_{i=0}^k iq_{in}\Big) = \frac{h}{a_n} g_Y(0)\exp(o(1)), \tag{B.9}$$

$$\mathbb{P}\Big(S_n = n-1\Big) = \frac{h}{a_n} g_Y(0)\exp(o(1)), \tag{B.10}$$

and hence, since $g_Y(0) > 0$ (see e.g. [32, Remark 4 after Theorem 2.2.3, pp. 79–80]),

$$\frac{\mathbb{P}\Big(S_{n-\sum_{i=0}^k q_{in}} = n-1-\sum_{i=0}^k iq_{in}\Big)}{\mathbb{P}(S_n = n-1)} \to 1, \tag{B.11}$$

as $n \to \infty$. It follows from (B.5) that (B.8) holds in this case too, now with $\sigma_{\mathbf{w}}^2 = \infty$, and the proof in case (ii) is completed as above by Theorem A.1. □

Remark B.1. We see from the proof above that Theorem 7.6 holds for any weight sequence $\mathbf{w}$ such that the probability distribution $\theta(\mathbf{w})$ satisfies a nice local limit theorem for the sums $(S_n, n \geq 0)$. We do not know whether that holds in full generality, but we note that any extension of the local limit theorems used above gives a partial answer to Problem 7.8. However, it is conceivable that there are weight sequences for which such a local limit theorem fails, but nevertheless Theorem 7.6 holds.

References

[1] L. Addario-Berry and S. Donderwinkel, *Random trees have height $O(\sqrt{n})$*, arXiv e-prints (2022), arXiv:2201.11773.

[2] D. Aldous, *Asymptotic fringe distributions for general families of random trees*, Ann. Appl. Probab. **1** (1991), no. 2, 228–266. MR 1102319

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

[4] X. S. Cai and L. Devroye, *A study of large fringe and non-fringe subtrees in conditional Galton-Watson trees*, ALEA Lat. Am. J. Probab. Math. Stat. **14** (2017), no. 1, 579–611. MR 3667924

[5] H. Chang and M. Fuchs, *Limit theorems for patterns in phylogenetic trees*, J. Math. Biol. **60** (2010), no. 4, 481–512. MR 2587587

[6] L. Devroye, *Limit laws for local counters in random binary search trees*, Random Structures Algorithms **2** (1991), no. 3, 303–315. MR 1109697

[7] ______, *Limit laws for sums of functions of subtrees of random binary search trees*, SIAM J. Comput. **32** (2002/03), no. 1, 152–171. MR 1954858

[8] M. Drmota, *Random trees, an interplay between combinatorics and probability*, SpringerWienNewYork, Vienna, 2009. MR 2484382

[9] W. Feller, *An introduction to probability theory and its applications. Vol. II*, second ed., John Wiley & Sons, Inc., New York-London-Sydney, 1971. MR 0270403

[10] Q. Feng and H. M. Mahmoud, *On the variety of shapes on the fringe of a random recursive tree*, J. Appl. Probab. **47** (2010), no. 1, 191–200. MR 2654767

[11] P. Flajolet, X. Gourdon, and C. Martínez, *Patterns in random binary search trees*, Random Structures Algorithms **11** (1997), no. 3, 223–244. MR 1609509

[12] M. Fuchs, *Limit theorems for subtree size profiles of increasing trees*, Combin. Probab. Comput. **21** (2012), no. 3, 412–441. MR 2912789

[13] Z. Gao and N. C. Wormald, *Asymptotic normality determined by high moments, and submap counts of random maps*, Probab. Theory Related Fields **130** (2004), no. 3, 368–376. MR 2095934

[14] B. V. Gnedenko and A. N. Kolmogorov, *Limit distributions for sums of independent random variables*, Addison-Wesley Publishing Co., Inc., Cambridge, Mass., 1954, Translated and annotated by K. L. Chung. MR 0062975

[15] P. Hitczenko and N. Wormald, *Multivariate asymptotic normality determined by high moments*, Preprint (2023).

[16] C. Holmgren and S. Janson, *Limit laws for functions of fringe trees for binary search trees and random recursive trees*, Electron. J. Probab. **20** (2015), no. 4, 51. MR 3311217

[17] ______, *Fringe trees, Crump-Mode-Jagers branching processes and m-ary search trees*, Probab. Surv. **14** (2017), 53–154. MR 3626585

[18] C. Holmgren, S. Janson, and M. Šileikis, *Multivariate normal limit laws for the numbers of fringe subtrees in m-ary search trees and preferential attachment trees*, Electron. J. Combin. **24** (2017), no. 2, Paper No. 2.51, 49. MR 3690263

[19] L. Holst, *A unified approach to limit theorems for urn models*, J. Appl. Probab. **16** (1979), no. 1, 154–162. MR 520945

[20] S. Janson, *Moment convergence in conditional limit theorems*, J. Appl. Probab. **38** (2001), no. 2, 421–437. MR 1834751

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

[22] ———, *Asymptotic normality of fringe subtrees and additive functionals in conditioned Galton-Watson trees*, Random Structures Algorithms **48** (2016), no. 1, 57–101. MR 3432572

[23] ———, *Central limit theorems for additive functionals and fringe trees in tries*, Electron. J. Probab. **27** (2022), Paper No. 47, 1–63. MR 4408126

[24] O. Kallenberg, *Foundations of modern probability*, second ed., Probability and its Applications (New York), Springer-Verlag, New York, 2002. MR 1876169

[25] V. F. Kolchin, *Random mappings*, Translation Series in Mathematics and Engineering, Optimization Software, Inc., Publications Division, New York, 1986, Translated from the Russian. MR 865130

[26] N. Minami, *On the number of vertices with a given degree in a Galton-Watson tree*, Adv. in Appl. Probab. **37** (2005), no. 1, 229–264. MR 2135161

[27] V. V. Petrov, *Sums of independent random variables*, Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 82, Springer-Verlag, New York-Heidelberg, 1975, Translated from the Russian by A. A. Brown. MR 0388499

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

[29] D. Ralaivaosaona, M. Šileikis, and S. Wagner, *A central limit theorem for almost local additive tree functionals*, Algorithmica **82** (2020), no. 3, 642–679. MR 4058420

[30] L. Takács, *A generalization of the ballot problem and its application in the theory of queues*, J. Amer. Statist. Assoc. **57** (1962), 327–337. MR 138139

[31] P. Thévenin, *Vertices with fixed outdegrees in large Galton-Watson trees*, Electron. J. Probab. **25** (2020), Paper No. 64, 25. MR 4115733

[32] V. M. Zolotarev, *One-dimensional stable distributions*, Translations of Mathematical Monographs, vol. 65, American Mathematical Society, Providence, RI, 1986, Translated from the Russian by H. H. McFaden, Translation edited by Ben Silver. MR 854867