A study of large fringe and non-fringe subtrees in conditional Galton-Watson trees

Xing Shi Cai and Luc Devroye

School of Computer Science, McGill University, Montreal, Canada.
E-mail address: `xingshi.cai@mail.mcgill.ca`

School of Computer Science, McGill University, Montreal, Canada.
E-mail address: `lucdevroye@gmail.com`

Abstract. We study the conditions for families of fringe or non-fringe subtrees to exist with high probability (whp) in $\mathcal{T}_n^{\mathbf{gw}}$, a Galton-Walton tree of size n. We first give a Poisson approximation of fringe subtree counts in $\mathcal{T}_n^{\mathbf{gw}}$, which permits us to determine the height of the maximal complete r-ary fringe subtree. Then we determine the maximal K_n such that every tree of size at most K_n appears as a fringe subtree in $\mathcal{T}_n^{\mathbf{gw}}$ whp. Finally, we show that non-fringe subtree counts are concentrated and determine, as an application, the height of the maximal complete r-ary non-fringe subtree in $\mathcal{T}_n^{\mathbf{gw}}$.

1. Introduction

In this paper, we study the conditions for families of fringe or non-fringe subtrees to exist whp (with high probability) in a Galton-Walton tree conditional to be of size n. In particular, we want to find the height of the maximal complete r-ary fringe and non-fringe subtrees. We also want to determine the threshold k_n such that all trees of size at most k_n appear as fringe subtrees. In doing so, we extend Janson (2016) result on fringe subtrees counts and prove a new concentration theorem for non-fringe subtree counts.

Let $\mathfrak{T}$ be the set of all rooted, ordered, and unlabeled trees, which we refer to as *plane* trees. All trees considered in this paper belong to $\mathfrak{T}$. (See Janson, 2012, sec. 2.1 for details.)

Given a tree $T \in \mathfrak{T}$ and a node $v \in T$, let T_v denote the subtree rooted at v. We call T_v a *fringe* subtree of T. If T_v is isomorphic to some tree $T' \in \mathfrak{T}$, then we write $T' = T_v$ and say that T has a *fringe* subtree of shape T' rooted at v, or simply T contains T' as a fringe subtree.

On the other hand, if T_v can be made isomorphic to T' by replacing some or none of its own fringe subtrees with leaves (nodes without children), then we write $T' \prec T_v$ and say that T has a *non-fringe* subtree of shape T' rooted at v, or simply T contains T' as a non-fringe subtree. (Note that $T' = T_v$ implies that $T' \prec T_v$.) We also use the notation $T' \prec T$ to denote that T has a non-fringe subtree of shape T' at its root.

Let ξ be a non-negative integer-valued random variable. The Galton-Watson tree $\mathcal{T}^{\mathbf{gw}}$ with offspring distribution ξ is the random tree generated by starting from the root and independently giving each node a random number of children, where the numbers of children are all distributed as ξ. The conditional Galton-Watson tree $\mathcal{T}_n^{\mathbf{gw}}$ is $\mathcal{T}^{\mathbf{gw}}$ restricted to the event $|\mathcal{T}^{\mathbf{gw}}| = n$, i.e., $\mathcal{T}^{\mathbf{gw}}$ has n nodes. The comprehensive survey by Janson (2012) describes the history and the basic properties of these trees.

In the study of conditional Galton-Watson trees, the following condition is assumed throughout the paper:

Condition A. Let $\mathcal{T}_n^{\mathbf{gw}}$ be a conditional Galton-Watson tree of size n with offspring distribution ξ, such that $\mathbb{E}\xi = 1$ and $0 < \sigma^2 := \mathrm{Var}\,(\xi) < \infty$. Let $\mathcal{T}^{\mathbf{gw}}$ be the corresponding unconditional Galton-Watson tree.

We summarize our notation:

· $\mathfrak{T}$ — the set of all rooted, ordered and unlabeled trees (plane trees)
· T — a tree in $\mathfrak{T}$
· T_v — a fringe subtree of T rooted at node $v \in T$
· ξ — a non-negative integer-valued random variable with $\mathbb{E}\xi = 1$ and $0 < \sigma^2 :=$ $\mathrm{Var}\,(\xi) < \infty$
· p_i — $\mathbb{P}\,\{\xi = i\}$
· h — the span of ξ, i.e., $\gcd\{i \geqslant 1 : p_i > 0\}$
· $\mathcal{T}^{\mathbf{gw}}$ — an unconditional Galton-Watson tree with offspring distribution ξ
· $\mathcal{T}_n^{\mathbf{gw}}$ — $\mathcal{T}^{\mathbf{gw}}$ given that $|\mathcal{T}^{\mathbf{gw}}| = n$
· $\mathcal{T}_{n,v}^{\mathbf{gw}}$ — a fringe subtree of $\mathcal{T}_n^{\mathbf{gw}}$ rooted at node $v \in \mathcal{T}_n^{\mathbf{gw}}$
· $\mathcal{T}_{n,*}^{\mathbf{gw}}$ — a fringe subtree of $\mathcal{T}_n^{\mathbf{gw}}$ rooted at a uniform random node of $\mathcal{T}_n^{\mathbf{gw}}$
· $(T_n)_{n \geqslant 1}$ — a sequence of trees
· $\mathfrak{T}_n$ — the set of all trees of size n
· $\mathcal{S}$ — a set of trees
· $\mathfrak{T}_{\leqslant n}^+$ — the set $\{T \in \mathfrak{T} : |T| \leqslant n, \mathbb{P}\,\{\mathcal{T}^{\mathbf{gw}} = T\} > 0\}$
· $(\mathcal{A}_n)_{n \geqslant 1}$ — a sequence of sets of trees
· $N_{\mathcal{S}}(\mathcal{T}_n^{\mathbf{gw}})$ — the number of fringe subtrees of $\mathcal{T}_n^{\mathbf{gw}}$ that belong to $\mathcal{S}$
· $\pi(\mathcal{S})$ — $\mathbb{P}\,\{\mathcal{T}^{\mathbf{gw}} \in \mathcal{S}\}$
· $N_T^{\mathbf{nf}}(\mathcal{T}_n^{\mathbf{gw}})$ — the number of non-fringe subtrees of $\mathcal{T}_n^{\mathbf{gw}}$ of shape T
· $\pi^{\mathbf{nf}}(T)$ — $\mathbb{P}\,\{T \prec \mathcal{T}^{\mathbf{gw}}\}$, the probability that $\mathcal{T}^{\mathbf{gw}}$ has a non-fringe subtree T at its root

Remark 1.1. If $p_1 = 0$, then there exist positive integers n such that $\mathbb{P}\,\{|\mathcal{T}^{\mathbf{gw}}| = n\}$ = 0. For such n, $\mathcal{T}_n^{\mathbf{gw}}$ is not well-defined. But it is easy to show that $\mathbb{P}\,\{|\mathcal{T}^{\mathbf{gw}}| = n\}$ > 0 for all $n \geqslant n_0$ with $n - 1 \equiv 0 \pmod{h}$, where h is span of ξ and n_0 depends only on ξ (Janson, 2012, cor. 15.6). Therefore, in this paper, for all asymptotic results about $\mathcal{T}^{\mathbf{gw}}$ and $\mathcal{T}_n^{\mathbf{gw}}$, the limits are always taken along the subsequence with $n - 1 \equiv 0 \pmod{h}$.

Extending a result by Aldous (1991), Janson (2012, thm. 7.12) proved the following theorem:

Theorem 1.2. *Assume Condition A. The conditional distribution* $\mathcal{L}(\mathcal{T}_{n,*}^{\mathbf{gw}}|\mathcal{T}_n^{\mathbf{gw}})$ *converges in probability to* $\mathcal{L}(\mathcal{T}^{\mathbf{gw}})$. *In other words, for all* $T \in \mathfrak{T}$, *as* $n \to \infty$,

$$\frac{N_T(\mathcal{T}_n^{\mathbf{gw}})}{n} = \mathbb{P}\left\{\mathcal{T}_{n,*}^{\mathbf{gw}} = T|\mathcal{T}_n^{\mathbf{gw}}\right\} \xrightarrow{p} \mathbb{P}\left\{\mathcal{T}^{\mathbf{gw}} = T\right\}. \tag{1.1}$$

Later Janson (2016) strengthened the above result, proving the asymptotic normality of $N_T(\mathcal{T}_n^{\mathbf{gw}})$ by studying additive functionals on $\mathcal{T}_n^{\mathbf{gw}}$ (Janson, 2016).

A natural generalization of $N_T(\mathcal{T}_n^{\mathbf{gw}})$ is the fringe subtree counts $N_{T_n}(\mathcal{T}_n^{\mathbf{gw}})$ where $T_n \in \mathfrak{T}$ is a sequence of trees instead of a fixed tree T. Let $\mathrm{Po}(\lambda)$ denote a Poisson random variable with mean λ. We have:

Theorem 1.3. *Assume Condition A. Let* $\pi(T) := \mathbb{P}\{\mathcal{T}^{\mathbf{gw}} = T\}$ *and let* $k_n \to \infty, k_n = o(n)$. *Then*

$$\sup_{T:|T|=k_n} d_{TV}\left(N_T(\mathcal{T}_n^{\mathbf{gw}}), \mathrm{Po}(n\pi(T))\right) = O\left(p_{\mathbf{max}}^{k_n} k_n^{3/2}\right) = o(1), \tag{1.2}$$

where $p_{\mathbf{max}} := \max_{i \geq 0} p_i$ *and* $d_{TV}(\cdot, \cdot)$ *denotes the total variation distance. Therefore, letting* T_n *be a sequence of trees with* $|T_n| = k_n$, *we have as* $n \to \infty$:

(i) *If* $n\pi(T_n) \to 0$, *then* $N_{T_n}(\mathcal{T}_n^{\mathbf{gw}}) = 0$ *whp.*

(ii) *If* $n\pi(T_n) \to \mu \in (0, \infty)$, *then* $N_{T_n}(\mathcal{T}_n^{\mathbf{gw}}) \xrightarrow{d} \mathrm{Po}(\mu)$.

(iii) *If* $n\pi(T_n) \to \infty$, *then*

$$\frac{N_{T_n}(\mathcal{T}_n^{\mathbf{gw}}) - n\pi(T_n)}{\sqrt{n\pi(T_n)}} \xrightarrow{d} N(0, 1),$$

where $N(0, 1)$ *denotes the standard normal distribution, and* $\xrightarrow{d}$ *denotes convergence in distribution.*

Theorem 1.3 can be partially generalized as follows:

Theorem 1.4. *Assume Condition A. Let* $\mathfrak{T}_{k_n}$ *be the set of all trees of size* k_n, *where* $k_n \to \infty$ *and* $k_n = o(n)$. *For* $\mathcal{S} \subseteq \mathfrak{T}_{k_n}$, *let* $\pi(\mathcal{S}) := \mathbb{P}\{\mathcal{T}^{\mathbf{gw}} \in \mathcal{S}\}$ *and* $N_{\mathcal{S}}(\mathcal{T}_n^{\mathbf{gw}}) := \sum_{v \in \mathcal{T}_n^{\mathbf{gw}}} [\![\mathcal{T}_{n,v}^{\mathbf{gw}} \in \mathcal{S}]\!]$. *Therefore, letting* $(\mathcal{A}_n)_{n \geq 1}$ *be a sequence of sets of trees with* $\mathcal{A}_n \subseteq \mathfrak{T}_{k_n}$, *we have:*

(i) *If* $n\pi(\mathcal{A}_n) \to 0$, *then* $N_{\mathcal{A}_n}(\mathcal{T}_n^{\mathbf{gw}}) = 0$ *whp.*

(ii) *If* $n\pi(\mathcal{A}_n) \to \mu \in (0, \infty)$, *then* $N_{\mathcal{A}_n}(\mathcal{T}_n^{\mathbf{gw}}) \xrightarrow{d} \mathrm{Po}(\mu)$.

(iii) *If* $n\pi(\mathcal{A}_n) \to \infty$, *then*

$$\frac{N_{\mathcal{A}_n}(\mathcal{T}_n^{\mathbf{gw}}) - n\pi(\mathcal{A}_n)}{\sqrt{n\pi(\mathcal{A}_n)}} \xrightarrow{d} N(0, 1).$$

(iv) *If* $\pi(\mathcal{A}_n)/\pi(\mathfrak{T}_{k_n}) \to 0$, *then*

$$\lim_{n \to \infty} d_{TV}\left(N_{\mathcal{A}_n}(\mathcal{T}_n^{\mathbf{gw}}), \mathrm{Po}(n\pi(\mathcal{A}_n))\right) = 0.$$

Remark 1.5. $N_{\mathcal{S}}(\mathcal{T}_n^{\mathbf{gw}})$ can also be interpreted as the number of fringe subtrees with certain properties. For example, $N_{\mathfrak{T}_k}(\mathcal{T}_n^{\mathbf{gw}})$ is the number of fringe subtrees of size k. The above theorem together with Lemma 2.4 shows that, as long as $k_n = o(n^{3/2})$ then we have a central limit theorem for $N_{\mathfrak{T}_{k_n}}(\mathcal{T}_n^{\mathbf{gw}})$.

The proof of Theorem 1.3 is given in Section 3. It uses many ingredients from previous results on fringe subtrees, especially from Janson (2016). In particular, Lemma 6.2 of Janson (2016) makes the computation of the variance of $N_{\mathfrak{T}_k}(\mathcal{T}_n^{\mathbf{gw}})$ quite easy, which is crucial for the proof.

However, this approach cannot be adapted to prove the convergence of total variation distance for $N_{\mathcal{A}_n}(\mathcal{T}_n^{\mathbf{gw}})$ in (iv) of Theorem 1.4 without assuming that $\pi(\mathcal{A}_n)/\pi(\mathfrak{T}_{k_n}) \to 0$. In particular, it does not work for $N_{\mathfrak{T}_{k_n}}(\mathcal{T}_n^{\mathbf{gw}})$, i.e., the number of fringe subtrees of size k_n. Therefore, to show (i)–(iii) of Theorem 1.4, we instead compute the factorial moments of $N_{\mathcal{A}_n}(\mathcal{T}_n^{\mathbf{gw}})$. We sketch the proof of the theorem at the end of Section 3.

Binary search trees and *recursive trees* are also well-studied random tree models (see Drmota (2009)). Many authors have found results similar to Theorem 1.3 for these two types of trees, see, e.g., Feng et al. (2008); Fuchs (2008); Devroye (1991, 2002/03); Flajolet et al. (1997). For recent developments, see Holmgren and Janson (2015).

We say that a tree T is *possible* if $\mathbb{P}\{\mathcal{T}^{\mathbf{gw}} = T\} > 0$. As an application of Theorems 1.3 and 1.4, we ask the following question — when does $\mathcal{T}_n^{\mathbf{gw}}$ contain all possible trees within a family of trees (possibly depending on n). As shown in Subsection 4.1, this is essentially a variation of the coupon collector problem.

In Subsection 4.2 we answer the above question for the set of complete r-ary trees. Let $H_{n,r}$ be the maximal integer such that $\mathcal{T}_n^{\mathbf{gw}}$ contains all complete r-ary trees of height at most $H_{n,r}$ as fringe subtrees. Lemma 4.2 shows that $H_{n,r} - \log_r \log n$ converges in probability to an explicit constant.

Let $\mathfrak{T}_{\leqslant k}^+$ be the set of all possible trees of size at most k. Let $K_n = \max\{k : \mathfrak{T}_{\leqslant k}^+ \subseteq \cup_{v \in \mathcal{T}_n^{\mathbf{gw}}} \mathcal{T}_{n,v}^{\mathbf{gw}}\}$, i.e., K_n is the maximal k such that every tree in $\mathfrak{T}_{\leqslant k}^+$ appears in $\mathcal{T}_n^{\mathbf{gw}}$ as fringe subtrees. In Subsection 4.3, we show that, roughly speaking, if the tail of the offspring distribution does not drop off too quickly, $K_n/\log n$ converges in probability to a positive constant. Otherwise, we have $K_n/\log n \xrightarrow{p} 0$. For example, for a random Cayley tree, we have $K_n \log\log(n)/\log(n) \xrightarrow{p} 1$. For many well-known Galton-Watson trees, we also give the second order asymptotic term of K_n.

Non-fringe subtrees are more complicated to analyze. However since on average fringe subtrees in $\mathcal{T}_n^{\mathbf{gw}}$ behave like unconditional Galton-Watson trees when n is large, the number of non-fringe subtrees of shape T should be more or less $n\mathbb{P}\{T \prec \mathcal{T}^{\mathbf{gw}}\}$. The following theorem is a precise version of this intuition.

Theorem 1.6. *Assume Condition A. Let $\pi^{\mathbf{nf}}(T) := \mathbb{P}\{T \prec \mathcal{T}^{\mathbf{gw}}\}$. Let $N_T^{\mathbf{nf}}(\mathcal{T}_n^{\mathbf{gw}}) := \sum_{v \in \mathcal{T}_n^{\mathbf{gw}}} [\![T \prec \mathcal{T}_{n,v}^{\mathbf{gw}}]\!]$. Let T_n be a sequence of trees with $|T_n| = k_n$ where $k_n \to \infty$ and $k_n = o(n)$. We have*

(i) *If $n\pi^{\mathbf{nf}}(T_n) \to 0$, then $N_{T_n}^{\mathbf{nf}}(\mathcal{T}_n^{\mathbf{gw}}) \xrightarrow{p} 0$.*

(ii) *If $n\pi^{\mathbf{nf}}(T_n) \to \infty$, then $N_{T_n}^{\mathbf{nf}}(\mathcal{T}_n^{\mathbf{gw}})/(n\pi^{\mathbf{nf}}(T_n)) \xrightarrow{p} 1$.*

Chyzak et al. (2008) studied non-fringe subtrees for various random trees, including *simply generated trees*. They proved that if for all n we have $T_n = T$ where T is fixed, then $N_{T_n}^{\mathbf{nf}}(\mathcal{T}_n^{\mathbf{gw}})$ has a central limit theorem. However, Theorem 1.6 cannot be simply derived from their result as our T_n depends upon n.

Remark 1.7. It is tempting to try to prove that if $n\pi^{\mathbf{nf}}(T_n) \to \mu \in (0, \infty)$, then $N_{T_n}^{\mathbf{nf}}(\mathcal{T}_n^{\mathbf{gw}}) \xrightarrow{d} \mathrm{Po}(\mu)$. This is true for fringe subtrees. Unfortunately, it is not true in general for non-fringe subtrees. See Lemma 5.5 in Section 5.2.

In Section 5, we prove Theorem 1.6 and apply it to study the maximal complete r-ary non-fringe subtree in $\mathcal{T}_n^{\mathbf{gw}}$. The paper ends with some open questions in Section 6.

2. Notations and Preliminaries

2.1. *Conditional Galton-Watson trees.* The *preorder* of nodes in a tree T is the order in which they are visited through the following Depth-First-Search procedure:

(1) Let $\mathcal{S}$ be an empty stack.
(2) Put the root of T at the top of $\mathcal{S}$.
(3) Remove the node (v) at the top of $\mathcal{S}$.
(4) Put the children of v at the top of $\mathcal{S}$ in order of appearance.
(5) If $\mathcal{S}$ is empty, terminate. Otherwise go to step 3.

Let $v_1, \ldots, v_k$ be the nodes of T in preorder. Let d_i be the *degree* (the number of children) of v_i. We call $(d_1, d_2, \ldots, d_k)$ the *preorder degree sequence* of T. Let $\mathbb{N} := \{1, 2, \ldots\}$ and let $\mathbb{N}_0 := \{0\} \cup \mathbb{N}$. It is well-known that (see Janson (2012, lem. 15.2)):

Lemma 2.1. *A sequence $(d_1, d_2, \ldots, d_k) \in \mathbb{N}_0^k$ is the preorder degree sequence of some tree if and only if it satisfies*

$$\begin{cases} \sum_{i=1}^{j} d_i \geqslant j, & (1 \leqslant j \leqslant k-1) \\ \sum_{i=1}^{k} d_i = k-1. \end{cases} \tag{2.1}$$

Figure 2.1 gives a demonstration of Lemma 2.1.

The degree sequence
$(d_1, \ldots, d_7) = (2, 1, 0, 3, 0, 0, 0)$

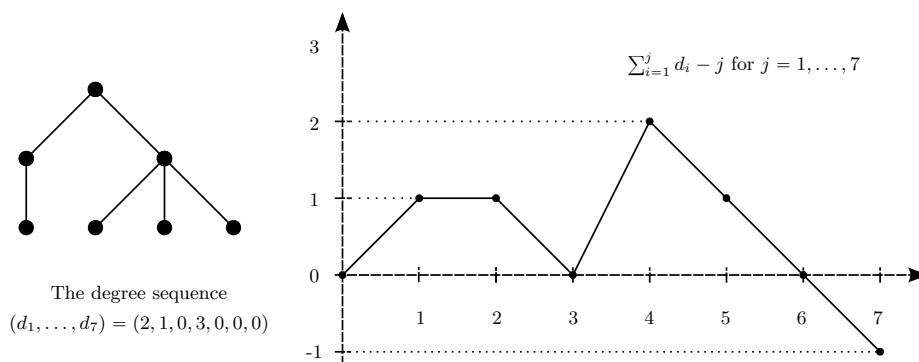$\sum_{i=1}^{j} d_i - j$ for $j = 1, \ldots, 7$

FIGURE 2.1. Example of preorder tree degree sequence.

Let $\mathcal{D}_k \subseteq \mathbb{N}_0^k$ be the set of all preorder degree sequences of length k. Observe:

Corollary 2.2. *If $(d_1, d_2, \ldots, d_k) \in \mathcal{D}_k$, then it is impossible that there exists $1 \leqslant k' < k$ such that $(d_1, d_2, \ldots, d_{k'}) \in \mathcal{D}_{k'}$.*

Let $\boldsymbol{\xi^n} := (\xi_1^{\mathbf{n}}, \ldots, \xi_n^{\mathbf{n}})$ be the preorder degree sequence of $\mathcal{T}_n^{\mathbf{gw}}$. Let $\widetilde{\boldsymbol{\xi}}^{\mathbf{n}} := (\widetilde{\xi}_1^{\mathbf{n}}, \ldots, \widetilde{\xi}_n^{\mathbf{n}})$ be a uniform random cyclic rotation of $\boldsymbol{\xi^n}$. Let $\xi_1, \xi_2, \ldots$ be i.i.d. copies of ξ. Let $S_n := \sum_{i=1}^{n} \xi_i$. The next lemma is a well-known connection between $\widetilde{\boldsymbol{\xi}}^{\mathbf{n}}$ and $\xi_1, \ldots, \xi_n$ (see, e.g., Otter, 1948, Kolchin, 1986, Dwass, 1969 and Pitman, 1998). For a complete proof, see Janson (2012, cor. 15.4).

Lemma 2.3. *Assume that* $\mathbb{P}\left\{|\mathcal{T}^{\mathbf{gw}}| = n\right\} > 0$. *We have*

$$(\widetilde{\xi}_1^{\mathbf{n}}, \ldots, \widetilde{\xi}_n^{\mathbf{n}}) \overset{\mathcal{L}}{=} (\xi_1, \ldots, \xi_n \mid S_n = n - 1),$$

where $\overset{\mathcal{L}}{=}$ *denotes "identically distributed" and the right-hand-side denotes* $(\xi_1, \ldots, \xi_n)$ *restricted to the event that* $S_n = n - 1$.

Let $p_i := \mathbb{P}\left\{\xi = i\right\}$. Let h be the *span* of ξ, i.e., $h := \gcd\{i \geqslant 1 : p_i > 0\}$. We recall the following result (see Janson, 2016, 4.3 or Kolchin, 1986):

Lemma 2.4. *Assume Condition A. We have*

$$\mathbb{P}\left\{|\mathcal{T}^{\mathbf{gw}}| = n\right\} \sim \frac{h}{\sqrt{2\pi\sigma^2}} n^{-3/2},$$

as $n \to \infty$ *with* $n - 1 \equiv 0 \pmod{h}$.

The following lemma is a special case of Janson (2016, lem. 5.1). We nonetheless give a short proof for later reference in the paper.

Lemma 2.5. *Assume* $\mathbb{P}\left\{|\mathcal{T}^{\mathbf{gw}}| = n\right\} > 0$. *Let* $T \in \mathfrak{T}_k$ *with* $1 \leqslant k \leqslant n$.

(i) *Let* $N_T(\mathcal{T}_n^{\mathbf{gw}}) := \sum_{v \in \mathcal{T}_n^{\mathbf{gw}}} \llbracket \mathcal{T}_{n,v}^{\mathbf{gw}} = T \rrbracket$. *Then*

$$\frac{\mathbb{E}\left[N_T(\mathcal{T}_n^{\mathbf{gw}})\right]}{n} = \pi(T)\frac{\mathbb{P}\left\{S_{n-k} = n - k\right\}}{\mathbb{P}\left\{S_n = n - 1\right\}}.$$

(ii) *Let* $N_{\mathfrak{T}_k}(\mathcal{T}_n^{\mathbf{gw}}) = \sum_{v \in \mathcal{T}_n^{\mathbf{gw}}} \llbracket |\mathcal{T}_{n,v}^{\mathbf{gw}}| = k \rrbracket$. *Then*

$$\frac{\mathbb{E}\left[N_{\mathfrak{T}_k}(\mathcal{T}_n^{\mathbf{gw}})\right]}{n} = \pi(\mathfrak{T}_k)\frac{\mathbb{P}\left\{S_{n-k} = n - k\right\}}{\mathbb{P}\left\{S_n = n - 1\right\}}.$$

Proof: Let $(d_1, \ldots, d_k)$ be the preorder degree sequence of T. Recall that $(\xi_1^{\mathbf{n}}, \ldots, \xi_n^{\mathbf{n}})$ is the preorder degree sequence of $\mathcal{T}_n^{\mathbf{gw}}$. Let

$$I_i = \llbracket \xi_i^{\mathbf{n}} = d_1, \xi_{i+1}^{\mathbf{n}} = d_2, \ldots, \xi_{i+k-1}^{\mathbf{n}} = d_k \rrbracket,$$

where the indices are taken modulo n.

Note that if $n - k + 1 < i \leqslant n$, then it is impossible that $I_i = 1$, because the length of the preorder degree sequence of the fringe subtree $\mathcal{T}_{n,v_i}^{\mathbf{gw}}$ must be strictly less than k. Therefore, if $I_i > 1$ and $n - k + 1 < i \leqslant n$, then there exists a $k' < k$ such that $(d_1, d_2, \ldots, d'_k)$ is also a preorder degree sequence, which is impossible by Corollary 2.2.

Therefore for all $1 \leqslant i \leqslant n$, $I_i = \llbracket T_{v_i} = T \rrbracket$ and $N_T(\mathcal{T}_n^{\mathbf{gw}}) = \sum_{i=1}^n I_i$. Recalling that $(\widetilde{\xi}_1^{\mathbf{n}}, \ldots, \widetilde{\xi}_n^{\mathbf{n}})$ is a uniform random rotation of $(\xi_1^{\mathbf{n}}, \ldots, \xi_n^{\mathbf{n}})$ and using Lemma 2.3, we have

$$\mathbb{E}\left[N_T(\mathcal{T}_n^{\mathbf{gw}})\right] = \mathbb{E}\left[\sum_{i=1}^n I_i\right] = \sum_{i=1}^n \mathbb{P}\left\{\xi_i^{\mathbf{n}} = d_1, \xi_{i+1}^{\mathbf{n}} = d_2, \ldots, \xi_{i+k-1}^{\mathbf{n}} = d_k\right\}$$

$$= \sum_{i=1}^n \mathbb{P}\left\{\xi_i = d_1, \xi_{i+1} = d_2, \ldots, \xi_{i+k-1} = d_k \mid S_n = n - 1\right\}$$

$$= n\mathbb{P}\left\{\xi_1 = d_1, \xi_2 = d_2, \ldots, \xi_k = d_k \mid S_n = n - 1\right\}$$

$$= n\frac{\mathbb{P}\left\{[\xi_1 = d_1, \xi_2 = d_2, \ldots, \xi_k = d_k] \cap [S_n = n - 1]\right\}}{\mathbb{P}\left\{S_n = n - 1\right\}}.$$

Since $(d_1, \ldots d_k)$ is a preorder degree sequence, $\sum_{i=1}^{k} d_i = k - 1$ by Lemma 2.1. Therefore, using the fact that $\xi_1, \ldots, \xi_n$ are independent, the last expression equals

$$n \frac{\mathbb{P}\left\{[\xi_1 = d_1, \ldots, \xi_k = d_k] \cap [S_{n-k} = n - k]\right\}}{\mathbb{P}\left\{S_n = n - 1\right\}} = n\mathbb{P}\left\{\mathcal{T}^{\mathbf{gw}} = T\right\} \frac{\mathbb{P}\left\{S_{n-k} = n - k\right\}}{\mathbb{P}\left\{S_n = n - 1\right\}}.$$

Thus part (i) is proved. Part (ii) follows by summing the equality in (i) over all $T \in \mathfrak{T}_k$. □

The following approximations are useful for estimating the expectation and the variance of the number of fringe subtrees.

Lemma 2.6 (Lemma 5.2 and 6.2 of Janson, 2016). *Assume Condition A and let ξ have span 1. We have:*
(i) Uniformly for all k with $1 \leqslant k \leqslant n/2$,

$$\frac{\mathbb{P}\left\{S_{n-k} = n - k\right\}}{\mathbb{P}\left\{S_n = n - 1\right\}} = 1 + O\left(\frac{k}{n}\right) + o\left(n^{-1/2}\right).$$

(ii) Uniformly for all k with $1 \leqslant k \leqslant n/4$,

$$\frac{\mathbb{P}\left\{S_{n-2k} = n - 2k + 1\right\}}{\mathbb{P}\left\{S_n = n - 1\right\}} - \left(\frac{\mathbb{P}\left\{S_{n-k} = n - k\right\}}{\mathbb{P}\left\{S_n = n - 1\right\}}\right)^2$$
$$= -\frac{1}{\sigma^2 n} + o\left(\frac{1}{n}\right) + O\left(\frac{k}{n^{3/2}} + \frac{k^2}{n^2}\right). \qquad (2.2)$$

Remark 2.7. As shown in the proof of Lemma 2.5, $N_T(\mathcal{T}_n^{\mathbf{gw}})$ is equivalent to the number of patterns $d_1, \ldots, d_{|T|}$ in the cycle $\widetilde{\xi}_1^{\mathbf{n}}, \ldots, \widetilde{\xi}_n^{\mathbf{n}}$. Thus if h (the span of ξ) is greater than one, we can divide $d_1, \ldots, d_{|T|}$ and $\widetilde{\xi}_1^{\mathbf{n}}, \ldots, \widetilde{\xi}_n^{\mathbf{n}}$ by h without changing the value of $N_T(\mathcal{T}_n^{\mathbf{gw}})$. Therefore, when studying subtree counts, we can always assume that $h = 1$.

2.2. *Poisson Approximation.* Let $\mathrm{Bi}(n, p)$ denote binomial (n, p) distribution. It is well known that if $X_n \overset{\mathcal{L}}{=} \mathrm{Bi}(n, \lambda/n)$, then X_n converges in distribution to $\mathrm{Po}(\lambda)$. This follows from the following stronger result (see Barbour et al., 1992, pp. 8 for a proof):

Lemma 2.8. *If $X \overset{\mathcal{L}}{=} \mathrm{Bi}(n, p)$, then*

$$d_{\mathrm{TV}}\left(X, \mathrm{Po}(\mathbb{E}X)\right) \leqslant p.$$

The following Lemma is a special case of Roos (2003, thm. 1), which applies to mixed Poisson distributions. Barbour et al. (1992, thm. 1.C) proved a similar result using Stein's method. We include our proof for its simplicity.

Lemma 2.9. *If $X \overset{\mathcal{L}}{=} \mathrm{Po}(\mu)$ and $Y \overset{\mathcal{L}}{=} \mathrm{Po}(\nu)$, then*

$$d_{\mathrm{TV}}\left(X, Y\right) \leqslant \left|\sqrt{\mu} - \sqrt{\nu}\right| = \frac{|\mu - \nu|}{\sqrt{\mu} + \sqrt{\nu}}.$$

Proof: Let $x_i = \mathbb{P}\{X = i\}$ and $y_i = \mathbb{P}\{Y = i\}$. We have

$$d_{\mathrm{TV}}(X, Y) = \frac{1}{2}\sum_{i=1}^{\infty}|x_i - y_i| = \frac{1}{2}\sum_{i=1}^{\infty}|\sqrt{x_i} - \sqrt{y_i}|(\sqrt{x_i} + \sqrt{y_i})$$

$$\leqslant \frac{1}{2}\left(\sum_{i=1}^{\infty}|\sqrt{x_i} - \sqrt{y_i}|^2 \sum_{j=1}^{\infty}(\sqrt{x_j} + \sqrt{y_j})^2\right)^{1/2}$$

$$= \frac{1}{2}\left(\left(2 - \sum_{i=1}^{\infty}2\sqrt{x_i y_i}\right)\left(2 + \sum_{j=1}^{\infty}2\sqrt{x_j y_j}\right)\right)^{1/2}$$

$$= \left(1 - \left(\sum_{i=1}^{\infty}\sqrt{x_i y_i}\right)^2\right)^{1/2},$$

where the second step uses the Cauchy-Schwartz inequality. An easy calculation shows that

$$\sum_{i=1}^{\infty}\sqrt{x_i y_i} = \sum_{i=1}^{\infty}e^{-\frac{\mu+\nu}{2}}\frac{(\mu\nu)^{i/2}}{i!} = \exp\left(\sqrt{\mu\nu} - \frac{\mu+\nu}{2}\right) = \exp\left(-\frac{(\sqrt{\mu} - \sqrt{\nu})^2}{2}\right).$$

Thus we have

$$d_{\mathrm{TV}}(X, Y) \leqslant \sqrt{1 - \exp\left(-(\sqrt{\mu} - \sqrt{\nu})^2\right)}$$

$$\leqslant \sqrt{(\sqrt{\mu} - \sqrt{\nu})^2} \qquad (\text{by } 1 - e^{-x} \leqslant x)$$

$$= \left|\sqrt{\mu} - \sqrt{\nu}\right|. \qquad\qquad \square$$

Combining the two, we have:

Lemma 2.10. *Let X and M be non-negative integer-valued random variables. If conditioned on the event $M = m$, X is binomial (m, p), then*

$$d_{\mathrm{TV}}(X, \mathrm{Po}(\mathbb{E}X)) \leqslant p + \sqrt{p\frac{\mathrm{Var}(M)}{\mathbb{E}M}}.$$

Proof: Let $Z \overset{\mathcal{L}}{=} \mathrm{Po}(\mathbb{E}X)$. We have

$$d_{\mathrm{TV}}(X, Z) = \frac{1}{2}\sum_{i\geqslant 0}|\mathbb{P}\{X = i\} - \mathbb{P}\{Z = i\}|$$

$$\leqslant \frac{1}{2}\sum_{i\geqslant 0}\sum_{m\geqslant 0}\mathbb{P}\{M = m\}|\mathbb{P}\{X = i \mid M = m\} - \mathbb{P}\{Z = i\}|$$

$$= \sum_{m\geqslant 0}\mathbb{P}\{M = m\}\,d_{\mathrm{TV}}(\mathrm{Bi}(m, p), Z)$$

$$\leqslant \sum_{m\geqslant 0}\mathbb{P}\{M = m\}\,d_{\mathrm{TV}}(\mathrm{Bi}(m, p), \mathrm{Po}(mp))$$

$$+ \sum_{m\geqslant 0}\mathbb{P}\{M = m\}\,d_{\mathrm{TV}}(\mathrm{Po}(mp), \mathrm{Po}(\mathbb{E}X)).$$

By Lemma 2.8, the first sum is at most p. Using Lemma 2.9 and the fact that $\mathbb{E}X = p\mathbb{E}M$, the second sum is at most

$$\sum_{m \geqslant 0} \mathbb{P}\{M = m\} \frac{|\mathbb{E}X - mp|}{\sqrt{\mathbb{E}X}} = \sqrt{p}\sum_{m \geqslant 0} \mathbb{P}\{M = m\}\frac{|\mathbb{E}M - m|}{\sqrt{\mathbb{E}M}}$$

$$= \sqrt{p}\frac{\mathbb{E}\left[|\mathbb{E}M - M|\right]}{\sqrt{\mathbb{E}M}}$$

$$\leqslant \sqrt{p}\sqrt{\frac{\mathrm{Var}\,(M)}{\mathbb{E}M}}. \qquad \square$$

3. Sequences of fringe subtrees

In this section we prove Theorem 1.3 and sketch the proof of Theorem 1.4. Recall that $\pi(T) := \mathbb{P}\{\mathcal{T}^{\mathbf{gw}} = T\}$, and $N_T(n) := [\![\sum_{v \in \mathcal{T}_n^{\mathbf{gw}}} \mathcal{T}_{n,v}^{\mathbf{gw}} = T]\!]$. Let $k = k_n$. Theorem 1.3 states that

$$\sup_{T \in \mathfrak{T}_k} d_{\mathrm{TV}}\left(N_T(\mathcal{T}_n^{\mathbf{gw}}), \mathrm{Po}(n\pi(T))\right) = o(1),$$

whenever $k = o(n)$ and $k \to \infty$. If $\pi(T) = 0$, then $d_{\mathrm{TV}}\left(N_T(\mathcal{T}_n^{\mathbf{gw}}), \mathrm{Po}(n\pi(T))\right) = 0$ deterministically. Thus we can assume that $\mathbb{P}\{\mathcal{T}^{\mathbf{gw}} \in \mathfrak{T}_k\} = \mathbb{P}\{|\mathcal{T}^{\mathbf{gw}}| = k\} > 0$ for all n, and that the above supremum is taken over all $T \in \mathfrak{T}_k$ with $\pi(T) > 0$.

Recall that $N_{\mathfrak{T}_k}(\mathcal{T}_n^{\mathbf{gw}}) := [\![\sum_{v \in \mathcal{T}_n^{\mathbf{gw}}} T_v \in \mathfrak{T}_k]\!]$, i.e., $N_{\mathfrak{T}_k}(\mathcal{T}_n^{\mathbf{gw}})$ is the number of fringe subtrees of size k in $\mathcal{T}_n^{\mathbf{gw}}$. Also recall that $\pi(\mathcal{S}) := \mathbb{P}\{\mathcal{T}^{\mathbf{gw}} \in \mathcal{S}\}$. We first compute the expectation and variance of $N_{\mathfrak{T}_k}(\mathcal{T}_n^{\mathbf{gw}})$. Then Lemma 2.10 can be applied to $N_{\mathfrak{T}_k}(\mathcal{T}_n^{\mathbf{gw}})$ and $N_{\mathcal{S}}(\mathcal{T}_n^{\mathbf{gw}})$ for $\mathcal{S} \subseteq \mathfrak{T}_k$ to show the following lemma, from which Theorem 1.3 follows easily:

Lemma 3.1. *Assume that $k = k_n = o(n)$ and $k \to \infty$. We have as $n \to \infty$,*

$$\sup_{\mathcal{S} \subseteq \mathfrak{T}_k} \frac{d_{\mathrm{TV}}\left(N_{\mathcal{S}}(\mathcal{T}_n^{\mathbf{gw}}), \mathrm{Po}(n\pi(\mathcal{S}))\right)}{\pi(\mathcal{S})/\pi(\mathfrak{T}_k) + \sqrt{\pi(\mathcal{S})/\pi(\mathfrak{T}_k)}} \leqslant 1 + o(k^{-3/2}) + O\left(\frac{k^{1/4}}{\sqrt{n}}\right).$$

Lemma 3.2. *Let $k = k_n = o(n)$. We have*

$$\sup_{T \in \mathfrak{T}_k}\left|\frac{\mathbb{E}N_T(\mathcal{T}_n^{\mathbf{gw}})}{n\pi(T)} - 1\right| = O\left(\frac{k}{n}\right) + o\left(n^{-1/2}\right),$$

and

$$\sup_{\mathcal{S} \subseteq \mathfrak{T}_k}\left|\frac{\mathbb{E}N_{\mathcal{S}}(\mathcal{T}_n^{\mathbf{gw}})}{n\pi(\mathcal{S})} - 1\right| = O\left(\frac{k}{n}\right) + o\left(n^{-1/2}\right).$$

Proof: Since $k = o(n)$, we have $k < n/2$ for n large. Thus by Lemma 2.5 and 2.6, uniformly for all $T \in \mathfrak{T}_k$,

$$\left|\frac{\mathbb{E}N_T(\mathcal{T}_n^{\mathbf{gw}})}{n} - \pi(T)\right| = \pi(T)\left|\frac{\mathbb{P}\{S_{n-k} = n - k\}}{\mathbb{P}\{S_n = n - 1\}} - 1\right|$$

$$= \pi(T)\left(O\left(\frac{k}{n}\right) + o(n^{-1/2})\right).$$

Summing over all trees T with $T \in \mathcal{S}$ gives the second part of the lemma. $\qquad \square$

Lemma 3.3. *Assume that $k = k_n = o(n)$ and $k \to \infty$. We have*

$$\frac{\mathrm{Var}\left(N_{\mathfrak{T}_k}(\mathcal{T}_n^{\mathbf{gw}})\right)}{\mathbb{E}N_{\mathfrak{T}_k}(\mathcal{T}_n^{\mathbf{gw}})} \leqslant 1 + o(k^{-3/2}) + O\left(\sqrt{\frac{k}{n}}\right).$$

Proof: Recall that $\widetilde{\boldsymbol{\xi}}^{\mathbf{n}} := (\widetilde{\xi}_1^{\mathbf{n}}, \ldots, \widetilde{\xi}_n^{\mathbf{n}})$ is a uniform random rotation of $\boldsymbol{\xi}^{\mathbf{n}}$, the preorder degree sequence of $\mathcal{T}_n^{\mathbf{gw}}$. Let

$$\widetilde{J}_i := [\![(\widetilde{\xi}_i^{\mathbf{n}}, \widetilde{\xi}_{i+1}^{\mathbf{n}}, \ldots \widetilde{\xi}_{i+k-1}^{\mathbf{n}}) \in \mathcal{D}_k]\!].$$

Then with an argument similar to Lemma 2.5, we have $N_{\mathfrak{T}_k}(\mathcal{T}_n^{\mathbf{gw}}) = \sum_{i=1}^n \widetilde{J}_i$.

Using the fact that $\widetilde{J}_1, \ldots, \widetilde{J}_n$ are identically distributed, Lemma 2.5 and 2.6, we have

$$\mathbb{E}\widetilde{J}_1 = \frac{1}{n}\mathbb{E}N_{\mathfrak{T}_k}(\mathcal{T}_n^{\mathbf{gw}}) = \mathbb{P}\left\{|\mathcal{T}^{\mathbf{gw}}| = k\right\}\frac{\mathbb{P}\left\{S_{n-k} = n - k\right\}}{\mathbb{P}\left\{S_n = n - 1\right\}}.$$

Similar to the proof of Lemma 2.5, we have

$$\begin{aligned}
\mathbb{E}\widetilde{J}_1\widetilde{J}_{k+1} &= \mathbb{P}\left\{(\widetilde{\xi}_1^{\mathbf{n}}, \ldots, \widetilde{\xi}_k^{\mathbf{n}}) \in \mathcal{D}_k, (\widetilde{\xi}_{k+1}^{\mathbf{n}}, \ldots, \widetilde{\xi}_{2k}^{\mathbf{n}}) \in \mathcal{D}_k\right\} \\
&= \mathbb{P}\left\{(\xi_1, \ldots, \xi_k) \in \mathcal{D}_k, (\xi_{k+1}, \ldots, \xi_{2k}) \in \mathcal{D}_k \mid S_n = n - 1\right\} \\
&= \pi(\mathfrak{T}_k)^2\frac{\mathbb{P}\left\{S_{n-2k} = n - 2k + 1\right\}}{\mathbb{P}\left\{S_n = n - 1\right\}},
\end{aligned}$$

where $\xi_1, \xi_2, \ldots$ are i.i.d. copies of ξ and $S_m := \sum_{i=1}^m \xi_i$.

Consider two indices $i \neq j$. If $|i - j| < k$ or $|i + n - j| < k$, then $\mathbb{E}\widetilde{J}_i\widetilde{J}_j = 0$. This is because two fringe subtrees of size k cannot overlap. So for such i and j we have $\mathrm{Cov}\left(\widetilde{J}_i, \widetilde{J}_j\right) = \mathbb{E}\left[\widetilde{J}_i\widetilde{J}_j\right] - \mathbb{E}\left[\widetilde{J}_i\right]\mathbb{E}\left[\widetilde{J}_j\right] \leqslant 0$.

On the other hand, if $|i - j| > k$ and $|i + n - j| > k$, i.e., $(\widetilde{\xi}_i^{\mathbf{n}}, \ldots, \widetilde{\xi}_{i+k-1}^{\mathbf{n}})$ and $(\widetilde{\xi}_j^{\mathbf{n}}, \ldots, \widetilde{\xi}_{j+k-1}^{\mathbf{n}})$ do not overlap, then $\mathrm{Cov}\left(\widetilde{J}_i\widetilde{J}_j\right) = \mathrm{Cov}\left(\widetilde{J}_1\widetilde{J}_{k+1}\right)$ since $\widetilde{\boldsymbol{\xi}}^{\mathbf{n}}$ is permutation invariant. By Lemma 2.4, we have $\pi(\mathfrak{T}_k) = \Theta(k^{-3/2})$. Therefore

$$\begin{aligned}
\mathrm{Cov}\left(\widetilde{J}_1, \widetilde{J}_{k+1}\right) &= \mathbb{E}\left[\widetilde{J}_1\widetilde{J}_{k+1}\right] - \mathbb{E}\left[\widetilde{J}_1\right]\mathbb{E}\left[\widetilde{J}_{k+1}\right] \\
&= \frac{\mathbb{P}\left\{S_{n-2k} = n - 2k + 1\right\}}{\mathbb{P}\left\{S_n = n - 1\right\}} - \left(\frac{\mathbb{P}\left\{S_{n-k} = n - k\right\}}{\mathbb{P}\left\{S_n = n - 1\right\}}\right)^2 \\
&= \pi(\mathfrak{T}_k)^2\left[-\frac{1}{\sigma^2 n} + o\left(\frac{1}{n}\right) + O\left(\frac{k}{n^{3/2}} + \frac{k^2}{n^2}\right)\right] \qquad \text{(Lemma 2.6)} \\
&\leqslant \frac{\pi(\mathfrak{T}_k)}{n}\left[o(k^{-3/2}) + O\left(\frac{k^{-3/2}k}{\sqrt{n}} + \frac{k^{-3/2}k^2}{n}\right)\right] \\
&= \frac{\mathbb{E}\widetilde{J}_1}{n}\left(o\left(k^{-3/2}\right) + O\left(\sqrt{\frac{k}{n}}\right)\right).
\end{aligned}$$

Therefore,

$$\begin{aligned}
\mathrm{Var}\left(N_{\mathfrak{T}_k}(\mathcal{T}_n^{\mathbf{gw}})\right) &= \sum_{1 \leqslant i \neq j \leqslant n} \mathrm{Cov}\left(\tilde{J}_i, \tilde{J}_j\right) + \sum_{i=1}^n \mathrm{Var}\left(\tilde{J}_i\right) \\
&\leqslant n^2 \mathrm{Cov}\left(\tilde{J}_1, \tilde{J}_{k+1}\right) + n\mathbb{E}\left[\tilde{J}_1\right]\left(1 - \mathbb{E}\left[\tilde{J}_1\right]\right) \\
&\leqslant n^2 \frac{\mathbb{E}\tilde{J}_1}{n}\left(o\left(k^{-3/2}\right) + O\left(\sqrt{k/n}\right)\right) + n\mathbb{E}\left[\tilde{J}_1\right] \\
&= \left(1 + o\left(k^{-3/2}\right) + O\left(\sqrt{k/n}\right)\right)\mathbb{E}\left[N_{\mathfrak{T}_k}(\mathcal{T}_n^{\mathbf{gw}})\right]. \qquad \square
\end{aligned}$$

The following observation allows us to apply Lemma 2.10 to finish the proof.

Lemma 3.4. *Let event $E_m = [N_{\mathfrak{T}_k}(\mathcal{T}_n^{\mathbf{gw}}) = m]$. Conditional on E_m, the m fringe subtrees of size k in $\mathcal{T}_n^{\mathbf{gw}}$ has the distribution of m i.i.d. copies of $\mathcal{T}_k^{\mathbf{gw}}$. Thus for $\mathcal{S} \subseteq \mathfrak{T}_k$, conditional on E_m, $N_{\mathcal{S}}(\mathcal{T}_n^{\mathbf{gw}})$ is binomial $(m, \pi(\mathcal{S})/\pi(\mathfrak{T}_k))$.*

Proof: Conditional on E_1, the probability that $\mathcal{T}_n^{\mathbf{gw}}$ has $T \in \mathfrak{T}_k$ as the only fringe subtree of size k must be proportional to $\pi(T)$. And since this fringe subtree can only have size k, this probability in fact must be $\pi(T)/\pi(\mathfrak{T}_k)$. In other words, this fringe subtree has the distribution of $\mathcal{T}_k^{\mathbf{gw}}$. It is not difficult to extend this argument to E_m with $m > 1$. $\qquad \square$

Proof of Lemma 3.1: Let $\mathcal{S} \subseteq \mathfrak{T}_k$. Let $X = N_{\mathcal{S}}(\mathcal{T}_n^{\mathbf{gw}})$, $M = N_{\mathfrak{T}_k}(\mathcal{T}_n^{\mathbf{gw}})$ and $p = \pi(\mathcal{S})/\pi(\mathfrak{T}_k)$. By Lemmas 2.10, 3.3, and 3.4, we have

$$d_{\mathrm{TV}}\left(X, \mathrm{Po}(\mathbb{E}X)\right) \leqslant p + \sqrt{p}\sqrt{\frac{\mathrm{Var}\left(M\right)}{\mathbb{E}M}} \leqslant (p + \sqrt{p})\left(1 + o(k^{-3/2}) + O\left(\sqrt{\frac{k}{n}}\right)\right).$$

By Lemma 2.9, we have

$$\begin{aligned}
d_{\mathrm{TV}}\left(\mathrm{Po}(\mathbb{E}X), \mathrm{Po}(n\pi(\mathcal{S}))\right) &\leqslant \frac{|n\pi(\mathcal{S}) - \mathbb{E}X|}{\sqrt{n\pi(\mathcal{S})}} \\
&= \sqrt{n\pi(\mathcal{S})}\left(O\left(\frac{k}{n}\right) + o\left(n^{-1/2}\right)\right) \qquad \text{(By Lemma 3.2)} \\
&\leqslant \sqrt{n\pi(\mathfrak{T}_k)}\left(O\left(\frac{k}{n}\right) + o\left(n^{-1/2}\right)\right) \\
&= O\left(\frac{k^{1/4}}{\sqrt{n}}\right) + o(k^{-3/4}) \qquad\qquad \text{(By Lemma 2.4).}
\end{aligned}$$

The lemma follows from triangle inequality. $\qquad \square$

Proof of Theorem 1.3: Let $k = k_n$. For $T \in \mathfrak{T}_k$, we have $\pi(T) \leqslant p_{\mathbf{max}}^k$. Therefore, by Lemma 2.4, $\pi(T)/\pi(\mathfrak{T}_k) \leqslant p_{\mathbf{max}}^k/\Theta(k^{-3/2}) = o(1)$. It follows from Lemma 3.1 that

$$d_{\mathrm{TV}}\left(N_T(\mathcal{T}_n^{\mathbf{gw}}), \mathrm{Po}(n\pi(T))\right) \leqslant (1 + o(1))\left(\frac{\pi(T)}{\pi(\mathfrak{T}_k)} + \sqrt{\frac{\pi(T)}{\pi(\mathfrak{T}_k)}}\right) = o(1).$$

Statements (i)–(iii) follows immediately. $\qquad \square$

3.1. *The sketch of the proof of Theorem 1.4.* Part (iv) of Theorem 2.4 follows directly from Lemma 2.9. However, to show (i)-(iii) of Theorem 1.4, we instead need the following lemma (whose proof is quite similar to that of Lemma 3.3):

Lemma 3.5. *Let $r = r_n = o(\sqrt{n})$ and $k = k_n = o(n/r_n^2)$ with $k \to \infty$. We have*

$$\sup_{\mathcal{S} \subseteq \mathfrak{T}_k} \sup_{s \leqslant r} \left| \frac{\mathbb{E}(N_{\mathcal{S}}(\mathcal{T}_n^{\mathbf{gw}}))_s}{(n\pi(\mathcal{S}))^s} - 1 \right| = o(1),$$

where $(x)_s := x(x-1)\cdots(x-s+1)$.

Thus if $n\pi(\mathcal{A}_n) \to \mu \in [0, \infty)$, then $\mathbb{E}(N_{\mathcal{A}_n}(\mathcal{T}_n^{\mathbf{gw}}))_s \to \mu^s$, for all fixed s. It is well-known that this implieslp$N_{\mathcal{A}_n}(\mathcal{T}_n^{\mathbf{gw}}) \xrightarrow{d} \text{Po}(\mu)$ (see van der Hofstad, 2017, thm. 2.4). Part (i) and (ii) of Theorem 1.4 follows immediately. And part (iii) comes from the following result:

Lemma 3.6 (Gao and Wormald, 2004, thm. 1)**.** *Let $(X_n)_{n \geqslant 1}$ be a sequence of integer-valued random variables. If there exists a sequence $\mu_n \to \infty$ such that*

$$\sup_{s \leqslant \sqrt{\mu_n}} \left| \frac{\mathbb{E}(X_n)_s}{\mu_n^s} - 1 \right| \to 0,$$

then we have

$$\frac{X_n - \mu_n}{\sqrt{\mu_n}} \xrightarrow{d} N(0, 1).$$

4. Families of fringe subtrees

In this section, we apply Theorem 1.3 and 1.4 to study the conditions for $\mathcal{T}_n^{\mathbf{gw}}$ to contain every tree that belongs to a family of trees.

4.1. *Coupon collector problem.* As shown later, our problem is essentially a variation of the famous coupon collector problem—if in every draw we get a coupon with a uniform random type among n types, how many draws do we need to collect all n types of coupons? The next lemma is about a generalization of this problem needed later. For the original problem, see Erdős and Rényi (1961) and Flajolet et al. (1992). For more about the generalized version defined below, see Neal (2008).

Lemma 4.1 (Generalized coupon collector)**.** *Let X_n be a random variable that takes values in $\{1, \ldots, n\}$. Let $p_{n,i} := \mathbb{P}\{X_n = i\}$. Assume that $p_{n,i} > 0$ for all $1 \leqslant i \leqslant n$. Let $X_{n,1}, X_{n,2}, \ldots$ be i.i.d. copies of X_n. Let*

$$N_n := \inf\{i \geqslant 1 : |\{X_{n,1}, X_{n,2}, \ldots, X_{n,i}\}| = n\}.$$

Let m_n be a sequence of real numbers. We have

$$1 - \sum_{i=1}^{n}(1 - p_{n,i})^{m_n} \leqslant \mathbb{P}\{N_n \leqslant m_n\} \leqslant \frac{1}{\sum_{i=1}^{n}(1 - p_{n,i})^{m_n}}.$$

Proof: Let $m = m_n$. Let $Z_{n,i} = [\![i \notin \{X_{n,1}, \ldots, X_{n,m}\}]\!]$. Then $N_n \leqslant m$ if and only if $Z_n := \sum_{i=1}^{n} Z_{n,i} = 0$, i.e., $\mathbb{P}\{N_n \leqslant m\} = \mathbb{P}\{Z_n = 0\} = 1 - \mathbb{P}\{Z_n \geqslant 1\}$.

The first inequality of this lemma follows from the following:

$$\mathbb{P}\{Z_n \geqslant 1\} \leqslant \mathbb{E}Z_n = \sum_{i=1}^{n} \mathbb{E}Z_{n,i} = \sum_{i=1}^{n} \mathbb{P}\{\cap_{j=1}^{m} X_{n,j} \neq i\} = \sum_{i=1}^{n}(1 - p_{n,i})^m.$$

For $1 \leqslant i \neq j \leqslant n$, we have

$$
\begin{aligned}
\mathbb{E}\left[Z_{n,i} Z_{n,j}\right] - \mathbb{E}\left[Z_{n,i}\right] \mathbb{E}\left[Z_{n,j}\right] &= (1 - p_{n,i} - p_{n,j})^m - (1 - p_{n,i})^m (1 - p_{n,j})^m \\
&= (1 - p_{n,i})^m \left[\left(1 - \frac{p_{n,j}}{1 - p_{n,i}}\right)^m - (1 - p_{n,j})^m\right] \\
&< 0.
\end{aligned}
$$

Therefore

$$
\begin{aligned}
\operatorname{Var}\left(Z_n\right) &= \sum_{1 \leqslant i,j \leqslant n} \mathbb{E}\left[Z_{n,i} Z_{n,j}\right] - \mathbb{E}\left[Z_{n,i}\right] \mathbb{E}\left[Z_{n,j}\right] \\
&= \sum_{1 \leqslant i \neq j \leqslant n} \left(\mathbb{E}\left[Z_{n,i} Z_{n,j}\right] - \mathbb{E}\left[Z_{n,i}\right] \mathbb{E}\left[Z_{n,j}\right]\right) \\
&\quad + \sum_{1 \leqslant i \leqslant n} \left(\mathbb{E}\left[Z_{n,i}\right] - \mathbb{E}\left[Z_{n,i}\right]^2\right) \\
&\leqslant \mathbb{E} Z_n.
\end{aligned}
$$

Thus by Chebyshev's inequality, as in the second moment method (see e.g., Alon and Spencer, 2008, chap. 4), we have

$$
\mathbb{P}\left\{Z_n = 0\right\} \leqslant \mathbb{P}\left\{|Z_n - \mathbb{E} Z_n| \geqslant \mathbb{E} Z_n\right\} \leqslant \frac{\operatorname{Var}\left(Z_n\right)}{(\mathbb{E} Z_n)^2} \leqslant \frac{1}{\mathbb{E} Z_n} = \frac{1}{\sum_{i=1}^n (1 - p_{n,i})^{m_n}}. \quad \square
$$

4.2. *Complete r-ary fringe subtrees.* A tree T is called possible if $\pi(T) > 0$. Let $r > 0$ be a fixed integer and h_n be a sequence of positive integers. A simple application of Theorem 1.3 is to find sufficient conditions such that whp every (or not every) possible complete r-ary tree appears in $\mathcal{T}_n^{\mathbf{gw}}$ as fringe subtrees.

Let $h_n \to \infty$ be a sequence of positive integers. Let $\mathcal{A}_{h_n,r}$ be the set of all *possible* complete r-ary trees of height at most h_n. Let

$$
H_{n,r} = \max\left\{h : \mathcal{T}_n^{\mathbf{gw}} \text{ contains all trees in } \mathcal{A}_{h_n,r} \text{ as fringe subtrees}\right\}.
$$

Lemma 4.2. *Assume Condition A and $p_r > 0$ for some $r \geqslant 2$. Let*

$$
\alpha_r = \log_r \left(\log \frac{1}{p_0} + \frac{1}{r-1} \log \frac{1}{p_r}\right).
$$

Let $\omega_n \to \infty$ be an arbitrary sequence.

(i) *If $h_n \leqslant \log_r(\log n - \omega_n) - \alpha_r$, then whp $\mathcal{T}_n^{\mathbf{gw}}$ contains all trees in $\mathcal{A}_{h_n,r}$ as fringe subtrees.*

(ii) *If $h_n \geqslant \log_r(\log n + \omega_n) - \alpha_r$, then whp $\mathcal{T}_n^{\mathbf{gw}}$ does not contain all trees in $\mathcal{A}_{h_n,r}$ as fringe subtrees.*

Also,

$$
H_{n,r} - \log_r \log n \xrightarrow{p} -\alpha_r.
$$

Proof: Let $T_{h_n}^{\mathbf{r\text{-}ary}}$ denote the complete r-ary tree of height h_n. Note that if $T_{h_n}^{\mathbf{r\text{-}ary}}$ appears in $\mathcal{T}_n^{\mathbf{gw}}$ as a fringe subtree, then every tree in $\mathcal{A}_{h_n,r}$ also appears in $\mathcal{T}_n^{\mathbf{gw}}$ as a fringe subtree. The tree $T_n^{\mathbf{r\text{-}ary}}$ has $\ell_n := r^{h_n}$ leaves and $v_n := (r^{h_n} - 1)/(r-1) = (\ell_n - 1)/(r-1)$ internal vertices, which all have degree r. Thus we have

$$
\pi(T_{h_n}^{\mathbf{r\text{-}ary}}) := \mathbb{P}\left\{\mathcal{T}^{\mathbf{gw}} = T_{h_n}^{\mathbf{r\text{-}ary}}\right\} = p_r^{v_n} p_0^{\ell_n}.
$$

If $h_n \leqslant \log_r(\log n - \omega_n) - \alpha_r$, then

$$\ell_n = r^{h_n} \leqslant \frac{\log n - \omega_n}{r^{\alpha_r}}.$$

Therefore

$$\begin{aligned}
\log \frac{1}{\pi(T_{h_n}^{\mathbf{r\text{-}ary}})} &= v_n \log \frac{1}{p_r} + \ell_n \log \frac{1}{p_0} \\
&= \frac{\ell_n - 1}{r - 1} \log \frac{1}{p_r} + \ell_n \log \frac{1}{p_0} \\
&= \ell_n \left(\frac{1}{r - 1} \log \frac{1}{p_r} + \log \frac{1}{p_0} \right) + O(1) \\
&\leqslant \frac{\log n - \omega_n}{r^{\alpha_r}} r^{\alpha_r} + O(1) \\
&= \log n - \omega_n + O(1).
\end{aligned}$$

Thus $\log(n\pi(T_{h_n}^{\mathbf{r\text{-}ary}})) \geqslant \omega_n + O(1) \to \infty$, which implies that $n\pi(T_{h_n}^{\mathbf{r\text{-}ary}}) \to \infty$. It follows from Theorem 1.3 that $N_{T_{h_n}^{\mathbf{r\text{-}ary}}}(\mathcal{T}_n^{\mathbf{gw}}) \xrightarrow{p} \infty$. Thus (i) is proved.

Similar computations show that the assumptions of (ii) implies that $n\pi(T_{h_n}^{\mathbf{r\text{-}ary}}) \to 0$. Thus $N_{T_{h_n}^{\mathbf{r\text{-}ary}}}(\mathcal{T}_n^{\mathbf{gw}}) \xrightarrow{p} 0$ by Theorem 1.3. The last statement of the lemma follows directly from (i) and (ii). □

We have a similar result for the set of 1-ary trees (chains) of height at most h. The proof is virtually identical to the previous lemma and we leave it to the reader.

Lemma 4.3. *Assume Condition A and $p_1 > 0$. Let $\omega_n \to \infty$ be an arbitrary sequence. We have:*

  (i) *If $h_n \leqslant (\log n - \omega_n)/\log \frac{1}{p_1}$, then whp $\mathcal{T}_n^{\mathbf{gw}}$ contains all trees in $\mathcal{A}_{h_n,1}$ as fringe subtrees.*
  (ii) *If $h_n \geqslant (\log n + \omega_n)/\log \frac{1}{p_1}$, then whp $\mathcal{T}_n^{\mathbf{gw}}$ does not contain all trees in $\mathcal{A}_{h_n,1}$ as fringe subtrees.*

Therefore

$$\frac{H_{n,1}}{\log_{1/p_1}(n)} \xrightarrow{p} 1.$$

4.2.1. *Binary trees.* Consider $\mathcal{T}_n^{\mathbf{gw}}$ with a binomial $(2, 1/2)$ offspring distribution, i.e., $p_0 = p_2 = 1/4$ and $p_1 = 1/2$. Let $\mathcal{T}_n^{\mathbf{bin}}$ be $\mathcal{T}_n^{\mathbf{gw}}$ with each degree-one node labeled of having a left or a right child uniformly and independently at random. Then $\mathcal{T}_n^{\mathbf{bin}}$ is a tree in which nodes have a left position and a right position where child nodes can attach, and each position can be occupied by at most one child. We call such a tree a *binary tree*.

Let $T_n^{\mathbf{bin}}$ be a binary tree of size n that has n_0, n_1, n_2 nodes of degree $0, 1, 2$ respectively, where $n_0 + n_1 + n_2 = n$. Let T_n be $T_n^{\mathbf{bin}}$ with the difference between left and right children being forgotten. We have $\mathbb{P}\left\{\mathcal{T}_n^{\mathbf{bin}} = T_n^{\mathbf{bin}} | \mathcal{T}_n^{\mathbf{gw}} = T_n\right\} = 1/2^{n_1}$,

since there are in total 2^{n_1} ways to label the n_1 degree-one nodes of T_n. Therefore

$$
\begin{aligned}
\mathbb{P}\left\{\mathcal{T}_n^{\mathbf{bin}} = T_n^{\mathbf{bin}}\right\} &= \mathbb{P}\left\{\mathcal{T}_n^{\mathbf{bin}} = T_n^{\mathbf{bin}} | \mathcal{T}_n^{\mathbf{gw}} = T_n\right\} \mathbb{P}\left\{\mathcal{T}_n^{\mathbf{gw}} = T_n\right\} \\
&= \mathbb{P}\left\{\mathcal{T}_n^{\mathbf{bin}} = T_n^{\mathbf{bin}} | \mathcal{T}_n^{\mathbf{gw}} = T_n\right\} \frac{\mathbb{P}\left\{\mathcal{T}^{\mathbf{gw}} = T_n\right\}}{\mathbb{P}\left\{|\mathcal{T}^{\mathbf{gw}}| = n\right\}} \\
&= \frac{1}{2^{n_1}} \frac{1}{4^{n_0 n_2} 2^{n_1}} \frac{1}{\mathbb{P}\left\{|\mathcal{T}^{\mathbf{gw}}| = n\right\}} = \frac{1}{4^n \mathbb{P}\left\{|\mathcal{T}^{\mathbf{gw}}| = n\right\}}.
\end{aligned}
$$

In other words, as is well-known from the connection between simply generated trees and Galton-Watson trees of size n, see, e.g., Janson (2012, pp. 132), $\mathcal{T}_n^{\mathbf{bin}}$ is uniformly distributed among all binary trees of size n.

Thus our analysis of maximum r-ary fringe subtree in $\mathcal{T}_n^{\mathbf{gw}}$ can be easily adapted to uniform random binary trees. For example, an argument similar to Lemma 4.3 shows that the maximum one-ary fringe subtree (chain) in $\mathcal{T}_n^{\mathbf{bin}}$ that consists of only left children has length about $\log_4 n$.

4.3. *All possible fringe subtrees.* Recall that $\mathfrak{T}_{\leqslant k}^+$ denotes the set of all possible trees of size at most k, i.e.,

$$
\mathfrak{T}_{\leqslant k}^+ := \{T \in \mathfrak{T} : |T| \leqslant k, \mathbb{P}\left\{\mathcal{T}^{\mathbf{gw}} = T\right\} > 0\}.
$$

Also recall that

$$
K_n := \max\{k : \mathfrak{T}_{\leqslant k}^+ \subseteq \cup_{v \in \mathcal{T}_n^{\mathbf{gw}}} \{\mathcal{T}_{n,v}^{\mathbf{gw}}\}\}.
$$

We would like to study the growth of K_n with n.

Let $\mathrm{Ge}(p)$ denote Geometric p distribution, i.e., $\mathbb{P}\left\{\mathrm{Ge}(p) = i\right\} = p(1-p)^i$ for $i \in \mathbb{N}_0$. Let $\mathrm{Be}(p)$ denote Bernoulli p distribution and let $\mathrm{Bi}(d, p)$ denote Binomial (d, p) distribution. Recall that $\mathrm{Po}(\lambda)$ is a Poisson(λ) random variable. Table 4.1 shows five types of well-known conditional Galton-Watson trees. See Janson (2012, sec. 10) for more examples.

Name	Definition	
Plane trees	$\xi \overset{\mathcal{L}}{=} \mathrm{Ge}(1/2)$	$p_i = 1/2^{i+1} \quad (i \geqslant 0)$
Full binary trees	$\xi \overset{\mathcal{L}}{=} 2\mathrm{Be}(1/2)$	$p_0 = p_2 = 1/2$
Motzkin trees	ξ uniform in $\{0, 1, 2\}$	$p_0 = p_1 = p_2 = 1/3$
d-ary trees	$\xi \overset{\mathcal{L}}{=} \mathrm{Bi}(d, 1/d)$	$p_i = \binom{d}{i}\left(1 - \frac{1}{d}\right)^i \left(\frac{1}{d}\right)^{d-i} \quad (0 \leqslant i \leqslant d)$
Labeled trees	$\xi \overset{\mathcal{L}}{=} \mathrm{Po}(1)$	$p_i = e^{-1}/i!$

TABLE 4.1. Some well-known conditional Galton-Watson trees.

We can assume that $\mathbb{P}\left\{|\mathcal{T}^{\mathbf{gw}}| = k_n\right\} > 0$ for all $n \in \mathbb{N}$. Otherwise, let $k_n' := \max\{i \leqslant k_n : \mathbb{P}\left\{|\mathcal{T}^{\mathbf{gw}}| = i\right\} > 0\}$. It is not difficult to show that $k_n - k_n' \leqslant h = O(1)$ for k_n large. (See Janson, 2012, lem. 12.3 for details.) Thus this assumption does not change results in this subsection.

Janson (2012, thm. 7.12) showed that $\mathcal{T}_{n,*}^{\mathbf{gw}} \overset{d}{\to} \mathcal{T}^{\mathbf{gw}}$. In other words, fringe subtrees on average behave like unconditional Galton-Watson trees. Let $T_k^{\mathbf{min}}$ be a tree $T \in \mathfrak{T}_{\leqslant k}^+$ that minimizes $\mathbb{P}\left\{\mathcal{T}^{\mathbf{gw}} = T\right\}$. Then $T_k^{\mathbf{min}}$ also is the least likely tree to appear in $\mathcal{T}_n^{\mathbf{gw}}$ as fringe subtree among all trees in $\mathfrak{T}_{\leqslant k}^+$ when n is large. So intuitively if whp $T_k^{\mathbf{min}}$ appears in $\mathcal{T}_n^{\mathbf{gw}}$, then every tree in $\mathfrak{T}_{\leqslant k}^+$ should also appears

whp. And if whp $T_k^{\mathbf{min}}$ does not appear, then whp there is at least one tree in $\mathfrak{T}_{\leqslant k}^+$ that is missing. Therefore, the problem can be reduced to finding

$$p_k^{\mathbf{min}} := \mathbb{P}\left\{\mathcal{T}^{\mathbf{gw}} = T_k^{\mathbf{min}}\right\} = \min_{T \in \mathfrak{T}_{\leqslant k}^+} \mathbb{P}\left\{\mathcal{T}^{\mathbf{gw}} = T\right\}.$$

Lemma 4.4. *Assume Condition A. If $k_n \to \infty$ and $np_{k_n}^{\mathbf{min}}/k_n \to \infty$, then whp every tree in $\mathfrak{T}_{\leqslant k_n}^+$ appears in $\mathcal{T}_n^{\mathbf{gw}}$.*

Proof: Let $k = k_n$. Recall that $p_{\mathbf{max}} := \max_{i \geqslant 0} p_i$ and $p_{\mathbf{max}} < 1$. Therefore $p_k^{\mathbf{min}} \leqslant p_{\mathbf{max}}^k$. Thus we can assume that $k \leqslant 2\log(n)/\log(p_{\mathbf{max}}^{-1})$ when k is large. Otherwise we have $np_k^{\mathbf{min}} \leqslant n^{-1} \to 0$, which contradicts the assumption.

Thus by Theorem 1.4, $N_{\mathfrak{T}_k}(\mathcal{T}_n^{\mathbf{gw}}) \geqslant y_n := \lceil \frac{1}{2}n\mathbb{P}\{\mathcal{T}^{\mathbf{gw}} = |k|\}\rceil$ whp. Let A_n be the event that $\mathcal{T}_n^{\mathbf{gw}}$ contains all possible trees of size k as fringe subtrees. Let $B_n(i)$ be the event that $N_{\mathfrak{T}_k}(\mathcal{T}_n^{\mathbf{gw}}) = i$ for given $i \geqslant y_n$. Thus

$$\mathbb{P}\left\{A_n\right\} = \mathbb{P}\left\{A_n \cap [N_{\mathfrak{T}_k}(\mathcal{T}_n^{\mathbf{gw}}) < y_n]\right\} + \sum_{i \geqslant y_n} \mathbb{P}\left\{A_n | B_n(i)\right\} \mathbb{P}\left\{B_n(i)\right\}$$

$$\geqslant \mathbb{P}\left\{A_n | B_n(y_n)\right\} \sum_{i \geqslant y_n} \mathbb{P}\left\{B_n(i)\right\},$$

$$= \mathbb{P}\left\{A_n | B_n(y_n)\right\}(1 - o(1)),$$

where the inequality comes from the obvious fact that $\mathbb{P}\left\{A_n | B_n(a)\right\} \geqslant \mathbb{P}\left\{A_n | B_n(b)\right\}$ given that $a \geqslant b$. So it suffices to prove that $\mathbb{P}\left\{A_n^c | B_n(y_n)\right\} \to 0$.

This is in fact equivalent to a coupon collector problem. Let $\mathcal{T}_n^{\mathbf{gw}} | B_n(y_n)$ be $\mathcal{T}_n^{\mathbf{gw}}$ restricted to the event $B_n(y_n)$. Let $\mathcal{T}^*$ be a random tree distributed as $\mathcal{T}_n^{\mathbf{gw}} | B_n(y_n)$. Replace each of its y_n subtrees with an independent copy of $\mathcal{T}_k^{\mathbf{gw}}$. By Lemma 3.4, the result is still a random tree distributed as $\mathcal{T}_n^{\mathbf{gw}} | B_n(y_n)$.

So $\mathbb{P}\left\{A_n | B_n(y_n)\right\}$ equals the probability of that y_n independent copies of $\mathcal{T}_k^{\mathbf{gw}}$ contain every tree in $\mathfrak{T}_{\leqslant k}^+$. It follows from Lemma 4.1 (the coupon collector) that

$$\mathbb{P}\left\{A_n^c | B_n(y_n)\right\} \leqslant \sum_{T \in \mathfrak{T}_{\leqslant k}^+} (1 - \mathbb{P}\{\mathcal{T}_k^{\mathbf{gw}} = T\})^{y_n} \leqslant \sum_{T \in \mathfrak{T}_{\leqslant k}^+} \exp\left\{-y_n\mathbb{P}\{\mathcal{T}_k^{\mathbf{gw}} = T\}\right\}$$

$$\leqslant \sum_{T \in \mathfrak{T}_{\leqslant k}^+} \exp\left\{-\frac{1}{2}n\mathbb{P}\{|\mathcal{T}^{\mathbf{gw}}| = k\}\frac{\mathbb{P}\{\mathcal{T}^{\mathbf{gw}} = T\}}{\mathbb{P}\{|\mathcal{T}^{\mathbf{gw}}| = k\}}\right\}$$

$$\leqslant O(|\mathfrak{T}_{\leqslant k}^+|) \exp\left\{-np_k^{\mathbf{min}}\right\}.$$

It is well-known that the number of plane trees of size exactly k is $4^{k-1}/\sqrt{\pi k^3}(1 + o(1))$. See, e.g., Flajolet and Sedgewick (2009, pp. 406). It follows that there exists a constant C such that $|\mathfrak{T}_{\leqslant k}^+| \leqslant \sum_{j=1}^k C4^{j-1} \leqslant C4^k$ for all $k \in \mathbb{N}$. Thus for large enough k, the last expression above is at most

$$C4^k \exp\left\{-np_k^{\mathbf{min}}\right\} = C\exp\left\{k\log(4) - np_k^{\mathbf{min}}\right\} \to 0.$$

Therefore we have $\mathbb{P}\left\{A_n^c | B_n(y_n)\right\} \to 0$. $\square$

Theorem 4.5. *Assume Condition A. Assume that as $k \to \infty$,*

$$\log(1/p_k^{\mathbf{min}}) \sim \gamma k^\alpha (\log k)^\beta,$$

where $\alpha \geqslant 1$, $\beta \geqslant 0$, $\gamma > 0$ are constants. Let $k_n \to \infty$ be a sequence of positive integers. Let $m = \log n$. Then for all constants $\delta > 0$, we have:

(i) If $k_n \leqslant (1 - \delta)[m/\gamma(\log m^{1/\alpha})^\beta]^{1/\alpha}$, then whp $\mathcal{T}_n^{\mathbf{gw}}$ contains all trees in $\mathfrak{T}_{\leqslant k_n}^+$ as fringe subtrees.

(ii) If $k_n \geqslant (1+\delta)[m/\gamma(\log m^{1/\alpha})^\beta]^{1/\alpha}$, then whp $\mathcal{T}_n^{\mathbf{gw}}$ does not contain all trees in $\mathfrak{T}_{\leqslant k_n}^+$ as fringe subtrees.

As a result,

$$\frac{K_n}{(\log n/(\log \log n)^\beta)^{1/\alpha}} \xrightarrow{p} \left(\frac{\alpha^\beta}{\gamma}\right)^{1/\alpha}.$$

The behavior of $p_k^{\mathbf{min}}$ varies for different offspring distributions. But as mentioned in the introduction, the types of trees that we are interested in all have $p_k^{\mathbf{min}}$ that are covered by Theorem 4.5.

Proof: Let $k = k_n$. Part (i) assumes that

$$k \leqslant (1 - \delta)\left[\frac{m}{\gamma(\log m^{1/\alpha})^\beta}\right]^{1/\alpha}.$$

Taking a logarithm, we have

$$\log k \leqslant \log(1 - \delta) + \frac{1}{\alpha}\log\frac{m}{\gamma(\log m^{1/\alpha})^\beta} = (1 + o(1))\log m^{1/\alpha}.$$

Thus

$$\log\frac{1}{p_k^{\mathbf{min}}} \sim \gamma k^\alpha(\log k)^\beta \leqslant (\gamma + o(1))\frac{(1 - \delta)^\alpha m}{\gamma(\log m^{1/\alpha})^\beta}(\log m^{1/\alpha})^\beta \sim (1 - \delta)^\alpha m.$$

Therefore, recalling $m = \log n$,

$$\log np_k^{\mathbf{min}} = m - \log\frac{1}{p_k^{\mathbf{min}}} \geqslant m - (1 + o(1))(1 - \delta)^\alpha m = \Omega(m).$$

It follows that

$$\log k - \log np_k^{\mathbf{min}} = O(\log m) - \Omega(m) \to -\infty.$$

Thus $np_k^{\mathbf{min}}/k \to \infty$ and it follows from Lemma 4.4 that whp $\mathcal{T}_n^{\mathbf{gw}}$ contains every tree in $\mathfrak{T}_{\leqslant k}^+$ as a fringe subtree.

Similar computations show that if $k \geqslant (1+\delta)[m/\gamma(\log m^{1/\alpha})^\beta]^{1/\alpha}$ then $np_k^{\mathbf{min}} \to 0$. It follows from Theorem 1.3 that $N_{T_k^{\mathbf{min}}}(\mathcal{T}_n^{\mathbf{gw}}) \xrightarrow{p} 0$. Thus whp $\mathcal{T}_n^{\mathbf{gw}}$ does not contain every tree in $\mathfrak{T}_{\leqslant k}^+$ as a fringe subtree. $\qquad\square$

Remark 4.6. The above coupon collector approach can be also used for studying the sufficient conditions for $\mathcal{T}_n^{\mathbf{gw}}$ to contain all r-ary trees of size k (not necessarily complete). We can think of these r-ary trees as types of coupon that we need to collect and $N_{\mathfrak{T}_{k_n}}(\mathcal{T}_n^{\mathbf{gw}})$ as the number of draws of coupons that we are allow to carry out.

The rest of this section is organized as follows. In the next subsection, we give a general method of finding $p_k^{\mathbf{min}}$. Then we divide offspring distributions in two categories and show that Theorem 4.5 is applicable to all the Galton-Watson trees listed in Table 4.1.

4.3.1. *Computing $p_k^{\mathbf{min}}$.* Let $\mathcal{I}_k := \{j : 1 \leqslant j \leqslant k, p_j > 0\}$. Let $L_1 := p_0$ and for $k \geqslant 2$ let

$$L_k := \min\left\{ p_0 \left(\frac{p_i}{p_0} \right)^{1/i} : i \in \mathcal{I}_{k-1} \right\}.$$

Since L_k is non-increasing, $L := \lim_{k \to \infty} L_k$ exists. Equivalently, we have

$$L := \inf\left\{ p_0 \left(\frac{p_i}{p_0} \right)^{1/i} : i \in \mathbb{N}, p_i > 0 \right\}.$$

Theorem 4.7. *Assume Condition* A. *We have* $(p_k^{\mathbf{min}})^{1/k} \to L$ *as* $k \to \infty$, *where the limit is taken along the subsequence k with* $\mathbb{P}\{|\mathcal{T}^{\mathbf{gw}}| = k\} > 0$. *As a result, we have* $L < 1$.

In fact, we have a stronger result for the upper bound of $(p_k^{\mathbf{min}})^{1/k}$.

Lemma 4.8. *Assume Condition* A. *For all fixed i with $p_i > 0$, there exist constants $C_i > 1$ and $C_i', C_i'', k(i) > 0$ such that for all $k \geqslant k(i)$ with $\mathbb{P}\{|\mathcal{T}^{\mathbf{gw}}| = k\} > 0$, there are at least $k^{-C_i'} C_i^k$ trees T of size k with*

$$0 < \mathbb{P}\{\mathcal{T}^{\mathbf{gw}} = T\} \leqslant C_i'' \left[p_0 \left(\frac{p_i}{p_0} \right)^{1/i} \right]^k.$$

Proof: We give a proof assuming that there exists a j such that $p_j > 0$ and that i and j are coprime. The proof of the general case is similar. Let $x = (k-1) \bmod i$. By the Chinese reminder theorem, there exists a smallest non-negative integer y such that

$$\begin{cases} y \equiv x \pmod{i}, \\ y \equiv 0 \pmod{j}. \end{cases}$$

Note that y depends only on i. Therefore, if $k \geqslant k(i) := y + 1$, we can choose

$$n_0 = k - n_i - n_j, \qquad n_j = \frac{y}{j}, \qquad n_i = \frac{k - 1 - y}{i},$$

such that n_0, n_i, n_j are all non-negative integers with

$$n_0 + n_i + n_j = k, \qquad \text{and} \qquad i n_i + j n_j = k - 1.$$

Let $\mathfrak{T}_k(n_0, n_i, n_j)$ be the set of plane trees of size k that has n_0, n_i and n_j nodes with degree 0, i and j respectively. It is well-known that when the above two conditions hold, we have

$$|\mathfrak{T}_k(n_0, n_i, n_j)| := \frac{1}{k} \binom{k}{n_0, n_i, n_j} = \frac{1}{k} \frac{k!}{n_0! n_i! n_j!}.$$

(See Flajolet and Sedgewick, 2009, pp. 194.) Since i is a constant and y only depends on i, there exists a constant C_i^* such that

$$|n_0 - k(1 - 1/i)| \leqslant C_i^*, \qquad |n_i - k/i| \leqslant C_i^*, \qquad n_j \leqslant C_i^*.$$

Using these inequalities and Stirling's approximation (Flajolet and Sedgewick, 2009, pp. 407), it is easy to verify that there exists a constant $C_i' > 0$ such that

$$|\mathfrak{T}_k(n_0, n_i, n_j)| \geqslant k^{-C_i'} \left(\left(\frac{1}{i} \right)^{1/i} \left(1 - \frac{1}{i} \right)^{1 - 1/i} \right)^{-k} := k^{-C_i'} C_i^k.$$

And for every $T \in \mathcal{S}(n_0, n_i, n_j)$, we have

$$\mathbb{P}\{\mathcal{T}^{\mathbf{gw}} = T\} \leqslant p_i^{n_i} p_0^{n_0} \leqslant p_i^{-C_i^*} p_0^{-C_i^*} p_i^{k/i} p_0^{k(1-1/i)} := C_i'' \left[p_0 \left(\frac{p_i}{p_0} \right)^{1/i} \right]^k. \qquad \square$$

Proof of Theorem 4.7: Let T be a tree with $|T| = k$ and $\mathbb{P}\{\mathcal{T}^{\mathbf{gw}} = T\} > 0$, i.e., $T \in \mathfrak{T}_{\leqslant k}^+$. Let n_i be the number of nodes of degree i in T. Note that if $i > 0$ and $i \notin \mathcal{I}_{k-1}$, then $n_i = 0$. Since by (2.1) the sum of the degrees in a preorder degree sequence equals the size of the tree minus one, we have

$$n_0 + n_1 + \ldots n_{k-1} = k, \qquad \text{and} \qquad n_1 + 2n_2 \ldots + (k-1)n_{k-1} = k - 1.$$

Using the convention that $0^0 = 1$, we have for $k \geqslant 2$

$$\begin{aligned} \mathbb{P}\{\mathcal{T}^{\mathbf{gw}} = T\} &= p_0^{n_0} p_1^{n_1} \cdots p_{k-1}^{n_{k-1}} \\ &= p_0^{n_0 + n_1 + \ldots + n_{k-1}} \left(\frac{p_1}{p_0} \right)^{n_1} \left(\frac{p_2}{p_0} \right)^{n_2} \cdots + \left(\frac{p_{k-1}}{p_0} \right)^{n_{k-1}} \\ &= p_0^k \prod_{i \in \mathcal{I}_{k-1}} \left[\left(\frac{p_i}{p_0} \right)^{\frac{1}{i}} \right]^{i n_i} \geqslant p_0^k \left[\min_{i \in \mathcal{I}_{k-1}} \left(\frac{p_i}{p_0} \right)^{\frac{1}{i}} \right]^{\sum_{i=1}^{k-1} i n_i} \\ &= p_0 L_k^{k-1} \geqslant p_0 L^{k-1}. \end{aligned} \qquad (4.1)$$

As a result $\liminf_{k \to \infty} (p_k^{\mathbf{min}})^{1/k} \geqslant L$.

To show the other way, let $\varepsilon > 0$ be a constant, and let $\alpha = \min\{i : L_{i+1} \leqslant L + \varepsilon\}$. Therefore $0 < p_0(p_\alpha/p_0)^{1/\alpha} \leqslant L + \varepsilon$. By Lemma 4.8, there is at least one tree T of size k such that

$$p_k^{\mathbf{min}} \leqslant \mathbb{P}\{\mathcal{T}^{\mathbf{gw}} = T\} \leqslant C_\alpha \left[p_0 \left(\frac{p_\alpha}{p_0} \right)^{\frac{1}{\alpha}} \right]^k \leqslant C_\alpha (L + \varepsilon)^k,$$

where $C_\alpha > 0$ is constant. Thus $\limsup_{k \to \infty} (p_k^{\mathbf{min}})^{1/k} \leqslant L + \varepsilon$. Since ε is arbitrary, we have $\limsup_{k \to \infty} (p_k^{\mathbf{min}})^{1/k} \leqslant L$.

Recall that $p_{\mathbf{max}} := \max_{i \geqslant 0} p_i < 1$. For all trees T with size k, we have $\mathbb{P}\{\mathcal{T}^{\mathbf{gw}} = T\} \leqslant p_{\mathbf{max}}^k$, i.e., $(p_k^{\mathbf{min}})^{1/k} \leqslant p_{\mathbf{max}}$. Thus $L = \lim_{k \to \infty} (p_k^{\mathbf{min}})^{1/k} \leqslant p_{\mathbf{max}} < 1$. $\qquad \square$

4.3.2. *When $L > 0$.* If $L > 0$, then by Theorem 4.7, $\log(1/(p_k^{\mathbf{min}})^{1/k}) \sim \log(1/L)k = \log(1/L)k(\log k)^0$. Thus we can apply Theorem 4.5 with $\gamma = \log(1/L)$, $\alpha = 1$ and $\beta = 0$ to get

$$\frac{K_n}{\log(n)} \overset{p}{\to} \frac{1}{\log(1/L)}.$$

The following Lemma computes L for some well-known Galton-Watson trees. See Janson (2012, sec. 10) for more about these trees.

Lemma 4.9. *(i) Full binary tree: If $\xi \overset{\mathcal{L}}{=} 2\operatorname{Be}(1/2)$, then $L = 1/2$. (ii) Motzkin tree: If $p_0 = p_1 = p_2 = 1/3$, then $L = 1/3$. (iii) d-ary tree: If $\xi \overset{\mathcal{L}}{=} \operatorname{Bi}(d, 1/d)$ for $d \geqslant 2$, then $L = (d-1)^{d-1}/d^d$. (iv) Plane tree: If $\xi \overset{\mathcal{L}}{=} \operatorname{Ge}(1/2)$, then $L = 1/4$.*

Proof: (i): If $\xi \sim 2\,\mathrm{Be}(1/2)$, then $p_0 = 1/2$, $p_2 = 1/2$ and $p_i = 0$ for $i \notin \{0, 2\}$. Thus for $k \geqslant 3$, we have

$$L_k = \min_{i:i<k,p_i>0} p_0 \left(\frac{p_i}{p_0}\right)^{1/i} = p_0 \left(\frac{p_2}{p_0}\right)^{1/2} = 1/2.$$

Therefore $L = \lim_{k\to\infty} L_k = 1/2$. (ii) and (iii) follow from similar simple calculations.

(iv): For all $i \geqslant 1$, we have

$$p_0 \left(\frac{p_i}{p_0}\right)^{1/i} = \frac{1}{2}\left(\frac{1}{2^i}\right)^{1/i} = 1/4.$$

Therefore $L_k = 1/4$ for all $k \geqslant 1$, and $L = 1/4$. $\qquad\square$

Define

$$\kappa = \begin{cases} \min\{i \in \mathbb{N} : L_{i+1} = L\} & \text{if } L_j = L \text{ for some } j, \\ \infty & \text{otherwise.} \end{cases}$$

If $\kappa < \infty$, then we call the Galton-Watson tree *well-behaved*. Examples of such trees include those for which ξ is bounded, and those for which ξ has a polynomial or sub-exponential tail. The case $\xi \stackrel{\mathcal{L}}{=} \mathrm{Ge}(1/2)$ is also well-behaved. Thus the four types of Galton-Watson trees in Lemma 4.9 are well-behaved. The following theorem gives better thresholds than Theorem 4.5.

Theorem 4.10. *Assume Condition A and let the Galton-Watson tree be well-behaved. Then for all constants $\delta > 0$, we have:*

(i) *If $k_n \leqslant (\log n - (1+\delta)\log\log n)/\log\frac{1}{L}$, then whp $\mathcal{T}_n^{\mathbf{gw}}$ contains all trees in $\mathfrak{T}_{\leqslant k_n}^+$ as fringe subtrees.*

(ii) *If $k_n \geqslant (\log n - (1-\delta)\log\log n)/\log\frac{1}{L}$, then whp $\mathcal{T}_n^{\mathbf{gw}}$ does not contain all trees in $\mathfrak{T}_{\leqslant k_n}^+$ as fringe subtrees.*

Thus as $n \to \infty$, we have

$$\frac{K_n \log(1/L) - \log n}{\log\log n} \stackrel{p}{\to} -1.$$

Remark 4.11. As can be seen from the proof of Lemma 4.8, among well-behaved Galton-Watson trees of size k_n, the least possible are those contain maximal number of degree κ nodes. These are the trees that do not appear in case (ii).

The main idea is that when $\kappa < \infty$, there are exponentially many trees of size k that have small probability to appear as fringe subtrees in $\mathcal{T}_n^{\mathbf{gw}}$. Then we can use Lemma 4.1 (the coupon collector) to find the sufficient condition for one of them to not to appear whp.

Proof: (i): Write $m = \log n$ and $k = k_n$. Using (4.1), it is easy to verify that in this case $np_k^{\mathbf{min}}/k \to \infty$. Thus (i) follows from Lemma 4.4.

(ii): The proof is similar to the one of Lemma 4.4. As in that proof, we can assume that $k = O(\log n)$. Thus by Theorem 1.4, whp $N_{\mathfrak{T}_k}(\mathcal{T}_n^{\mathbf{gw}}) \leqslant y_n := \lfloor \frac{3}{2} n \mathbb{P}\{|\mathcal{T}^{\mathbf{gw}}| = k\}\rfloor$. Let A_n be the event that $\mathcal{T}_n^{\mathbf{gw}}$ contains all possible trees of size

k as fringe subtrees. Let $B_n(i)$ be the event that $N_{\mathfrak{T}_k}(\mathcal{T}_n^{\mathbf{gw}}) = i$ for some $i \leqslant y_n$. Then

$$\mathbb{P}\{A_n\} \leqslant \mathbb{P}\{N_{\mathfrak{T}_k}(\mathcal{T}_n^{\mathbf{gw}}) > y_n\} + \sum_{i \leqslant y_n} \mathbb{P}\{A_n | B_n(i)\} \mathbb{P}\{B_n(i)\}$$

$$\leqslant o(1) + \mathbb{P}\{A_n | B_n(y_n)\} \sum_{i \leqslant y_n} \mathbb{P}\{B_n(i)\},$$

$$\leqslant o(1) + \mathbb{P}\{A_n | B_n(y_n)\}.$$

Thus it suffices to prove that $\mathbb{P}\{A_n | B_n(y_n)\} \to 0$.

Using the same coupling as in the proof of Lemma 4.4, we have $\mathbb{P}\{A_n | B_n(y_n)\}$ equals the probability that y_n independent copies of $\mathcal{T}_k^{\mathbf{gw}}$ do not contains all trees in $\mathfrak{T}_{\leqslant k}^+$. It follows from Lemma 4.1 (the coupon collector) that $\mathbb{P}\{A_n | B_n(y_n)\} \to 0$ if $\sum_{T \in \mathfrak{T}_{\leqslant k}^+}(1 - \mathbb{P}\{\mathcal{T}_k^{\mathbf{gw}} = T\})^{y_n}$ goes to infinity.

By definition of κ, we have $L = p_0(p_\kappa/p_0)^{1/\kappa}$. It follows from Lemma 4.8 that there exists constants $C_\kappa > 1$ and $C_\kappa', C_\kappa'' > 0$ such that there are at least $k^{-C_\kappa'} C_\kappa^k$ trees T in $\mathfrak{T}_{\leqslant k}^+$ with

$$\mathbb{P}\{\mathcal{T}_k^{\mathbf{gw}} = T\} = \frac{\mathbb{P}\{\mathcal{T}^{\mathbf{gw}} = T\}}{\mathbb{P}\{|\mathcal{T}^{\mathbf{gw}}| = k\}} \leqslant C_\kappa'' \frac{(p_0(p_\kappa/p_0)^{1/\kappa})^k}{\mathbb{P}\{|\mathcal{T}^{\mathbf{gw}}| = k\}} = \frac{C_\kappa'' L^k}{\mathbb{P}\{|\mathcal{T}^{\mathbf{gw}}| = k\}}.$$

Therefore

$$\sum_{T \in \mathfrak{T}_{\leqslant k}^+}(1 - \mathbb{P}\{\mathcal{T}_k^{\mathbf{gw}} = T\})^{y_n} \geqslant k^{-C_\kappa'} C_\kappa^k \left(1 - \frac{C_\kappa'' L^k}{\mathbb{P}\{|\mathcal{T}^{\mathbf{gw}}| = k\}}\right)^{y_n}.$$

Since $L < 1$ and $\mathbb{P}\{|\mathcal{T}^{\mathbf{gw}}| = k\} = \Theta(k^{-3/2})$, we have $L^k/\mathbb{P}\{|\mathcal{T}^{\mathbf{gw}}| = k\} = o(1)$. Thus for k large enough, the logarithm of the above is

$$k\log(C_\kappa) - C_\kappa'\log(k) + y_n \log\left(1 - \frac{C_\kappa'' L^k}{\mathbb{P}\{|\mathcal{T}^{\mathbf{gw}}| = k\}}\right)$$

$$\geqslant \frac{1}{2}k\log(C_\kappa) - y_n \frac{C_\kappa'' L^k}{\mathbb{P}\{|\mathcal{T}^{\mathbf{gw}}| = k\}}$$

$$\geqslant \frac{1}{2}k\log(C_\kappa) - \frac{3}{2}nC_\kappa'' L^k = \frac{1}{2}k\log(C_\kappa) - O(nL^k).$$

By our assumptions, $k = \Omega(\log n)$ and $L^k \leqslant (\log n)^{1-\delta}/n$. Since $C_\kappa > 0$, we have

$$\frac{1}{2}k\log(C_\kappa) - O(nL^k) \geqslant \Omega(\log n) - O\left(n\frac{(\log n)^{1-\delta}}{n}\right) \to \infty,$$

which implies $\mathbb{P}\{A_n | B_n(y_n)\} \to 0$. □

Remark 4.12. If $L > 0$ and $\kappa = \infty$, then Theorem 4.5 shows that

$$K_n/\log(n) \xrightarrow{p} 1/\log(1/L).$$

But the second order term of K_n is sensitive to small modifications of the offspring distribution, which makes it slightly more challenging to analyze the second order term.

4.3.3. *When $L = 0$.* It is clear that $L = 0$ if and only if ξ has infinite support and $\liminf_{i\to\infty} p_0(p_i/p_0)^{1/i} = 0$, which implies $\limsup_{i\to\infty} \log(1/p_i)/i = \infty$, along the subsequence with $p_i > 0$. If in addition we have $p_i > 0$ for all $i \geqslant 0$ and $\log(1/p_i) \sim f(i)$ for some $f : [0, \infty) \to [0, \infty)$ with $f(i)/i \uparrow \infty$, then we say that ξ has an *f-super-exponential tail*. We have the following threshold for Galton-Watson trees with such a property.

Theorem 4.13. *Assume Condition A and let ξ have an f-super-exponential tail. Let f^{-1} denote the inverse of f. Then for all constants $\delta > 0$, we have*

(1) *If $k_n \leqslant f^{-1}((1 - \delta) \log n) + 1$, then whp $\mathcal{T}_n^{\mathbf{gw}}$ contains all trees in $\mathfrak{T}_{\leqslant k_n}^+$ as fringe subtrees.*

(2) *If $k_n \geqslant f^{-1}((1 + \delta) \log n) + 1$, then whp $\mathcal{T}_n^{\mathbf{gw}}$ does not contain all trees in $\mathfrak{T}_{\leqslant k_n}^+$ as fringe subtrees.*

Therefore,

$$\frac{K_n}{f^{-1}(\log n)} \xrightarrow{p} 1.$$

Proof: (i): Let $k = k_n$. Choose $\varepsilon > 0$ such that $(1 - \delta)(1 + \varepsilon) < (1 - \delta/2)$. Since $\log(1/p_i) \sim f(i)$, there exists an integer $i(\varepsilon)$ such that for all $i > i(\varepsilon)$,

$$\log p_i \geqslant -(1 + \varepsilon/2) f(i).$$

Let $w_i := \log p_0 (p_i/p_0)^{1/i}$. We have as $k \to \infty$,

$$\min_{i(\varepsilon) < i < k} w_i = \min_{i(\varepsilon) < i < k} \left\{ \left(1 - \frac{1}{i}\right) \log(p_0) + \frac{\log p_i}{i} \right\}$$

$$\geqslant \log(p_0) - \max_{i(\varepsilon) < i < k} \frac{(1 + \varepsilon/2) f(i)}{i}$$

$$= \log(p_0) - \frac{(1 + \varepsilon/2) f(k - 1)}{k - 1} \to -\infty,$$

where we use that $f(i)/i \uparrow \infty$. Since $\min_{1 \leqslant i \leqslant i(\varepsilon)} w_i$ is a constant, we have for large k,

$$\log L_k := \min_{1 \leqslant i < k} w_i \geqslant \log(p_0) - \frac{(1 + \varepsilon/2) f(k - 1)}{k - 1}.$$

It follows from (4.1) that for k large enough,

$$\log p_k^{\mathbf{min}} \geqslant \log(p_0 L_k^{k-1})$$

$$\geqslant \log(p_0) + (k - 1) \log p_0 - (1 + \varepsilon/2) f(k - 1).$$

$$\geqslant -(1 + \varepsilon) f(k - 1),$$

where the last step uses $f(k)/k \uparrow \infty$.

The assumption $k - 1 \leqslant f^{-1}((1 - \delta) \log n)$ implies that $f(k - 1) \leqslant (1 - \delta) \log n$ and $k = O(\log n)$. Thus

$$\log p_k^{\mathbf{min}} \geqslant -(1 + \varepsilon)(1 - \delta) \log n \geqslant -(1 - \delta/2) \log n.$$

Thus $n p_k^{\mathbf{min}} \geqslant n^{\delta/2}$. We have $n p_k^{\mathbf{min}}/k \to \infty$. It follows from Lemma 4.4 that $\mathcal{T}_n^{\mathbf{gw}}$ contains all possible trees of size at most k as fringe subtree whp.

(ii): Let $T_{k-1}^{\mathbf{star}}$ be the tree in which one node has degree $k - 1$ and all other nodes are leaves. Computations similar to above show that if $k - 1 \geqslant f^{-1}((1 + \delta) \log n)$, then $n\pi(T_{k-1}^{\mathbf{star}}) \to 0$. Therefore $\mathcal{T}_n^{\mathbf{gw}}$ does not contain $T_{k-1}^{\mathbf{star}}$ whp. □

Example 4.14 (The discrete Gaussian distribution). When $p_i = ce^{-c'i^2}$ for some appropriate positive normalization constants c and c', we have $L = 0$, and Theorem 4.13 applies. Then

$$\frac{K_n}{\sqrt{\log(n)}} \xrightarrow{p} \frac{1}{\sqrt{c'}},$$

as $n \to \infty$.

Example 4.15 (The Cayley trees). A better example is the Galton-Watson tree with offspring distribution $\xi \overset{\mathcal{L}}{=} \text{Po}(1)$, i.e., the Cayley tree. It has $p_i = e^{-1}/i!$ and $\log(1/p_i) \sim i \log(i)$. It is easy to see that

$$\frac{K_n \log \log n}{\log n} \xrightarrow{p} 1.$$

Using (4.1) it is not difficult to verify that the tail drops so fast that the least possible tree of size k is $T_{k-1}^{\mathbf{star}}$. This is a special case of the following general observation.

Lemma 4.16. *Assume Condition A. If $p_i > 0$ for all $i \geqslant 0$ and $p_i^{1/i} \downarrow 0$, then for k large enough, $p_k^{\mathbf{min}} = \mathbb{P}\left\{\mathcal{T}^{\mathbf{gw}} = T_{k-1}^{\mathbf{star}}\right\} = p_0^{k-1}p_{k-1}$. In particular, this is true for $\xi \overset{\mathcal{L}}{=} \text{Po}(1)$. In the latter case we have*

$$\log p_k^{\mathbf{min}} = \log(p_0^{k-1}p_{k-1}) = -k \log k (1 + O(1/k)).$$

5. Non-fringe subtrees

In this section we prove Theorem 1.6, the concentration of non-fringe subtree counts in conditional Galton-Watson trees.

Given a tree T, let $v(T)$ be the number of its internal nodes and let $\ell(T)$ be the number of its leaves. Recall that $N_T^{\mathbf{nf}}(\mathcal{T}_n^{\mathbf{gw}}) := \sum_{u \in \mathcal{T}_n^{\mathbf{gw}}} [\![T \prec \mathcal{T}_{n,u}^{\mathbf{gw}}]\!]$, and that $\widetilde{\boldsymbol{\xi}}^{\mathbf{n}} := (\widetilde{\xi}_1^{\mathbf{n}}, \ldots, \widetilde{\xi}_n^{\mathbf{n}})$ is a uniform random rotation of the preorder degree sequence of $\mathcal{T}_n^{\mathbf{gw}}$.

To simplify the notation, write $v := v(T)$ and $\ell := \ell(T)$. By Lemma 2.1, T has a preorder degree sequence of the form

$$(\boldsymbol{a}_1, 0, \boldsymbol{a}_2, 0, \ldots, \boldsymbol{a}_\ell, 0) :=$$
$$(a_{1,1}, a_{1,2}, \ldots, a_{1,r(1)}, 0, a_{2,1}, a_{2,2}, \ldots, a_{2,r(2)}, 0, \ldots, a_{\ell,1}, a_{2,2}, \ldots, a_{\ell,r(\ell)}, 0)$$
$$(5.1)$$

for non-negative integers $r(1), r(2), \ldots, r(\ell)$ and that

$$\sum_{s=1}^{\ell} r(s) = v, \qquad a_{s,t} > 0, \qquad \sum_{s=1}^{\ell} \boldsymbol{a}_s := \sum_{s=1}^{\ell} \sum_{t=1}^{r(s)} a_{s,t} = v + \ell - 1. \qquad (5.2)$$

Therefore, if $T \prec T'$, then T' has a preorder degree sequence of the form

$$(\boldsymbol{a}_1, \boldsymbol{b}_1, \boldsymbol{a}_2, \boldsymbol{b}_2, \ldots, \boldsymbol{a}_\ell, \boldsymbol{b}_\ell) \qquad (5.3)$$

where $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_\ell$ are preorder degree sequences of some plane trees. Thus each non-fringe subtree of shape T in $\mathcal{T}_n^{\mathbf{gw}}$ corresponds to a segment of $\widetilde{\boldsymbol{\xi}}^{\mathbf{n}}$ of the form of (5.3). If none of the segments overlap with each other, then we can permute them into the form $(\boldsymbol{a}_1, \ldots, \boldsymbol{a}_\ell, \boldsymbol{b}_1, \ldots, \boldsymbol{b}_\ell)$. Recall that $\widetilde{\boldsymbol{\xi}}^{\mathbf{n}}$ is permutation invariant, i.e., if

we permute $\widetilde{\boldsymbol{\xi}}^{\mathbf{n}}$, the result has the same distribution as $\widetilde{\boldsymbol{\xi}}^{\mathbf{n}}$. Thus $N_T^{\mathbf{nf}}(\mathcal{T}_n^{\mathbf{gw}})$ should be almost distributed like the number of the patterns $(\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_\ell)$ in $\widetilde{\boldsymbol{\xi}}^{\mathbf{n}}$.

The problem with this argument is that non-fringe subtrees can overlap. But as shown later in this section, under the assumptions of Theorem 1.6, the effect of such overlaps is negligible.

We will use $\mathcal{D}_n$ to denote the set of preorder degree sequences of trees with size n. Let $\widetilde{\mathcal{D}}_n$ be the set of sequences that are cyclic rotations of sequences in $\mathcal{D}_n$. Given $\boldsymbol{d} := (d_1, \ldots, d_n) \in \widetilde{\mathcal{D}}_n$, let $\deg_i(\boldsymbol{d}) := (d_i, d_{i+1}, \ldots, d_{i+k-1})$ such that $\deg_i(\boldsymbol{d}) \in \mathcal{D}_k$ for some $k \geqslant 1$, where the indices are all modulo n. Lemma 2.1 guarantees that such $\deg_i(\boldsymbol{d})$ exists and is unambiguous. Let $T_i(\boldsymbol{d})$ be the tree with the preorder degree sequence $\deg_i(\boldsymbol{d})$.

5.1. *Factorial moments.* Let $(x)_r := x(x-1)\cdots(x-r+1)$. For a random variable X, $\mathbb{E}(X)_r$ is called the *r-th factorial moment* of X. We give exact formulas for the first and second factorial moments of $N_T^{\mathbf{nf}}(\mathcal{T}_n^{\mathbf{gw}})$ in this subsection.

Lemma 5.1. *Assume that* $\mathbb{P}\{|\mathcal{T}^{\mathbf{gw}}| = n\} > 0$. *Let T be a tree. We have*

$$\frac{\mathbb{E}\left[N_T^{\mathbf{nf}}(\mathcal{T}_n^{\mathbf{gw}})\right]}{n} = \pi^{\mathbf{nf}}(T)\frac{\mathbb{P}\left\{S_{n-v(T)} = n - v(T) - \ell(T)\right\}}{\mathbb{P}\left\{S_n = n - 1\right\}}.$$

Proof: Let $v := v(T)$ and $\ell := \ell(T)$. Let $I_i = [\![T < T_i(\widetilde{\boldsymbol{\xi}}^{\mathbf{n}})]\!]$. Then $N_T^{\mathbf{nf}}(\mathcal{T}_n^{\mathbf{gw}}) = \sum_{i=1}^n I_i$. By the permutation invariance of $\widetilde{\boldsymbol{\xi}}^{\mathbf{n}}$, we have

$$\mathbb{E}\left[N_T^{\mathbf{nf}}(\mathcal{T}_n^{\mathbf{gw}})\right] = \mathbb{E}\left[\sum_{i=1}^n I_i\right] = n\mathbb{P}\{I_1 = 1\}.$$

Recall that T has a preorder degree sequence of the form $(\boldsymbol{a}_1, 0, \ldots, \boldsymbol{a}_\ell, 0)$ satisfying (5.2). Let $\mathcal{A} \subseteq \widetilde{\mathcal{D}}_n$ be the set of sequences such that $\widetilde{\boldsymbol{\xi}}^{\mathbf{n}} \in \mathcal{A}$ if and only if $I_1 = 1$. In other words, $\boldsymbol{d} := (d_1, d_2, \ldots, d_n) \in \mathcal{A}$ if and only if $\deg_1(\boldsymbol{d}) = (\boldsymbol{a}_1, \boldsymbol{b}_1, \ldots, \boldsymbol{a}_\ell, \boldsymbol{b}_\ell)$ for some $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_\ell$ which are preorder degree sequences of trees. By permuting $\deg_1(\boldsymbol{d})$ into $(\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_\ell, \boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_\ell)$, we get a new sequence $\boldsymbol{d}' := (d_1', d_2', \ldots, d_n') \in \mathcal{A}'$ where

$$\mathcal{A}' := \left\{(e_1, e_2, \ldots, e_n) \in \widetilde{\mathcal{D}}_n : (e_1, e_2, \ldots, e_v) = (\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_\ell)\right\}.$$

Such a permutation defines a mapping $f : \mathcal{A} \to \mathcal{A}'$.

For every $\boldsymbol{d}' \in \mathcal{A}'$, condition (5.2) implies that in $\boldsymbol{d}'$ after $(\boldsymbol{a}_1, \ldots, \boldsymbol{a}_\ell)$, there are at least ℓ consecutive segments that are preorder degree sequences of trees, i.e., there is a unique $\boldsymbol{d} \in \mathcal{A}$ with $f(\boldsymbol{d}) = \boldsymbol{d}'$. Thus f is a one-to-one mapping. If $\boldsymbol{d}' = f(\boldsymbol{d})$, then $\mathbb{P}\left\{\widetilde{\boldsymbol{\xi}}^{\mathbf{n}} = \boldsymbol{d}\right\} = \mathbb{P}\left\{\widetilde{\boldsymbol{\xi}}^{\mathbf{n}} = \boldsymbol{d}'\right\}$, since $\widetilde{\boldsymbol{\xi}}^{\mathbf{n}}$ is permutation invariant. Therefore we have

$$\mathbb{P}\{I_1 = 1\} = \mathbb{P}\left\{\widetilde{\boldsymbol{\xi}}^{\mathbf{n}} \in \mathcal{A}\right\} = \mathbb{P}\left\{\widetilde{\boldsymbol{\xi}}^{\mathbf{n}} \in \mathcal{A}'\right\}.$$

Recall that by Lemma 2.3, $\widetilde{\boldsymbol{\xi}}^{\mathbf{n}} \sim (\xi_1, \ldots, \xi_n | S_n = n - 1)$, where $\xi_1, \ldots, \xi_n$ are i.i.d. copies of ξ and $S_n = \sum_{s=1}^{n} \xi_s$. We have

$$
\begin{aligned}
\mathbb{P}\left\{\widetilde{\boldsymbol{\xi}}^{\mathbf{n}} \in \mathcal{A}'\right\} &= \mathbb{P}\left\{(\widetilde{\xi}_1^{\mathbf{n}}, \widetilde{\xi}_2^{\mathbf{n}}, \ldots, \widetilde{\xi}_v^{\mathbf{n}}) = (\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_\ell)\right\} \\
&= \frac{\mathbb{P}\left\{(\xi_1, \xi_2, \ldots, \xi_v) = (\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_\ell), S_n = n - 1\right\}}{\mathbb{P}\left\{S_n = n - 1\right\}} \\
&= \mathbb{P}\left\{T \prec \mathcal{T}^{\mathbf{gw}}\right\} \frac{\mathbb{P}\left\{S_{n-v} = n - v - \ell\right\}}{\mathbb{P}\left\{S_n = n - 1\right\}},
\end{aligned}
$$

where in the last step we use $\sum_{s=1}^{\ell} \boldsymbol{a}_s = v + \ell - 1$. □

To compute $\mathbb{E}(N_T^{\mathbf{nf}}(\mathcal{T}_n^{\mathbf{gw}}))_2$, we enumerate all the cases that T can appear as overlapping non-fringe subtrees by constructing a set of trees $\{T \boxplus T\}$ as follows. For trees T, S and node $v \in T$, let $T' = \mathrm{Splay}(T, v, S)$ denote tree T with subtree T_v replaced by S. Thus $T'_v = S$. Let $\mathcal{V}(T)$ denote the internal nodes of T. Then define the collection

$$
\{T \boxplus T\} = \left[\bigcup_{v \in \mathcal{V}(T) : T_v \prec T} \{\mathrm{Splay}(T, v, T)\}\right] \setminus \{T\}.
$$

Note that $|\{T \boxplus T\}| < v(T)$. Also note that given $T' \in \{T \boxplus T\}$ we can always find a unique node $v \in \mathcal{V}(T)$ such that $T' = \mathrm{Splay}(T, v, T)$.

Lemma 5.2. *Assume that* $\mathbb{P}\left\{|\mathcal{T}^{\mathbf{gw}}| = n\right\} > 0$. *Let* T *be a tree. We have*

$$
\begin{aligned}
\mathbb{E}\left[(N_T^{\mathbf{nf}}(\mathcal{T}_n^{\mathbf{gw}}))_2\right] &= n(n - 2v(T) + 1)\pi^{\mathbf{nf}}(T)^2 \\
&\quad \times \frac{\mathbb{P}\left\{S_{n-2v(T)} = n + 1 - 2(v(T) + \ell(T))\right\}}{\mathbb{P}\left\{S_n = n - 1\right\}} \\
&\quad + 2n \sum_{T' \in \{T \boxplus T\}} \pi^{\mathbf{nf}}(T') \frac{\mathbb{P}\left\{S_{n-v(T')} = n - v(T') - \ell(T')\right\}}{\mathbb{P}\left\{S_n = n - 1\right\}}.
\end{aligned}
$$

Proof: Let $v = v(T)$ and $\ell = \ell(T)$. Let I_i be defined as in the proof of Lemma 5.1. Since $I_1, \ldots, I_n$ are indicator random variables and permutation invariant, we have

$$
\mathbb{E}\left[(N_T^{\mathbf{nf}}(\mathcal{T}_n^{\mathbf{gw}}))_2\right] = \sum_{1 \leqslant i \neq j \leqslant n} \mathbb{E}\left[I_i I_j\right] = n \sum_{i=2}^{n} \mathbb{E}\left[I_1 I_i\right].
$$

The event $I_1 I_i = 1$ happens if and only if $T \prec T_1(\widetilde{\boldsymbol{\xi}}^{\mathbf{n}})$ and $T \prec T_i(\widetilde{\boldsymbol{\xi}}^{\mathbf{n}})$ both happen. Thus instead of summing $\mathbb{E}[I_1 I_i]$ over i, we can sum $\mathbb{P}\left\{\widetilde{\boldsymbol{\xi}}^{\mathbf{n}} = \boldsymbol{d}\right\}$ over pairs $(i, \boldsymbol{d}) \in \{2, \ldots, n\} \times \widetilde{\mathcal{D}}_n$ that satisfy $T \prec T_1(\boldsymbol{d})$ and $T \prec T_i(\boldsymbol{d})$, i.e.,

$$
\deg_1(\boldsymbol{d}) = (\boldsymbol{a}_1, \boldsymbol{b}_1, \boldsymbol{a}_2, \boldsymbol{b}_2, \ldots, \boldsymbol{a}_\ell, \boldsymbol{b}_\ell), \quad \text{and} \quad \deg_i(\boldsymbol{d}) = (\boldsymbol{a}_1, \boldsymbol{b}_1', \boldsymbol{a}_2, \boldsymbol{b}_2', \ldots, \boldsymbol{a}_\ell, \boldsymbol{b}_\ell'),
$$

where $(\boldsymbol{a}_1, 0, \ldots, \boldsymbol{a}_\ell, 0)$ is the preorder degree sequence of T and $\boldsymbol{b}_1, \boldsymbol{b}_1', \ldots, \boldsymbol{b}_\ell, \boldsymbol{b}_\ell'$ are preorder degree sequences of trees. Let $\mathcal{A}$ be the set of such pairs. Then $\sum_{i \geqslant 2} \mathbb{E}[I_1 I_i] = \sum_{(i, \boldsymbol{d}) \in \mathcal{A}} \mathbb{P}\left\{\widetilde{\boldsymbol{\xi}}^{\mathbf{n}} = \boldsymbol{d}\right\}$.

For $1 \leqslant r \leqslant n$, let $\mathcal{I}_r(\boldsymbol{d})$ be the set of positions in $\boldsymbol{d}$ that are occupied by $\deg_r(\boldsymbol{d})$, i.e.,

$$
\mathcal{I}_r(\boldsymbol{d}) := \{j \bmod n : r \leqslant j < r + |\deg_r(\boldsymbol{d})|\}.
$$

Let $\mathcal{I}_1^{\mathbf{in}}(\boldsymbol{d}) \subseteq \mathcal{I}_1(\boldsymbol{d})$ be the set of positions in $\boldsymbol{d}$ that are occupied by the parts of $\deg_1(\boldsymbol{d})$ that correspond to $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_\ell$. Let $\mathcal{I}_1^{\mathbf{out}}(\boldsymbol{d}) = \mathcal{I}_1(\boldsymbol{d}) \backslash \mathcal{I}_1^{\mathbf{in}}(\boldsymbol{d})$. Define $\mathcal{I}_i^{\mathbf{in}}(\boldsymbol{d})$ and $\mathcal{I}_i^{\mathbf{out}}(\boldsymbol{d})$ accordingly. Let $\mathcal{A}' \subseteq \mathcal{A}$ be the set of $(i, \boldsymbol{d})$ in $\mathcal{A}$ such that

$$\mathcal{A}' = \left\{ (i, \boldsymbol{d}) \in \mathcal{A} : \mathcal{I}_1^{\mathbf{in}}(\boldsymbol{d}) \cap \mathcal{I}_i^{\mathbf{in}}(\boldsymbol{d}) = \varnothing \right\}.$$

Let $\mathcal{A}'' := \mathcal{A} \backslash \mathcal{A}'$.

If $(i, \boldsymbol{d}) \in \mathcal{A}''$, then $\mathcal{I}_1^{\mathbf{in}}(\boldsymbol{d}) \cap \mathcal{I}_i^{\mathbf{in}}(\boldsymbol{d}) \neq \varnothing$. In other words, either T_i is fringe subtree of T_1 and T_i is rooted at a node that corresponds to an internal node of T (regarding that T_1 is a non-fringe subtree of the shape T), or vice versa. Thus there exists a $T' \in \{T \boxplus T\}$ such that either $T' \prec T_1(\boldsymbol{d})$ or $T' \prec T_i(\boldsymbol{d})$. By symmetry, we have

$$\sum_{(i,\boldsymbol{d}) \in \mathcal{A}''} \mathbb{P}\left\{ \tilde{\boldsymbol{\xi}}^{\mathbf{n}} = \boldsymbol{d} \right\} = 2 \sum_{T' \in \{T \boxplus T\}} \mathbb{P}\left\{ T' \prec T_1(\tilde{\boldsymbol{\xi}}^{\mathbf{n}}) \right\}$$

$$= 2 \sum_{T' \in \{T \boxplus T\}} \pi^{\mathbf{nf}}(T') \frac{\mathbb{P}\left\{ S_{n-v(T')} = n - v(T') - \ell(T') \right\}}{\mathbb{P}\left\{ S_n = n - 1 \right\}}, \quad (5.4)$$

where the last step follows from Lemma 5.1.

Now consider $(i, (d_1, \ldots, d_n)) \in \mathcal{A}'$. Arrange $(d_1, \ldots, d_n)$ in a cycle. Paint the segment $\deg_1((d_1, \ldots, d_n))$ red and the segment $\deg_i((d_1, \ldots, d_n))$ blue. One of the three cases must be true: (i) $\mathcal{I}_1(\boldsymbol{d}) \cap \mathcal{I}_i(\boldsymbol{d}) = \varnothing$ — The red segment and the blue segment do not overlap. (ii) $\mathcal{I}_i(\boldsymbol{d}) \subseteq \mathcal{I}_1(\boldsymbol{d})$ — The red segment contains the blue segment. (iii) $\mathcal{I}_1(\boldsymbol{d}) \subseteq \mathcal{I}_i(\boldsymbol{d})$ — The blue segment contains the red segment. (Since $\deg_1(\boldsymbol{d})$ and $\deg_i(\boldsymbol{d})$ are both preorder degree sequences, if $\mathcal{I}_1(\boldsymbol{d}) \cap \mathcal{I}_i(\boldsymbol{d}) \neq \varnothing$ then either (ii) or (iii) must happen. And since $i \neq 1$ we cannot have $\mathcal{I}_i(\boldsymbol{d}) = \mathcal{I}_1(\boldsymbol{d})$.) Figure 5.2 gives examples of the three cases.
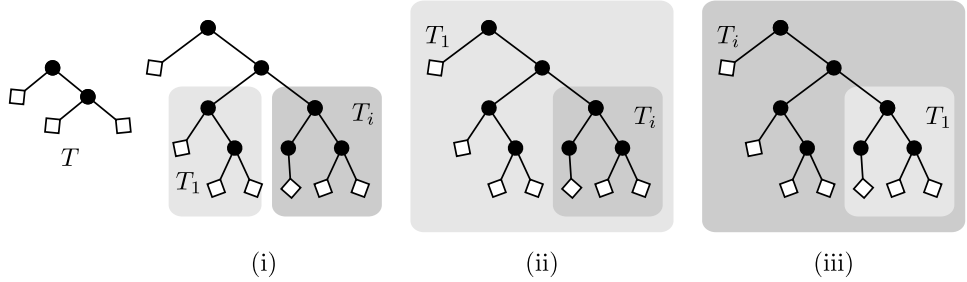

FIGURE 5.2. Examples of three cases in $\mathcal{A}''$.

We permute $(d_1, \ldots, d_n)$ as follows. For (i) and (ii), we first permute the red segment from $(\boldsymbol{a}_1, \boldsymbol{b}_1, \ldots, \boldsymbol{a}_\ell, \boldsymbol{b}_\ell)$ to $(\boldsymbol{a}_1, \ldots, \boldsymbol{a}_\ell, \boldsymbol{b}_1, \ldots, \boldsymbol{b}_\ell)$. Then we permute the blue segment of from $(\boldsymbol{a}_1, \boldsymbol{b}_1', \ldots, \boldsymbol{a}_\ell, \boldsymbol{b}_\ell')$ to $(\boldsymbol{a}_1, \ldots, \boldsymbol{a}_\ell, \boldsymbol{b}_1', \ldots, \boldsymbol{b}_\ell')$. It is clear this can be done in case (i). And it is not difficult to see that in case (ii) the positions that are occupied by the blue segment is completely contained by the positions that are occupied by $b_{\ell'}$ for some $1 \leqslant \ell' \leqslant \ell$. This means that T_i is a fringe subtree of T_1 and the root of T_i does not correspond to an internal node of T (regarding that T_1 is a non-fringe subtree in the shape of T). So the first step of the permutation moves the blue segment but does not change its contents and we can carry out the second step without problem.

In case (iii), we reverse the order of the two steps. After this the starting position of the red segment may have changed. We rotate the new sequence such that the red segment still starts from position 1.

In the end, we get a new pair $(i', (d'_1, \ldots, d'_n))$ such that $(d'_1, d'_2, \ldots, d'_v) = (\boldsymbol{a}_1, \ldots, \boldsymbol{a}_\ell)$ and $(d'_{i'}, d'_{i'+1}, \ldots, d'_{i'+v-1}) = (\boldsymbol{a}_1, \ldots, \boldsymbol{a}_\ell)$. Let $\mathcal{B} \subseteq \{v+1, \ldots, n\} \times \widetilde{\mathcal{D}}_n$ be the set of such pairs. The above permutation defines a mapping $f : \mathcal{A}' \to \mathcal{B}$. Since given $(i', \boldsymbol{d}') \in \mathcal{B}$, we can without ambiguity recover the red segment and blue segment of $\boldsymbol{d}'$, the mapping is reversible, i.e., f is one-to-one. Since $\widetilde{\boldsymbol{\xi}}^{\mathbf{n}}$ is permutation invariant, if $(i', \boldsymbol{d}') = f(i, \boldsymbol{d})$, then $\mathbb{P}\left\{\widetilde{\boldsymbol{\xi}}^{\mathbf{n}} = \boldsymbol{d}\right\} = \mathbb{P}\left\{\widetilde{\boldsymbol{\xi}}^{\mathbf{n}} = \boldsymbol{d}'\right\}$. Therefore

$$\sum_{(i,\boldsymbol{d}) \in \mathcal{A}'} \mathbb{P}\left\{\widetilde{\boldsymbol{\xi}}^{\mathbf{n}} = \boldsymbol{d}\right\} = \sum_{(i,\boldsymbol{d}) \in \mathcal{B}} \mathbb{P}\left\{\widetilde{\boldsymbol{\xi}}^{\mathbf{n}} = \boldsymbol{d}\right\}.$$

Given $(i, (d_1, \ldots, d_n)) \in \mathcal{B}$, we can move the segment $(d_i, \ldots, d_{i+v-1})$ to the position $v+1$ to get a new sequence $(d'_1, \ldots, d'_n) \in \mathcal{C}$, where

$$\mathcal{C} := \left\{(e_1, \ldots, e_n) \in \widetilde{\mathcal{D}}_n : (e_1, \ldots, e_{2v}) = (\boldsymbol{a}_1, \ldots, \boldsymbol{a}_\ell, \boldsymbol{a}_1, \ldots, \boldsymbol{a}_\ell)\right\}.$$

Since there are $n-2v+1$ possible values of i, this permutation gives us a $(n-2v+1)$-to-one mapping $h : \mathcal{B} \to \mathcal{C}$ and if $\boldsymbol{d}' = h(i, \boldsymbol{d})$, then $\mathbb{P}\left\{\widetilde{\boldsymbol{\xi}}^{\mathbf{n}} = \boldsymbol{d}\right\} = \mathbb{P}\left\{\widetilde{\boldsymbol{\xi}}^{\mathbf{n}} = \boldsymbol{d}'\right\}$.

We obtain as usual

$$\sum_{\boldsymbol{d} \in \mathcal{C}} \mathbb{P}\left\{\widetilde{\boldsymbol{\xi}}^{\mathbf{n}} = \boldsymbol{d}\right\} = \mathbb{P}\left\{(\widetilde{\xi}^{\mathbf{n}}_1, \ldots, \widetilde{\xi}^{\mathbf{n}}_{2v}) = (\boldsymbol{a}_1, \ldots, \boldsymbol{a}_\ell, \boldsymbol{a}_1, \ldots, \boldsymbol{a}_\ell)\right\}$$

$$= \mathbb{P}\left\{(\xi_1, \ldots, \xi_{2v}) = (\boldsymbol{a}_1, \ldots, \boldsymbol{a}_\ell, \boldsymbol{a}_1, \ldots, \boldsymbol{a}_\ell) \mid S_n = n-1\right\}$$

$$= \mathbb{P}\left\{(\xi_1, \ldots, \xi_{2v}) = (\boldsymbol{a}_1, \ldots, \boldsymbol{a}_\ell, \boldsymbol{a}_1, \ldots, \boldsymbol{a}_\ell)\right\}$$

$$\times \frac{\mathbb{P}\left\{S_{n-2v} = (n-1) - 2(v+\ell-1)\right\}}{\mathbb{P}\left\{S_n = n-1\right\}}$$

$$= \pi^{\mathbf{nf}}(T)^2 \frac{\mathbb{P}\left\{S_{n-2v} = n+1 - 2(v+\ell)\right\}}{\mathbb{P}\left\{S_n = n-1\right\}}.$$

It follows that

$$\sum_{(i,\boldsymbol{d}) \in \mathcal{A}'} \mathbb{P}\left\{\widetilde{\boldsymbol{\xi}}^{\mathbf{n}} = \boldsymbol{d}\right\} = \sum_{(i',\boldsymbol{d}) \in \mathcal{B}} \mathbb{P}\left\{\widetilde{\boldsymbol{\xi}}^{\mathbf{n}} = \boldsymbol{d}\right\} = (n-2v+1) \sum_{\boldsymbol{d} \in \mathcal{C}} \mathbb{P}\left\{\widetilde{\boldsymbol{\xi}}^{\mathbf{n}} = \boldsymbol{d}\right\}$$

$$= (n-2v+1)\pi^{\mathbf{nf}}(T)^2 \frac{\mathbb{P}\left\{S_{n-2v} = n+1 - 2(v+\ell)\right\}}{\mathbb{P}\left\{S_n = n-1\right\}}. \quad (5.5)$$

The lemma follows by combining (5.4) and (5.5) with the following:

$$\mathbb{E}(N_T^{\mathbf{nf}}(\mathcal{T}_n^{\mathbf{gw}}))_2 = n \sum_{i=2}^{n} \mathbb{E}[I_1 I_i] = n \sum_{(i,\boldsymbol{d}) \in \mathcal{A}} \mathbb{P}\left\{\widetilde{\boldsymbol{\xi}}^{\mathbf{n}} = \boldsymbol{d}\right\}$$

$$= n \sum_{(i,\boldsymbol{d}) \in \mathcal{A}'} \mathbb{P}\left\{\widetilde{\boldsymbol{\xi}}^{\mathbf{n}} = \boldsymbol{d}\right\} + n \sum_{(i,\boldsymbol{d}) \in \mathcal{A}''} \mathbb{P}\left\{\widetilde{\boldsymbol{\xi}}^{\mathbf{n}} = \boldsymbol{d}\right\}. \qquad \square$$

5.2. *Sequence of non-fringe subtrees.* Let T_n be a sequence of trees. Let $v_n := v(T_n)$ and $\ell_n := \ell(T_n)$. In this subsection we prove Theorem 1.6, the concentration of $N_{T_n}^{\mathbf{nf}}(\mathcal{T}_n^{\mathbf{gw}})$.

Lemma 5.3. *Assume Condition* A. *If* $|T_n| = o(n)$, *then*

$$\frac{\mathbb{E}\left[N_{T_n}^{\mathbf{nf}}(\mathcal{T}_n^{\mathbf{gw}})\right]}{n\pi^{\mathbf{nf}}(T_n)} \to 1.$$

Proof: $|T_n| = o(n)$ implies that $v_n = o(n)$ and $\ell_n = o(n)$. Therefore it follows Lemma 2.6 and 5.1 that

$$\frac{\mathbb{E}\left[N_{T_n}^{\mathbf{nf}}(\mathcal{T}_n^{\mathbf{gw}})\right]}{n\pi^{\mathbf{nf}}(T_n)} = \frac{\mathbb{P}\left\{S_{n-v_n} = n - v_n - \ell_n\right\}}{\mathbb{P}\left\{S_n = n - 1\right\}} \to 1. \qquad \square$$

Lemma 5.4. *Assume Condition* A. *If* $|T_n| = o(n)$ *and* $n\pi^{\mathbf{nf}}(T_n) \to \infty$, *then*

$$\frac{\mathbb{E}(N_{T_n}^{\mathbf{nf}}(\mathcal{T}_n^{\mathbf{gw}}))_2}{(n\pi^{\mathbf{nf}}(T_n))^2} \to 1.$$

Proof: Let $v = v_n$, $\ell = \ell_n$ and $T = T_n$. Since $|T_n| = v + \ell$, we have $v = o(n)$ and $\ell = o(n)$. If $T' \in \{T \boxplus T\}$, then $v(T') < 2v = o(n)$ and $\ell(T') < 2\ell = o(n)$. Therefore, it follows from Lemma 2.6 and 5.2 that

$$\mathbb{E}(N_T^{\mathbf{nf}}(\mathcal{T}_n^{\mathbf{gw}}))_2 = n(n - 2v + 1)\pi^{\mathbf{nf}}(T)^2 \frac{\mathbb{P}\left\{S_{n-2v} = n + 1 - 2(v + \ell)\right\}}{\mathbb{P}\left\{S_n = n - 1\right\}}$$

$$+ 2n \sum_{T' \in \{T \boxplus T\}} \pi^{\mathbf{nf}}(T') \frac{\mathbb{P}\left\{S_{n-v(T')} = n - v(T') - \ell(T')\right\}}{\mathbb{P}\left\{S_n = n - 1\right\}}$$

$$= (1 + o(1))(n\pi^{\mathbf{nf}}(T))^2 + O(n) \sum_{T' \in \{T \boxplus T\}} \pi^{\mathbf{nf}}(T').$$

Thus it suffices to show that $n\sum_{T' \in \{T \boxplus T\}} \pi^{\mathbf{nf}}(T') = o(n\pi^{\mathbf{nf}}(T))^2$.

Consider the superset $\mathcal{A}$ of $\{T \boxplus T\}$ that contains trees which can be obtained by replacing a proper non-leaf subtree of T with another copy of T. (We do not restrict where this replacement can happen as in the definition of $\{T \boxplus T\}$.) Note that $|\mathcal{A}| = v - 1$, since T has v internal nodes and one of them is the root.

If $T' \in \mathcal{A}$, then T' contains T as a fringe subtree. Thus $\pi^{\mathbf{nf}}(T') \leqslant \pi^{\mathbf{nf}}(T)$. In the case that v is bounded, we have

$$n \sum_{T' \in \mathcal{A}} \mathbb{P}\left\{T' < \mathcal{T}^{\mathbf{gw}}\right\} \leqslant nv\pi^{\mathbf{nf}}(T) = O(n\pi^{\mathbf{nf}}(T)) = o(n\pi^{\mathbf{nf}}(T))^2.$$

Thus we can assume that $v \to \infty$.

For $T' \in \mathcal{A}$, if T' has at least $3v/2$ internal nodes, call T' *big*, otherwise call it *small*. Let $\mathcal{A}_{\mathbf{b}}$ and $\mathcal{A}_{\mathbf{s}}$ be the sets of big and small trees in $\mathcal{A}$ respectively.

If $T' \in \mathcal{A}_{\mathbf{b}}$, then besides internal nodes that correspond to internal nodes of T, T' contains at least $v/2$ extra internal nodes. So we have $\mathbb{P}\left\{T' < \mathcal{T}^{\mathbf{gw}}\right\} \leqslant \pi^{\mathbf{nf}}(T)p_{\mathbf{max}}^{v/2}$. Since $v \to \infty$ and $p_{\mathbf{max}} < 1$, $vp_{\mathbf{max}}^{v/2} = o(1)$. Using that $|\mathcal{A}| < v$, we have

$$n \sum_{T' \in \mathcal{A}_{\mathbf{b}}} \mathbb{P}\left\{T' < \mathcal{T}^{\mathbf{gw}}\right\} \leqslant nv\pi^{\mathbf{nf}}(T)p_{\mathbf{max}}^{v/2} = o(n\pi^{\mathbf{nf}}(T)).$$

Let $T_{i,j}$ be a fringe subtree in T whose root is at depth i and is the j-th node of this level. If replacing $T_{i,j}$ with a copy of T makes a new tree T'_i that has strictly less than $3v/2$ internal nodes, then $T_{i,j}$ must contain more than $v/2$ internal nodes. Therefore, for each i, there is at most one possible such j. For an example of T'_i, see Figure 5.3.
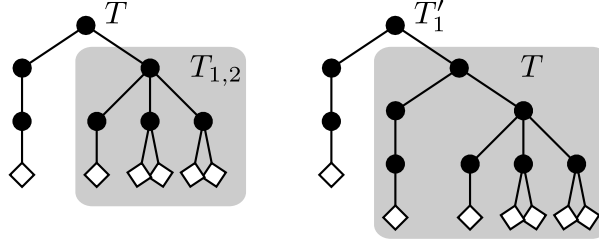
FIGURE 5.3. An example of T_1' for T with 7 internal nodes.

As T has v internal nodes, there are at most $v-1$ possible i that can make $T_{i,j}$ a proper and non-leaf subtree. Since T_i' has at least i internal nodes besides these in the copy of T that replaced $T_{i,j}$, we have $\pi^{\mathbf{nf}}(T_i') \leqslant \pi^{\mathbf{nf}}(T)p_{\mathbf{max}}^i$. In summary, we have

$$n \sum_{T' \in \mathcal{A}_{\mathbf{s}}} \pi^{\mathbf{nf}}(T') \leqslant n \sum_{i=1}^{v} \pi^{\mathbf{nf}}(T_i') \leqslant n \sum_{i=1}^{v} \pi^{\mathbf{nf}}(T)p_{\mathbf{max}}^i \leqslant O(n\pi^{\mathbf{nf}}(T)).$$

Therefore,

$$n \sum_{T' \in \{T \boxplus T\}} \mathbb{P}\left\{T' < \mathcal{T}^{\mathbf{gw}}\right\} \leqslant n \sum_{T' \in \mathcal{A}_{\mathbf{b}}} \mathbb{P}\left\{T' < \mathcal{T}^{\mathbf{gw}}\right\} + n \sum_{T' \in \mathcal{A}_{\mathbf{s}}} \mathbb{P}\left\{T' < \mathcal{T}^{\mathbf{gw}}\right\}$$
$$= o(n\pi^{\mathbf{nf}}(T)) + O(n\pi^{\mathbf{nf}}(T)) = o(n\pi^{\mathbf{nf}}(T))^2. \qquad \square$$

The condition that $n\pi^{\mathbf{nf}}(T_n) \to \infty$ is necessary, as shown by the following lemma.

Lemma 5.5. *Assume Condition A. Let $L_{h(n)}$ be a chain (complete 1-ary tree) of height $h(n)$. Let $X_n := N_{L_{h(n)}}^{\mathbf{nf}}(\mathcal{T}_n^{\mathbf{gw}})$. If $n\mathbb{P}\left\{L_{h(n)} < \mathcal{T}^{\mathbf{gw}}\right\} \to \mu \in (0,\infty)$ as $n \to \infty$, then*

$$\mathbb{E}X_n \to \mu, \qquad \mathrm{Var}\,(X_n) \to \mu\frac{1+p_1}{1-p_1}, \qquad \mathbb{E}\left[(X_n - \mathbb{E}X_n)^3\right] \to \mu\frac{3p_1^2 + 2p_1 + 1}{(1-p_1)^2}.$$

As a result, $\liminf_{n \to \infty} d_{\mathrm{TV}}\,(X_n, \mathrm{Po}(\mu)) > 0$.

Proof: Let $h = h(n)$. Since $n\mathbb{P}\{L_h < \mathcal{T}^{\mathbf{gw}}\} = np_1^h \to \mu \in (0,\infty)$, we have $h = \log_{1/p_1} n + O(1)$. L_h has h internal nodes and one leaf. Thus it follows from Lemma 5.1 and 2.6 that

$$\mathbb{E}X_n = n\pi^{\mathbf{nf}}(L_h)\frac{\mathbb{P}\{S_{n-h} = n-h-1\}}{\mathbb{P}\{S_n = n-1\}} \to \mu.$$

Since $\{L_h \oplus L_h\} = \{L_{h+i} : 1 \leqslant i \leqslant h-1\}$, by Lemma 2.6,

$$\zeta_1 := 2n \sum_{T' \in \{L_h \oplus L_h\}} \pi^{\mathbf{nf}}(T')\frac{\mathbb{P}\{S_{n-v(T')} = n-v(T')-\ell(T')\}}{\mathbb{P}\{S_n = n-1\}}$$
$$= 2n \sum_{i=1}^{h-1} \pi^{\mathbf{nf}}(L_{h+i})\frac{\mathbb{P}\{S_{n-h-i} = n-h-i-1\}}{\mathbb{P}\{S_n = n-1\}}$$
$$= (1 + o(1))2n \sum_{i=1}^{h-1} p_1^{h+i} = (1+o(1))2np_1^h \sum_{i=1}^{h-1} p_1^i \to 2\mu\frac{p_1}{1-p_1}.$$

We also have by Lemma 2.6,

$$\zeta_2 := n(n-2h)\pi^{\mathbf{nf}}(L_h)^2 \frac{\mathbb{P}\left\{S_{n-2h} = n+1-2(h+1)\right\}}{\mathbb{P}\left\{S_n = n-1\right\}} \to \mu^2.$$

Therefore, it follows from Lemma 5.2 that

$$\mathbb{E}(X_n)_2 = \zeta_1 + \zeta_2 \to 2\mu\frac{p_1}{1-p_1} + \mu^2.$$

Thus

$$\text{Var}\left(X_n\right) = \mathbb{E}\left[(X_n)_2\right] + \mathbb{E}\left[X_n\right] - \mathbb{E}\left[X_n\right]^2 \to \mu\frac{1+p_1}{1-p_1}.$$

So we have

$$\frac{\text{Var}\left(X_n\right)}{\mathbb{E}\left[X_n\right]} \to \frac{1+p_1}{1-p_1} > 1.$$

With an argument similar to Lemma 5.2, we can compute $\mathbb{E}\left[(X_n)_3\right]$, which yields

$$\mathbb{E}\left[(X_n - \mathbb{E}X_n)^3\right] \to \mu\frac{3p_1^2 + 2p_1 + 1}{(1-p_1)^2}.$$

Since

$$\mathbb{E}\left[|X_n - \mathbb{E}X_n|^3\right] = 2\mathbb{E}\left[|X_n - \mathbb{E}X_n|^3 \times [\![X_n < \mathbb{E}X_n]\!]\right] + \mathbb{E}\left[(X_n - \mathbb{E}X_n)^3\right]$$
$$\leqslant 2(\mathbb{E}X_n)^3 + \mathbb{E}\left[(X_n - \mathbb{E}X_n)^3\right],$$

the above limit implies that

$$C := \limsup_{n\to\infty} \mathbb{E}\left[|X_n - \mathbb{E}X_n|^3\right] < \infty.$$

Now we can finish by following the method of Barbour et al. (1992, thm. 3B). Let $Z_n \overset{\mathcal{L}}{=} \text{Po}(\mathbb{E}X_n)$ be a coupling of X_n that minimizes $\mathbb{P}\left\{Z_n \neq X_n\right\}$. Therefore we have $d_{\text{TV}}\left(X_n, \text{Po}(\mathbb{E}X_n)\right) = \mathbb{P}\left\{Z_n \neq X_n\right\}$. Thus

$$\begin{aligned}
\text{Var}\left(X_n\right) - \mathbb{E}\left[X_n\right] &= \mathbb{E}\left[(X_n - \mathbb{E}X_n)^2\right] - \mathbb{E}\left[(Z_n - \mathbb{E}X_n)^2\right] \\
&= \mathbb{E}\left[\left[(X_n - \mathbb{E}X_n)^2 - (Z_n - \mathbb{E}X_n)^2\right] \times [\![X_n \neq Z_n]\!]\right] \\
&\leqslant \mathbb{E}\left[(X_n - \mathbb{E}X_n)^2 \times [\![X_n \neq Z_n]\!]\right] \\
&\leqslant P(X_n \neq Z_n)^{1/3}(\mathbb{E}\left[|X_n - \mathbb{E}X_n|^3\right])^{2/3},
\end{aligned}$$

where in the last step we use Hölder's inequality (Gut, 2013, pp. 129). So

$$d_{\text{TV}}\left(X_n, \text{Po}(\mathbb{E}X_n)\right) = \mathbb{P}\left\{X_n \neq Z_n\right\} \geqslant \left(\frac{\text{Var}\left(X_n\right) - \mathbb{E}X_n}{E(|X_n - \mathbb{E}X_n|^3))^{2/3}}\right)^3.$$

Therefore

$$\liminf_{n\to\infty} d_{\text{TV}}\left(X_n, \text{Po}(\mathbb{E}X_n)\right) \geqslant \frac{1}{C^2}\left(\frac{1+p_1}{1-p_1}\right)^3 > 0.$$

Since $\mathbb{E}X_n \to \mu$, we also have $\liminf_{n\to\infty} d_{\text{TV}}\left(X_n, \text{Po}(\mu)\right) > 0.$ $\qquad\square$

Remark 5.6. For the Possion distribution to be a good approximation of a sum of indicator random variables, it is necessary that these indicators are almost independent. But for $L_{h(n)}$, this is not true—if we find a chain of length $h(n)$ at position i, then with probability p_1 we will find another such chain at position $i+1$. This explains why Poisson approximation fails in this case.

Proof of Theorem 1.6: (i): Since $|T_n| = v_n + \ell_n$, we have $v_n = o(n)$ and $\ell = o(n)$. So by Lemma 5.3, $\mathbb{E}N_{T_n}^{\mathbf{nf}}(\mathcal{T}_n^{\mathbf{gw}}) \sim n\pi^{\mathbf{nf}}(T_n)$. Thus $n\pi^{\mathbf{nf}}(T_n) \to 0$ implies $N_{T_n}^{\mathbf{nf}}(\mathcal{T}_n^{\mathbf{gw}}) \xrightarrow{p} 0$.

(ii): It follows from Lemma 5.3 and Lemma 5.4 that

$$\begin{aligned}
\mathrm{Var}\left(N_{T_n}^{\mathbf{nf}}(\mathcal{T}_n^{\mathbf{gw}})\right) &= \mathbb{E}(N_{T_n}^{\mathbf{nf}}(\mathcal{T}_n^{\mathbf{gw}}))_2 + \mathbb{E}N_{T_n}^{\mathbf{nf}}(\mathcal{T}_n^{\mathbf{gw}}) - (\mathbb{E}N_{T_n}^{\mathbf{nf}}(\mathcal{T}_n^{\mathbf{gw}}))^2 \\
&= (1 + o(1))(n\pi^{\mathbf{nf}}(T_n))^2 + (1 + o(1))(n\pi^{\mathbf{nf}}(T_n)) \\
&\quad - (1 + o(1))(n\pi^{\mathbf{nf}}(T_n))^2 \\
&= o(n\pi^{\mathbf{nf}}(T_n))^2 = o(\mathbb{E}N_{T_n}^{\mathbf{nf}}(\mathcal{T}_n^{\mathbf{gw}}))^2.
\end{aligned}$$

Thus $N_{T_n}^{\mathbf{nf}}(\mathcal{T}_n^{\mathbf{gw}})/(n\pi^{\mathbf{nf}}(T_n)) \xrightarrow{p} 1$. $\qquad\square$

5.3. *Complete r-ary non-fringe subtrees.* Theorem 1.6 allows us to find the maximal complete r-ary non-fringe subtree in $\mathcal{T}_n^{\mathbf{gw}}$. We omit the proofs of the following results due to their similarities to Lemma 4.2 and 4.3:

Lemma 5.7. *Assume Condition A and let $p_r > 0$ for some $r \geqslant 2$. Let $\bar{H}_{n,r}$ be the height of the maximal complete r-ary non-fringe subtree in $\mathcal{T}_n^{\mathbf{gw}}$. Then as $n \to \infty$,*

$$\frac{\bar{H}_{n,r}}{\log_r(\log n)} \xrightarrow{p} 1.$$

Lemma 5.8. *Assume Condition A and let $p_1 > 0$. Let $\bar{H}_{n,1}$ be the height of the maximal chain (complete 1-ary) non-fringe subtree in $\mathcal{T}_n^{\mathbf{gw}}$. Then as $n \to \infty$,*

$$\frac{\bar{H}_{n,1}}{\log_{1/p_1} n} \xrightarrow{p} 1.$$

Example 5.9 (The binary tree). Recall that when $p_0 = p_2 = 1/4$ and $p_1 = 1/2$, $\mathcal{T}_n^{\mathbf{gw}}$ is equivalent to a uniform random binary tree of size n. It follows from Lemma 5.8 that $\bar{H}_{n,1}/\log_2 n \xrightarrow{p} 1$. This result was previously proved by Devroye et al. (1999).

6. Open questions

Part (iv) of Theorem 1.4 shows that $d_{TV}\left(N_{\mathcal{A}_n}(\mathcal{T}_n^{\mathbf{gw}}), \mathrm{Po}(n\pi(\mathcal{A}_n))\right) = o(1)$, given that $\pi(\mathcal{A}_n)/\pi(\mathfrak{T}_{k_n}) \to 0$. We believe this condition may not be necessary. However, to prove it seems to require a very different method.

Theorem 1.6 shows that if $n\pi^{\mathbf{nf}}(T_n) \to \infty$, then $N_{T_n}^{\mathbf{nf}}(\mathcal{T}_n^{\mathbf{gw}})/n\pi^{\mathbf{nf}}(T_n) \xrightarrow{p} 1$. We believe that it is also true that $(N_{T_n}^{\mathbf{nf}}(\mathcal{T}_n^{\mathbf{gw}}) - n\pi^{\mathbf{nf}}(T_n))/\sqrt{n\pi^{\mathbf{nf}}(T_n)}$ converges in distribution to a standard normal distribution. (Janson, 2016, thm. 1.9 has shown that this is indeed the case when $v(n)$ is bounded.) As shown in Section 5.2, the overlapping of subtrees does not affect the second moment of non-fringe subtree counts much. Thus we may be able to compute high moments in similar ways and apply Lemma 3.6 to prove a central limit theorem.

Theorem 1.4 generalizes Theorem 1.3 by considering the number of fringe subtrees whose shapes belong to a set of trees $\mathfrak{T}_{k_n}$ instead of being a single tree T_n. It may be possible to generalize Theorem 1.6 in similar way, i.e., we consider the non-fringe subtrees whose shapes belong to a set of trees $\mathfrak{T}_{k_n}$ instead of being a single tree T_n.

Another problem may be of interest is to get a non-fringe version of Theorem 4.5, i.e., what are the sufficient conditions for all (or not all) trees of size at most k to appear in $\mathcal{T}_n^{\mathbf{gw}}$ as non-fringe subtrees.

Let T be a tree and v be a node of T. Recall that T_v denotes the fringe subtree rooted at v. If by removing some or none the subtrees of T_v, we can make it isomorphic to another tree T', then we say that T contains an *embedded subtree* of the shape T' at v. A more challenging open question is to determine the size of the maximum complete r-ary embedded subtree in large conditional Galton-Watson trees.

Acknowledgements

References

D. Aldous. Asymptotic fringe distributions for general families of random trees. *Ann. Appl. Probab.* **1** (2), 228–266 (1991). MR1102319.

N. Alon and J. H. Spencer. *The probabilistic method.* Wiley-Interscience Series in Discrete Mathematics and Optimization. John Wiley & Sons, Inc., Hoboken, NJ, third edition (2008). ISBN 978-0-470-17020-5. MR2437651.

A. D. Barbour, L. Holst and S. Janson. *Poisson approximation*, volume 2 of *Oxford Studies in Probability*. The Clarendon Press, Oxford University Press, New York (1992). ISBN 0-19-852235-5. MR1163825.

F. Chyzak, M. Drmota, T. Klausner and G. Kok. The distribution of patterns in random trees. *Combin. Probab. Comput.* **17** (1), 21–59 (2008). MR2376422.

L. Devroye. Limit laws for local counters in random binary search trees. *Random Structures Algorithms* **2** (3), 303–315 (1991). MR1109697.

L. Devroye. Limit laws for sums of functions of subtrees of random binary search trees. *SIAM J. Comput.* **32** (1), 152–171 (2002/03). MR1954858.

L. Devroye, P. Flajolet, F. Hurtado, M. Noy and W. Steiger. Properties of random triangulations and trees. *Discrete Comput. Geom.* **22** (1), 105–117 (1999). MR1692686.

M. Drmota. *Random trees.* SpringerWienNewYork, Vienna (2009). ISBN 978-3-211-75355-2. MR2484382.

M. Dwass. The total progeny in a branching process and a related random walk. *J. Appl. Probability* **6**, 682–686 (1969). MR0253433.

P. Erdős and A. Rényi. On a classical problem of probability theory. *Magyar Tud. Akad. Mat. Kutató Int. Közl.* **6**, 215–220 (1961). MR0150807.

Q. Feng, H. M. Mahmoud and A. Panholzer. Phase changes in subtree varieties in random recursive and binary search trees. *SIAM J. Discrete Math.* **22** (1), 160–184 (2008). MR2383234.

P. Flajolet, D. Gardy and L. Thimonier. Birthday paradox, coupon collectors, caching algorithms and self-organizing search. *Discrete Appl. Math.* **39** (3), 207–229 (1992). MR1189469.

P. Flajolet, X. Gourdon and C. Martínez. Patterns in random binary search trees. *Random Structures Algorithms* **11** (3), 223–244 (1997). MR1609509.

P. Flajolet and R. Sedgewick. *Analytic combinatorics.* Cambridge University Press, Cambridge (2009). ISBN 978-0-521-89806-5. MR2483235.

M. Fuchs. Subtree sizes in recursive trees and binary search trees: Berry-Esseen bounds and Poisson approximations. *Combin. Probab. Comput.* **17** (5), 661–680 (2008). MR2454562.

Z. Gao and N. C. Wormald. Asymptotic normality determined by high moments, and submap counts of random maps. *Probab. Theory Related Fields* **130** (3), 368–376 (2004). MR2095934.

A. Gut. *Probability: a graduate course.* Springer Texts in Statistics. Springer, New York, second edition (2013). ISBN 978-1-4614-4707-8; 978-1-4614-4708-5. MR2977961.

R. van der Hofstad. *Random graphs and complex networks. Vol. 1.* Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge (2017). ISBN 978-1-107-17287-6. MR3617364.

C. Holmgren and S. Janson. Limit laws for functions of fringe trees for binary search trees and random recursive trees. *Electron. J. Probab.* **20**, no. 4, 51 (2015). MR3311217.

S. Janson. Simply generated trees, conditioned Galton-Watson trees, random allocations and condensation. *Probab. Surv.* **9**, 103–252 (2012). MR2908619.

S. Janson. Asymptotic normality of fringe subtrees and additive functionals in conditioned Galton-Watson trees. *Random Structures Algorithms* **48** (1), 57–101 (2016). MR3432572.

V. F. Kolchin. *Random mappings.* Translation Series in Mathematics and Engineering. Optimization Software, Inc., Publications Division, New York (1986). ISBN 0-911575-16-2. MR865130.

P. Neal. The generalised coupon collector problem. *J. Appl. Probab.* **45** (3), 621–629 (2008). MR2455173.

R. Otter. The number of trees. *Ann. of Math. (2)* **49**, 583–599 (1948). MR0025715.

J. Pitman. Enumerations of trees and forests related to branching processes and random walks. In *Microsurveys in discrete probability (Princeton, NJ, 1997)*, volume 41 of *DIMACS Ser. Discrete Math. Theoret. Comput. Sci.*, pages 163–180. Amer. Math. Soc., Providence, RI (1998). MR1630413.

B. Roos. Improvements in the Poisson approximation of mixed Poisson distributions. *J. Statist. Plann. Inference* **113** (2), 467–483 (2003). MR1965122.