Depth of vertices with high degree in random recursive trees

Laura Eslava

Universidad Nacional Autonoma Mexico,
Instituto de Investigaciones en Matemáticas Aplicadas y en Sistemas
CDMX, 04510, Mexico.
E-mail address: `laura@sigma.iimas.unam.mx`
URL: http://sigma.iimas.unam.mx/laura

Abstract. Let T_n be a random recursive tree with n nodes. List vertices of T_n according to a decreasing order of their degrees as $v^{(1)}, \ldots, v^{(n)}$, and write $\deg(v^{(i)})$ and $h(v^{(i)})$ for the degree of $v^{(i)}$ and the distance of $v^{(i)}$ from the root, respectively. We prove that, as $n \to \infty$ along suitable subsequences,

$$\left(\deg(v^{(i)}) - \lfloor \log_2 n \rfloor, \frac{h(v^{(i)}) - \mu \log_e n}{\sqrt{\sigma^2 \log_e n}} \right) \to \left((P_i, i \geq 1), (N_i, i \geq 1) \right),$$

where $\mu = 1 - (\log_2 e)/2$, $\sigma^2 = 1 - (\log_2 e)/4$, $(P_i, i \geq 1)$ is a Poisson point process on $\mathbb{Z}$ and $(N_i, i \geq 1)$ is a sequence of independent standard Gaussians. We additionally establish joint normality for the depths of finitely many uniformly random vertices in T_n, which extends results from Devroye in 1988 and Mahmoud in 1991. The joint limit holds even if the random vertices are conditioned to have large degree; in particular, both the mean and variance of the conditional depths remain of orden $\ln n$.

Our results are based on a $n!$-to-1 correspondence between (a representation of) Kingman's n-coalescent and random recursive trees; a utility that was observed in work by Pittel 1994 and recovered by Addario-Berry and the author in 2018.

1. Introduction

Random recursive trees (RRTs) have been widely studied since their introduction, Na and Rapoport (1970), and are closely related to binary search trees, preferential attachment trees and increasing trees in general, see e.g. Bergeron et al. (1992); Drmota (2009). For $n \geq 1$, let $[n] := \{1, \ldots, n\}$ and let $\mathcal{I}_n$ be the class of increasing trees with vertex set $[n]$; that is, rooted,

labelled trees where labels are increasing along paths starting at the root. A random recursive tree on n vertices is uniformly distributed on $\mathcal{I}_n$ and may be regarded as a degenerate case of linear preferential attachment trees. We will be concerned with the interplay between degree and depth of vertices in RRTs and will contrast our results with known results on linear preferential attachment trees Bhamidi (2007); Móri (2005).

For $\alpha \in [0, \infty)$, the linear preferential attachment process $(T_{\alpha,n}; n \geq 1)$ is defined as follows. Let $T_{\alpha,1}$ be a single vertex labeled 1. For $n \in \mathbb{Z}_+$ let $T_{\alpha,n+1}$ be the tree obtained from $T_{\alpha,n}$ by adding an edge from a new vertex labelled $n+1$ to vertex $v_n \in [n]$; where, conditional on $T_{\alpha,n}$, $\mathbb{P}(v_n = v)$ is proportional to $\alpha \deg_{T_{\alpha,n}}(v) + 1$. Let

$$\Delta_{\alpha,n} := \max_{v \in [n]} \deg_{T_{\alpha,n}}(v), \tag{1.1}$$

$$\mathcal{M}_{\alpha,n} := \{v \in [n] : \deg_{T_{\alpha,n}} = \Delta_{\alpha,n}\}. \tag{1.2}$$

Random recursive trees correspond to $\alpha = 0$ and so the choice of each v_n is independent from the past and uniformly random on $[n-1]$. In what follows, we omit the index α from the notation of RRTs.

The maximum degree of a RRT satisfies $\Delta_n / \log n \to 1$ a.s. as $n \to \infty$ Devroye and Lu (1995) and the limiting law of $\Delta_n - \lfloor \log n \rfloor$ has, up to lattice effects, a (discrete) Gumbel distribution. This was first shown using singularity analysis of generating functions in Goh and Schmutz (2002). Addario-Berry and Eslava (2018) provides a probabilistic proof, while Eslava (2021) uses the Chen-Stein method to improve the tail bounds on the limiting distribution. For $\alpha > 0$, it has been proven that $\Delta_{\alpha,n} / n^{1/(2+1/\alpha)}$, converges a.s. and in L_p to a positive, finite random variable with absolutely continuous distribution, Móri (2005). Furthermore, almost surely there is a unique vertex attaining the maximum degree and its label v^* is finite; that is, $\mathcal{M}_{\alpha,n} = \{v^*\}$ and $v^* = O_p(1)$ a.s. Bhamidi (2007). Recently, the case of maximal degree vertices in a wider class of growing networks is analysed in Banerjee and Bhamidi (2021).

In this work we are concerned with the depths of vertices with extremal degree values. We say that v is a *high-degree vertex* if $\deg_{T_n}(v) \geq c \ln n$ for some $c > 0$ and that v is a *near-maximum degree vertex* if $|\deg_{T_n}(v) - \log n| < o(\ln n)$. Addario-Berry and the author describe the number of high-degree vertices in T_n via the sequence $(\deg_{T_n}(v^{(i)}) - \lfloor \log n \rfloor, v^{(i)} \in [n])$, where $(v^{(i)})_{i \in [n]}$ are listed in decreasing order of their degrees; see Addario-Berry and Eslava (2018, Theorem 1.2). They show that, along suitable subsequences, this sequence converges in distribution to a Poisson point process $\mathcal{N}$ in $\mathbb{Z} \cup \{\infty\}$ with $\mathbb{E}[|\mathcal{N} \cap [j, \infty]|] = \Theta(2^{-j})$ for all $j \in \mathbb{Z}$.

The main result of this paper extends such convergence to consider not only the degrees of the vertices but also their depths; see Theorem 1.2. A remarkable consequence of Theorem 1.2, in contrast to the behaviour of preferential attachement trees, is that for RRTs there is a random number of maximum-degree vertices and their labels are increasing with n; see Corollary 1.3. On the other hand, Theorem 1.2 is essentially based on Theorem 1.1 which states that the depths of any finite collection of uniformly chosen vertices, conditional on exceeding any given degree, are asymptotically normal and independent; this extends the known limiting distribution of the depth of a uniformly chosen vertex in T_n Devroye (1988); Mahmoud (1991). In particular, the distributional convergence holds when conditioning the vertices on being high-degree or near-maximum-degree vertices.

The central technique of this paper is the link between RRTs and a representation of Kingman's coalescent developed in Addario-Berry (2015); Addario-Berry and Eslava (2018); Eslava (2021) and it is based on the fact that bijections preserve the uniform measure on finite probability spaces. In this case, there is an $n!$-to-1 mapping between the space on n-coalescents (described as n-chains in Section 2.1) and increasing trees on n vertices. For a general description of Kingman's coalescent, see Berestycki (2009, Chapter 2) and for a comparison between distinct representations of Kingman's coalescent, see Eslava (2021).

In contrast to standard constructions, our representation of Kingman's coalescent is based on the data structure known as union-find trees. In fact, the (standard) binary tree representation of Kingman's coalescent had been previously used to study union-find trees, Devroye (1988), and the connection between the results therein and the height of RRTs is mentioned in Pittel (1994). Although the connection between Kingman's coalescent and random recursive trees had been observed, prior to the works in Addario-Berry and Eslava (2018); Eslava (2021), its utility in studying vertex degrees seems to have gone unremarked.

We underline two aspects of the correspondence between Kingman's coalescent and RRTs. First, it allows us to shift the perspective of vertices arriving at distinct times (and their depths and degrees being dependent on such arrival times) to a perspective where all vertices are exchangeable. Second, informally described, we may associate a random number of coin flips to each vertex in the coalescent, then its degree corresponds to the length of the first streak of heads while its depth corresponds to the total number of tails (see Section 2.1 for more details). The crux of the analysis for our results then relies on controlling the correlations that arise in considering several vertices at once.

Notation. We denote natural logarithms and logarithms with base 2 with $\ln(\cdot)$ and $\log(\cdot)$, respectively. We let $\mathbb{Z}_+ := \{1, 2, \ldots\}$ denote the natural number and use $\mathbb{Z}_{\geq 0} := \{0, 1, \ldots\}$ to include zero. The cumulative distribution function of the standard Gaussian distribution is denoted Φ; when there is no ambiguity, we also use Φ for the measure of the distribution. We write $\mathcal{B}_{\mathbb{R}}$ for the Borel sets of $\mathbb{R}$. For $x \in \mathbb{R}$ and $k \in \mathbb{Z}_+$ let $(x)_k = x(x-1)\cdots(x-k+1)$ and write $(x)_0 = 1$.

We consider labeled rooted trees $T = (V, E)$ where edges are naturally directed towards the root which is denoted $r(T)$. That is, we write $e = uv$ for an edge with tail u and head v. For a vertex $v \in V(T)$, we write $\deg_T(v)$ for the number of edges directed towards v in T and let $h_T(v)$ denote the distance between v and $r(T)$; we call $\deg_T(v)$ and $h_T(v)$ the degree and depth of v respectively. Finally, a forest F is a set of trees whose vertex sets are pairwise disjoint. Denote by $V(F)$ and $E(F)$, respectively, the union of the vertex and edge sets of the trees contained in F and for a vertex $v \in V(F)$, we let $\deg_F(v)$ and $h_F(v)$ be the degree and depth of v in the tree in F that contains v.

1.1. *Statement of results.* Our first result establishes the asymptotic independence of the depth of k distinct vertices uniformly chosen at random. The distribution of their depths is asymptotically normal and independent, even after conditioning on arbitrary lower bounds for the vertices' degrees.

Theorem 1.1. *Let $k \in \mathbb{Z}_+$. Let $(N_i)_{i \in [k]}$ be i.i.d. standard Gaussian variables and let $(v_i)_{i \in [k]}$ be k distinct vertices in T_n chosen uniformly at random. For every $(a_1, \ldots, a_k) \in [0, 2)^k$ and $(b_1, \ldots, b_k) \in \mathbb{Z}^k$, the conditional law of*

$$\left(\frac{h_{T_n}(v_i) - (1 - a_i/2) \ln n}{\sqrt{(1 - a_i/4) \ln n}}, \, i \in [k] \right),$$

given that $\deg_{T_n}(v_i) \geq \lfloor a_i \ln n \rfloor + b_i$ for all $i \in [k]$, converges to the law of $(N_i)_{i \in [k]}$.

The case of Theorem 1.1 with $k = 1$ and $a_1 = b_1 = 0$ is implicitly established in Devroye (1988); Mahmoud (1991) where, $h_{T_n}(n)$, the insertion depth in a RRT is studied.

For the next results, we assume that $(n_\ell)_{\ell \geq 1}$ satisfies $\log n_\ell - \lfloor \log n_\ell \rfloor \to \varepsilon$ as $\ell \to \infty$, for some $\varepsilon \in [0, 1]$. This condition is necessary due to a lattice effect caused by the fact that $(\deg_{T_n}(v))_{v \in [n]}$ are integer valued. Let $\mathcal{P}$ be a Poisson point process on $\mathbb{R}$ with intensity $\lambda(x) = 2^{-x} \ln 2$. Since $\mathcal{P} \cap [0, \infty) \overset{\mathcal{L}}{=} \text{Poi}(1)$ there exists a well defined ordering $\mathcal{P} = \{P_1, P_2, \ldots\}$ such that

$$|\mathcal{P} \cap [P_i, \infty)| = |\mathcal{P} \cap (P_i, \infty)| + 1 = i. \tag{1.3}$$

For $n \geq 1$, list vertices of T_n in decreasing order of degree as $v^{(1)}, \ldots, v^{(n)}$ (breaking ties uniformly at random).

Theorem 1.2. *Let $(N_i)_{i\geq 1}$ be i.i.d. standard Gaussian variables, independent of $\mathcal{P}$ as defined above. For $\varepsilon \in [0,1]$ and any increasing sequence of integers $(n_\ell)_{\ell\geq 1}$ for which $\log n_\ell - \lfloor \log n_\ell \rfloor \to \varepsilon$ as $\ell \to \infty$, then the sequence*

$$\left(\left(\deg_{T_{n_\ell}}(v^{(i)}) - \lfloor \log n_\ell \rfloor, \frac{h_{T_{n_\ell}}(v^{(i)}) - (1 - (\log e)/2)\ln n_\ell}{\sqrt{(1 - (\log e)/4)\ln n_\ell}} \right), i \in [n] \right)$$

converges in law to $((\lfloor P_i + \varepsilon \rfloor, N_i), i \geq 1))$.

For the next corollary of Theorem 1.2, recall the definition of $\mathcal{M}_{n_\ell}$ in (1.2). For $\varepsilon > 0$ let $\mathbf{M}_\varepsilon$ be defined, for each $k \geq 1$, by

$$\mathbb{P}\left(\mathbf{M}_\varepsilon = k \right) := \sum_{m \in \mathbb{Z}} e^{-2^{-m+\varepsilon}} \frac{2^{-(m+1-\varepsilon)k}}{k!}; \tag{1.4}$$

then $\mathbf{M}_\varepsilon$ is a mixture of Poisson random variables conditioned to being positive; see Lemma 5.1. The following result focuses on vertices attaining the maximum degree.

Corollary 1.3. *Let $(N_i)_{i\geq 1}$ be i.i.d. standard Gaussian variables. For $\varepsilon \in [0,1]$ and any increasing sequence of integers $(n_\ell)_{\ell\geq 1}$ for which $\log n_\ell - \lfloor \log n_\ell \rfloor \to \varepsilon$ as $\ell \to \infty$, then*

$$\left(\frac{h_{T_{n_\ell}}(i) - (1 - (\log e)/2)\ln n_\ell}{\sqrt{(1 - (\log e)/4)\ln n_\ell}}, i \in \mathcal{M}_{n_\ell} \right) \xrightarrow{\mathcal{L}} (N_i, 1 \leq i \leq \mathbf{M}_\varepsilon).$$

In other words, the depths of vertices attaining the maximum degree are asymptotically normal and independent. Furthermore, the number $|\mathcal{M}_n|$ of such vertices converges to a mixture of Poisson random variables conditioned to be strictly positive.

1.2. *Paper overview.* In the next section we provide an overview of the proof of the main theorems. Theorem 2.1 below is an equivalent statement to Theorem 1.1 in terms of Kingman's coalescent whose construction is given in Section 2.1; in addition, we lay out the main intermediate results (Proposition 2.2 and Lemmas 2.6–2.8) for the proof of Theorem 2.1. Section 3 is then concerned with proving Lemmas 2.7 and 2.8 while Section 4 contains the proof of Proposition 2.2. The proofs of Theorem 1.2 and Corollary 1.3 are given in Section 5.

2. Outline for the proof of Theorems 1.1 and 1.2

Let $T_n = ([n], E_n)$ be a RRT and let $\sigma : [n] \to [n]$ be an independent, uniformly random permutation. Define $T'_n = ([n], E'_n)$ where

$$E'_n = \{\sigma(u)\sigma(v) : uv \in E_n\}; \tag{2.1}$$

that is, T'_n is obtained from T_n by relabelling its vertices with a uniformly random permutation. By definition of T'_n, Theorem 2.1 below is equivalent to Theorem 1.1; its proof overview is detailed in Section 2.1.

Theorem 2.1. *Let $k \in \mathbb{Z}_+$. Let $(N_i)_{i\in[k]}$ be i.i.d. standard Gaussian variables. For every $(a_1, \ldots, a_k) \in [0,2)^k$ and $(b_1, \ldots, b_k) \in \mathbb{Z}^k$, the conditional law of*

$$\left(\frac{h_{T'_n}(i) - (1 - a_i/2)\ln n}{\sqrt{(1 - a_i/4)\ln n}}, i \in [k] \right),$$

given that $\deg_{T'_n}(i) \geq \lfloor a_i \ln n \rfloor + b_i$ for all $i \in [k]$, converges to the law of $(N_i)_{i\in[k]}$.

The advantage of shifting towards this statement is that T_n' may be analysed through a construction of Kingman's coalescent on n elements Addario-Berry and Eslava (2018); Eslava (2021). Theorem 2.1 is the first intermediate result for Theorem 1.2.

The next proposition is the second intermediate result for Theorem 1.2. For $n \geq 1$, denote

$$\mu_n := \left(1 - \frac{\log e}{2}\right) \ln n \qquad \text{and} \qquad \sigma_n^2 := \left(1 - \frac{\log e}{4}\right) \ln n, \tag{2.2}$$

and for $i \in \mathbb{Z}$ and $B \in \mathcal{B}_\mathbb{R}$ let

$$X_i^{(n)}(B) := \#\{v \in [n] : \deg_{T_n}(v) = \lfloor \log n \rfloor + i, \, (h(v) - \mu_n)/\sigma_n \in B\}, \tag{2.3}$$

$$X_{\geq i}^{(n)}(B) := \#\{v \in [n] : \deg_{T_n}(v) \geq \lfloor \log n \rfloor + i, \, (h(v) - \mu_n)/\sigma_n \in B\}. \tag{2.4}$$

Proposition 2.2. *Fix $K, M \in \mathbb{Z}_+$. Let $(a_k)_{k \in [K]}$ be non-negative integers such that $\sum_{k \in [K]} a_k = M$. Let $(j_k)_{k \in [K]}$ be a non-decreasing sequence of integers with $0 \leq K' = \min\{k : j_{k+1} = j_K\}$ and let $(B_k)_{k \in K}$ be sets in $\mathcal{B}_\mathbb{R}$ satisfying $B_k \cap B_\ell = \emptyset$ whenever $j_k = j_\ell$ and $k \neq \ell$. Then, we have*

$$\mathbb{E}\left[\prod_{k=1}^{K'} \left(X_{j_k}^{(n)}(B_k)\right)_{a_k} \prod_{k=K'+1}^{K} \left(X_{\geq j_K}^{(n)}(B_k)\right)_{a_k}\right]$$

$$= \prod_{k=1}^{K'} \left(2^{-j_k + \varepsilon_n - 1} \Phi(B_k)\right)^{a_k} \prod_{k=K'+1}^{K} \left(2^{-j_K + \varepsilon_n} \Phi(B_k)\right)^{a_k} (1 + o(1)).$$

Remark 2.3. By generalizing the definitions in (2.3) and (2.4) to include indices $i = i(n)$, one may extend the result in Proposition 2.2 to hold for any $(j_k)_{k \in [K]}$ such that $0 \leq j_1 + \log n \leq j_K + \log n < c \ln n$ where $c \in (0, 2)$.

The proof of Proposition 2.2 is given in Section 4. We briefly explain an ingredient that is key in proving Proposition 2.2 which is based on the exchangeability of vertices in the representation of T_n as a Kingman's coalescent. As T_n' is a relabelling of T_n, we have, jointly for all $i \in \mathbb{Z}$ and $j \in \mathbb{Z}_+$,

$$\#\{v \in [n] : \deg_{T_n}(v) = i, \, h_{T_n}(v) = j\} \stackrel{\mathcal{L}}{=} \#\{v \in [n] : \deg_{T_n'}(v) = i, \, h_{T_n'}(v) = j\}.$$

Note that each $X_j^{(n)}(B)$ and $X_{\geq j}^{(n)}(B)$ is a sum of indicator variables. Therefore, the factorial moments in Proposition 2.2 are reduced to a sum of probabilities in terms of the degrees and depths of vertices in T_n' which have the following form:

$$\mathbb{P}\left(\deg_{T_n'}(j) \geq m_j, \, h_{T_n'}(j) \in B_j, \, j \in [k]\right)$$

$$= \mathbb{P}\left(\deg_{T_n'}(j) \geq m_j, \, j \in [k]\right) \mathbb{P}\left(h_{T_n'}(j) \in B_j, \, j \in [k] \mid \deg_{T_n'}(j) \geq m_j, \, j \in [k]\right);$$

where $B_j \in \mathcal{B}_\mathbb{R}$ and $m_j < 2\ln n$ for each $j \in [k]$. The joint tails of finitely many vertices in T_n' have been analyzed in Addario-Berry and Eslava (2018) and we present the relevant result as Proposition 3.3. The law of the depths conditional on $\{\deg_{T_n'}(j) \geq m_j, \, j \in [k]\}$ is approximated using Theorem 2.1.

For Theorem 1.2, we next formulate the result as a statement about convergence of marked point processes. Let $\mathbb{Z}^* := \mathbb{Z} \cup \{\infty\}$. Endow $\mathbb{Z}^*$ with the metric defined by $d(i, j) := |2^{-i} - 2^{-j}|$ and $d(i, \infty) := 2^{-i}$ for $i, j \in \mathbb{Z}$. Recall that $\mathcal{P}$ is a Poisson point process on $\mathbb{R}$ with intensity rate $\lambda(x) = 2^{-x} \ln 2$ and let $(\xi_x)_{x \in \mathcal{P}}$ be an independent collection of i.i.d. standard Gaussian variables.

Let $\varepsilon \in [0, 1]$. We define a ground process $\mathcal{P}^\varepsilon$ on $\mathbb{Z}^*$, and a marked point process $\mathcal{MP}^\varepsilon$ on $\mathbb{Z}^* \times \mathbb{R}$ given by

$$\mathcal{P}^\varepsilon := \sum_{x \in \mathcal{P}} \delta_{\lfloor x + \varepsilon \rfloor}, \qquad\qquad \mathcal{MP}^\varepsilon := \sum_{x \in \mathcal{P}} \delta_{(\lfloor x + \varepsilon \rfloor, \xi_x)}, \tag{2.5}$$

Similarly, for all $n \in \mathbb{Z}_+$ we use T_n to define empirical point processes given by

$$\mathcal{P}^{(n)} := \sum_{v \in [n]} \delta_{\deg_{T_n}(v) - \lfloor \log n \rfloor}, \qquad \mathcal{MP}^{(n)} := \sum_{v \in [n]} \delta_{(\deg_{T_n}(v) - \lfloor \log n \rfloor, (h(v) - \mu_n)/\sigma_n)}. \qquad (2.6)$$

Let $\mathcal{M}_{\mathbb{Z}^*}^{\#}$ and $\mathcal{M}_{\mathbb{Z}^* \times \mathbb{R}}^{\#}$ be the space of boundedly finite measures of $\mathbb{Z}^*$ and $\mathbb{Z}^* \times \mathbb{R}$, respectively. It is clear that $\mathcal{MP}^\varepsilon$ and $\mathcal{MP}^{(n)}$ are elements of $\mathcal{M}_{\mathbb{Z}^* \times \mathbb{R}}^{\#}$ (see, e.g. Daley and Vere-Jones (2003, Definition 9.2.I)). The advantage of working on the state space $\mathbb{Z}^*$ is that intervals $[i, \infty]$ are compact and so the convergence of finite dimensional distributions of $\mathcal{MP}^{(n_\ell)}$ implies, e.g., convergence in distribution of $X_{\geq i}^{(n_\ell)}(B)$ for any $B \in \mathcal{B}_\mathbb{R}$ and $i \in \mathbb{Z}_+$. We note that the weak convergence $\mathcal{P}^{(n_\ell)} \to \mathcal{P}^\varepsilon$, for suitable sequences $(n_\ell)_{\ell \geq 1}$, had already been established in Addario-Berry and Eslava (2018).

Finally, the statement of Theorem 1.2 is equivalent to the weak convergence of $\mathcal{MP}^{(n_\ell)} \to \mathcal{MP}^\varepsilon$ in $\mathcal{M}_{\mathbb{Z}^* \times \mathbb{R}}^{\#}$, for suitable sequences $(n_\ell)_{\ell \geq 1}$. The joint distribution of the variables in Proposition 2.2 are finite dimensional distributions of the marked point process $\mathcal{MP}^{(n_\ell)}$. Therefore, Proposition 2.2 together with the method of moments establish the desired weak convergence. The complete details, together with the proof of Corollary 1.3, are given in Section 5.

2.1. *Kingman's coalescent approach for Theorem 2.1.* For each $n \geq 1$, we consider the set of forests $\mathcal{F}_n := \{F : V(F) = [n]\}$ with vertex set $[n]$. An n-chain is a sequence $(F_n, \ldots, F_1)$ of elements of $\mathcal{F}_n$ if for $1 < i \leq n$, F_{i-1} is obtained from F_i by adding an edge connecting two of the roots in F_i; note that F_n is consequently the forest with n one-vertex trees. We write $F_i = \{T_1^{(i)}, \ldots, T_i^{(i)}\}$, listing the trees according to an increasing order of their smallest-labeled vertex.

We next define the construction of Kingman's n-coalescent as a random n-chain $\mathbf{C} = (F_n, \ldots, F_1)$. Informally, at each step $1 \leq j < n$, two uniformly random trees of F_{n-j+1} are merged by adding an edge between their roots, the direction of the edge is determined independently and with equal probability. For an example of the process see Figure 2.1.

Definition 2.4 (Kingman's n-coalescent)**.** Let $n \in \mathbb{Z}_+$. For each $1 < i \leq n$, choose $\{a_i, b_i\} \subset \{\{a, b\} : 1 \leq a < b \leq i\}$ independently and uniformly at random; also let $(\xi_i)_{1 < i \leq n}$ be a sequence of independent Bernoulli$(1/2)$ random variables.

The n-chain $\mathbf{C} = (F_n, \ldots, F_1)$ is defined as follows. For $1 < i \leq n$, starting from F_i, add an edge e_{i-1} between the roots of $T_{a_i}^{(i)}$ and $T_{b_i}^{(i)}$; direct e_{i-1} towards $r(T_{a_i}^{(i)})$ if $\xi_i = 1$, and towards $r(T_{b_i}^{(i)})$ otherwise. Then F_{i-1} contains the new tree and the remaining $i - 2$ unaltered trees from F_i.

It is not difficult to see that $\mathbf{C}$ has a uniform distribution on the $n!(n-1)!$ possible n-chains representing Kingman's n-coalescent, Addario-Berry and Eslava (2018). On the other hand, recall that T_n has a uniform distribution on the $(n-1)!$ increasing trees in $\mathcal{I}_n$. The next result essentially follows since bijections preserve the uniform measure on finite probability spaces; for details see Addario-Berry and Eslava (2018, Proposition 3.3) and also Eslava (2021, Propositions 1.2 and 1.3).

Proposition 2.5 (Proposition 3.3 in Addario-Berry and Eslava (2018))**.** *Let $n \geq 1$. Let T_n' be defined as in (2.1) and let $T_1^{(1)}$ correspond to the final tree in $\mathbf{C}$ as constructed in Definition 2.4. Then $T_n' \overset{\mathcal{L}}{=} T_1^{(1)}$; that is, for any labelled tree T with vertex set $[n]$ we have $\mathbb{P}(T_n' = T) = \mathbb{P}\left(T_1^{(1)} = T\right)$.*

In the remainder of the paper, we analyse T_n' using the construction in Definition 2.4; that is, we analyse the n-chain $(F_n, \ldots, F_1)$ and identify the final tree $T_1^{(1)} = T_n'$. To ease the notation, write

$$d_n(v) = \deg_{T_n'}(v) \qquad \text{and} \qquad h_n(v) = h_{T_n'}(v), \qquad (2.7)$$

for the degrees and depths of vertices in T_n'.

For each vertex $v \in [n]$ and $1 < i \leq n$, let $T_i(v)$ be the tree in F_i that contains v. Let $(s_{i,v})_{1 < i \leq n}$ and $(h_{i,v})_{1 < i \leq n}$ be sequences where $s_{i,v}$ is the indicator that $T_i(v) \in \{T_{a_i}^{(i)}, T_{b_i}^{(i)}\}$ and $h_{i,v}$ is the
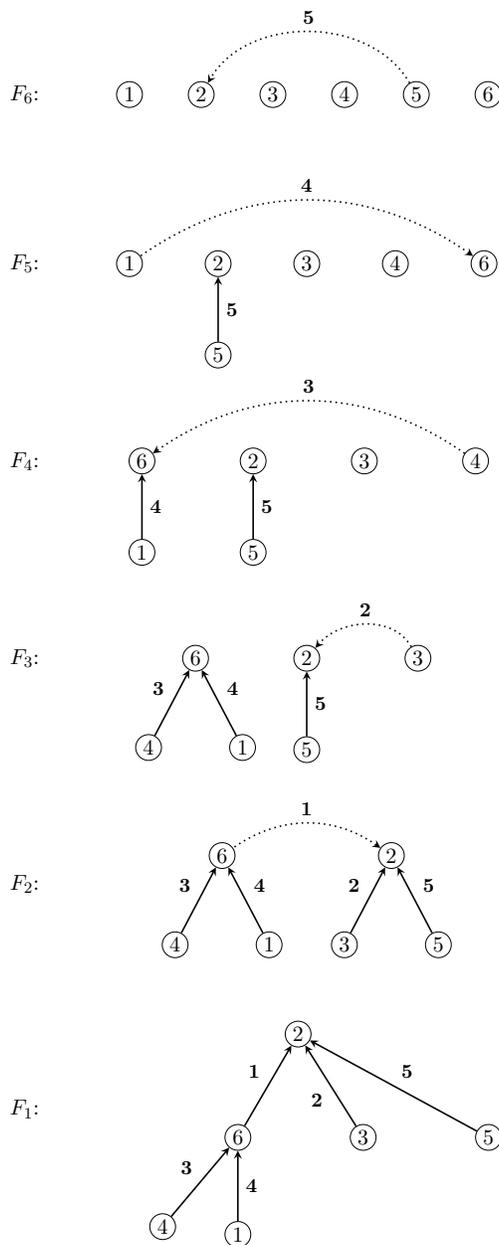
FIGURE 2.1. An example of Kingman's n-coalescent $\mathbf{C} = (F_n, \ldots, F_1)$ for $n = 6$. For $1 < i \leq n$, we represent the edge in $E(F_{i-1}) \setminus E(F_i)$ with a dotted line in F_i. Edges are marked with the index of the first forest where e is present. In this case, $\xi_6 = \xi_4 = \xi_3 = 1$, $\xi_5 = \xi_2 = 0$ and $\{a_6, b_6\} = \{2, 5\}$, $\{a_5, b_5\} = \{1, 5\}$, $\{a_4, b_4\} = \{1, 4\}$, $\{a_3, b_3\} = \{2, 3\}$, $\{a_2, b_2\} = \{1, 2\}$.

indicator that the tail of the edge e_{i-1} is $r(T_i(v))$. Note that $h_{i,v} = 1$ only if $s_{i,v} = 1$. Indeed, if $s_{i,v} = 0$ then, when constructing F_{i-1} from F_i, the trees that are merged do not include $T_i(v)$ and so necessarily $h_{i,v} = 0$.

Now, conditional on $s_{i,v} = 1$, there are two equally likely events regarding the merger of $T_i(v)$ to form F_{i-1}: if e_{i-1} is directed towards $r(T_i(v))$, then the degree of $r(T_i(v))$ increases by one in F_{i-1};

FIGURE 2.2. For $v \in [n]$ and $1 < i \leq n$ let $r_i = r(T_i(v))$ and suppose $i \in \mathcal{S}_n(v)$. Conditional on F_i, if e_i is directed towards r_i, then the degree of r_i increases by one in F_{i-1}. If e_i is directed outwards r_i, then the depth of each $u \in T_i(v)$ increases by one in F_{i-1}.

otherwise, e_{i-1} is directed out of $r(T_i(v))$ and all vertices in $T_i(v)$ increase their depth by one in F_{i-1}; see Figure 2.2.

The choice of trees to be merged at each step is both independent and uniform so, for fixed $v \in [n]$, the variables $(s_{i,v})_{1 < i \leq n}$ are independent Bernoulli random variables with $\mathbb{E}[s_{i,v}] = 2/i$. Moreover, since the variables $(\xi_i)_{i \in [n-1]}$ are independent from the selection of the trees to be merged, the variables $(h_{i,v})_{1 < i \leq n}$ are also independent Bernoulli random variables with $\mathbb{E}[h_{i,v}] = 1/i$.

For $v \in [n]$, let

$$\mathcal{S}_n(v) := \{1 < i \leq n \,:\, s_{i,v} = 1\}, \tag{2.8}$$

$$\mathcal{S}_{n,1}(v) := \{\ln^2 n < i \leq n \,:\, s_{i,v} = 1\}; \tag{2.9}$$

we call these the *selection* and *truncated selection sets* of v and write $S_n(v) := |\mathcal{S}_n(v)|$ and $S_{n,1}(v) := |\mathcal{S}_{n,1}(v)|$, respectively. Using Lindeberg's condition (e.g., see Durrett (1996, Theorem 3.4.5)) we can see that, as $n \to \infty$,

$$\frac{S_n(v) - 2\ln n}{\sqrt{2\ln n}} \xrightarrow{\mathcal{L}} N(0,1) \qquad \text{and} \qquad \frac{S_{n,1}(v) - 2\ln n}{\sqrt{2\ln n}} \xrightarrow{\mathcal{L}} N(0,1). \tag{2.10}$$

For each $1 < i \leq n$, the law of $h_{i,v}$ conditional on $s_{i,v} = 1$ is Bernoulli with mean $1/2$. Since $(s_{i,v})_{1 < i \leq n}$ are independent, we have that

$$h_n(v) = \sum_{i=2}^{n} h_{i,v} \stackrel{\mathcal{L}}{=} \sum_{j=1}^{S_n(v)} I_j, \tag{2.11}$$

where $(I_j)_{j \geq 1}$ are i.i.d. Bernoulli random variables with mean $1/2$ independent of $\mathcal{S}_n(v)$. Using the first expression of $h_n(v)$ in (2.11) and Lindeberg's condition, we have that $h_n(v)$ is asymptotically normal for $v \in [n]$ as $n \to \infty$. We may also establish the limiting distribution of $h_n(v)$ from the last expression in (2.11) and the following lemma; we include its short proof for completeness.

Lemma 2.6. *Let* $a \in [0,2)$, $b \in \mathbb{Z}$ *and set* $m := \lfloor a \ln n \rfloor + b$. *Let* $(Q_n)_{n \geq 1}$ *be integer valued random variables such that the limiting law of* $\frac{Q_n - 2\ln n}{\sqrt{2\ln n}}$ *is a standard Gaussian distribution. Let* N *be a standard Gaussian variable and let* $\hat{H}_n$ *have the law*[1] *of a* $\mathrm{Bin}(Q_n - m, 1/2)$ *conditional on* $|Q_n| \geq m$. *Then,*

$$\frac{\hat{H}_n - (1 - a/2)\ln n}{\sqrt{(1 - a/4)\ln n}} \xrightarrow{\mathcal{L}} N.$$

Proof: Write $\hat{Q}_n$ for the law of $Q_n - m$ conditionally given that $Q_n \geq m$. Since $a < 2$, we have that $\mathbb{P}(Q_n \geq m) \to 1$ as $n \to \infty$ and so the limiting law of $\frac{\hat{Q}_n - (2-a)\ln n}{\sqrt{2\ln n}}$ is a standard Gaussian

---

[1]Abusing notation we let $\mathrm{Bin}(0, p) \equiv 0$.

distribution. Let $(I_1^{(n)}, \ldots, I_n^{(n)})_{n \geq 1}$ be i.i.d. Bernoulli random variables with mean $1/2$ independent of $(\hat{Q}_n)_{n \geq 1}$ and let $(X_n, Y_n)_{n \geq 1}$ be defined as

$$X_n := \frac{2\sum_{j=1}^n I_j^{(n)} - n}{\sqrt{n}} \qquad \text{and} \qquad Y_n := \frac{\hat{Q}_n - (2-a)\ln n}{\sqrt{2\ln n}};$$

by the independence assumption we have that $(X_n, Y_n) \xrightarrow{\mathcal{L}} (N_1, N_2)$ where N_1 and N_2 are independent standard Gaussian variables. By the Skorohod embedding theorem (e.g., see Billingsley (1999, Theorem 6.7)), we may couple $(\hat{Q}_n)_{n \geq 1}$ and $(I_1^{(n)}, \ldots, I_n^{(n)})_{n \geq 1}$ in such a way that $(X_n, Y_n) \to (N_1, N_2)$ almost surely.

Let $\hat{H}_n = \sum_{j=1}^{\hat{Q}_n} I_j^{(\hat{Q}_n)}$ and let $F(x)$ be the cumulative distribution function of $\frac{\hat{H}_n - (1-a/2)\ln n}{\sqrt{(1-a/4)\ln n}}$. We have $\hat{Q}_n / \ln n \to 2 - a$ and, in particular, $\hat{Q}_n \to \infty$, then for $x \in \mathbb{R}$ we have

$$
\begin{aligned}
F(x) =& \mathbb{P}\left(2\sum_{j=1}^{\hat{Q}_n} I_j^{(\hat{Q}_n)} \leq x\sqrt{(4-a)\ln n} + (2-a)\ln n \right) \\
=& \mathbb{P}\left(\frac{2\sum_{j=1}^{\hat{Q}_n} I_j^{(\hat{Q}_n)} - \hat{Q}_n}{\sqrt{\hat{Q}_n}}\sqrt{\frac{\hat{Q}_n}{(4-a)\ln n}} + \frac{\hat{Q}_n - (2-a)\ln n}{\sqrt{2\ln n}}\sqrt{\frac{2}{(4-a)}} \leq x \right) \\
\to& \mathbb{P}\left(\sqrt{\frac{2-a}{4-a}}N_1 + \sqrt{\frac{2}{4-a}}N_2 \leq x \right) = \Phi(x);
\end{aligned}
$$

from which the desired convergence is established. $\qquad\square$

We will see in Section 3 that conditional on $d_n(v) \geq m$, we have $h_n(v) \overset{\mathcal{L}}{=} \mathrm{Bin}(S_n(v) - m, 1/2)$, which together with Lemma 2.6, yields the convergence in law of $h_n(v)$ conditional on $d_n(v) \geq m$ (assuming $2\ln n - m \to \infty$); for the details see (3.1) and the discussion thereafter. Briefly sketched, this is the case $k = 1$ of Theorem 2.1.

For the general case, we have to deal with the correlations between the selection sets of finitely many vertices. For example, let v and w be distinct vertices in T_n' and let $\lambda_{v,w} := \max\{1 < \ell \leq n : \ell \in \mathcal{S}_n(v) \cap \mathcal{S}_n(w)\}$; that is, in $F_{\lambda_{v,w}}$, vertices v and w belong to distinct trees that are merged in $F_{\lambda_{v,w}-1}$. The correlation between $(h_{i,v})_{2 \leq i \leq n}$ and $(h_{i,w})_{2 \leq i \leq n}$ is evident since $h_{v,\lambda_{v,w}} = 1 - h_{w,\lambda_{v,w}}$ while $h_{i,v} = h_{i,w}$ for $i < \lambda_{v,w}$. In words, exactly one of v or w increases its depth in $F_{\lambda_{v,w}-1}$ and, from then on, their depths increase simultaneously for the remainder of the process.

To circumvent this problem, we analyse the coalescent at $F_{\lfloor \ln^2 n \rfloor}$ where, for finitely many vertices, their selection sets –and so their depths– are asymptotically independent (see Lemma 2.7) and the main contribution to their depths in T_n' is already established (see Lemma 2.8).

For each $v \in [n]$, define[2]

$$h_{n,1}(v) := h_{F_{\lfloor \ln^2 n \rfloor}}(v) = \sum_{j \in \mathcal{S}_{n,1}(v)} h_{j,v} \quad \text{and} \quad h_{n,2}(v) := h_n(v) - h_{n,1}(v). \tag{2.12}$$

Lemma 2.7. *Let* $k \in \mathbb{Z}_+$. *Let* $(N_i)_{i \in [k]}$ *be i.i.d. standard Gaussian variables. For every* $(a_1, \ldots, a_k) \in [0, 2)^k$ *and* $(b_1, \ldots, b_k) \in \mathbb{Z}^k$, *the conditional law of*

$$\left(\frac{h_{n,1}(i) - (1 - a_i/2)\ln n}{\sqrt{(1 - a_i/4)\ln n}}, \; i \in [k] \right),$$

---

[2]If we were to define $h_{n,1}(i) = h_{F_j}(i)$ with $j = o(\sqrt{\ln n})$, then the statement of Lemma 2.8 would hold immediately; however, establishing Lemma 2.7 would become a much more delicate task.

given that $d_n(i) \geq \lfloor a_i \ln n \rfloor + b_i$ for all $i \in [k]$, converges to the law of $(N_i)_{i \in [k]}$.

Lemma 2.8. *Fix $k \in \mathbb{Z}_+$ and $c \in (0, 2)$. Let $m_i = m_i(n) < c \ln n$ for all $i \in [k]$. For each $\varepsilon > 0$ and any $j \in [k]$,*

$$\mathbb{P}\left(h_{n,2}(j) \geq \varepsilon \sqrt{\ln n} \,\middle|\, d_n(i) \geq m_i, \, i \in [k]\right) \to 0.$$

Theorem 2.1 follows directly from Lemmas 2.7 and 2.8 (e.g., see Billingsley (1999, Theorem 3.1)).

3. Depths and degrees via truncated selection sets

We briefly analyse the case of a single vertex. Write $\mathcal{S}_n(1) = \{j_1, \ldots, j_{S_n(1)}\}$ with $j_1 > j_2 > \cdots > j_{S_n(1)}$. Throughout the coalescent process, both the degree and depth of 1 are determined by the variables $(h_{1,j_\ell})_{j_\ell \in \mathcal{S}_n(1)}$ as follows. Vertex 1 is the root of $T_n(1)$ in F_n. For $j \notin \mathcal{S}_n(1)$, we have $T_j(1) = T_{j-1}(1)$. For $j \in \mathcal{S}_n(1)$, if $h_{1,j} = 0$ then the root of $T_j(1)$ increases its degree by one; otherwise, every vertex in $T_j(1)$ increases its depth by one. In particular, the degree of 1 in F_j corresponds to the first streak of values 0 in $(h_{v,j_\ell})_{j_\ell > j}$. More precisely, for $1 \leq j < n$, $\deg_{F_j}(1) = \max\{d \in [n] : j_d > j, \, h_{1,j_1} = h_{1,j_2} = \cdots = h_{1,j_d} = 0\}$ and

$$h_{F_j}(1) = \sum_{\ell=1}^{S_n(1) \setminus [j]} h_{1,j_\ell}. \tag{3.1}$$

Recall that $d_n(1) = \deg_{F_1}(1)$ and $h_n(1) = h_{F_1}(1)$. Let $m \in \mathbb{Z}_+$. The event $d_n(1) \geq m$ may occur only if $S_n(1) \geq m$ and, since $(h_{1,j_\ell})_{j_\ell \in \mathcal{S}_n(1)}$ are i.i.d. Bernoulli variables with mean $1/2$, we have that $\mathbb{P}\left(d_n(1) \geq m \,\middle|\, S_n(1) \geq m\right) = 2^{-m}$. In addition, conditional on $d_n(1) \geq m$ we have $h_n(1) \stackrel{\mathcal{L}}{=} \mathrm{Bin}(S_n(1) - m, 1/2)$; this is well-defined since $d_n(1) \geq m$ implies $S_n(1) \geq m$. Together with Lemma 2.6, the distributional equivalence establishes the case $k = 1$ of Theorem 2.1.

The proof of Lemma 2.7 adapts this argument to obtain the asymptotic limit of the variables $(h_{n,1}(i), \, i \in [k])$; that is, the depths of k distinct vertices in $F_{\lfloor \ln^2 n \rfloor}$. Recall the definition of $\mathcal{S}_{n,1}(v)$ in (2.9) and that $S_{n,1}(v) = |\mathcal{S}_{n,1}(v)|$. Let

$$\Omega_1 := \{J : J \subset \{\lfloor \ln^2 n \rfloor + 1, \ldots, n\}\}; \tag{3.2}$$

we use a superscript bar to denote vectors indexed by $[k]$; for example, we write $\bar{\mathcal{S}}_{n,1} := (\mathcal{S}_{n,1}(1), \ldots, \mathcal{S}_{n,1}(k)) \in \Omega_1^k$.

For $\bar{m} = (m_1, \ldots, m_k) \in \mathbb{Z}_+^k$ and $\delta \in (0, 2)$. Let

$$\mathcal{A}_{\bar{m}} := \{\bar{J} \in \Omega_1^k : \mathbb{P}\left(\bar{\mathcal{S}}_{n,1} = \bar{J}, \, d_n(i) \geq m_i, \, i \in [k]\right) > 0\} \tag{3.3}$$

contain the outcomes of the first k truncated sets $\bar{\mathcal{S}}_{n,1}$ on the event $\{d_n(i) \geq m_i, \, i \in [k]\}$ and let

$$\mathcal{B}_\delta := \{\bar{J} \in \Omega_1^k : (J_1, \ldots, J_k) \text{ are pairwise disjoint}, \, ||J_i| - 2 \ln n| \leq \delta \ln n, \, i \in [k]\}. \tag{3.4}$$

The following lemma gathers two observations about events conditional on $\bar{\mathcal{S}}_{n-1} = \bar{J}$ and a sufficient condition on $\bar{m}$ for $\mathcal{B}_\delta \subset \mathcal{A}_{\bar{m}}$ to hold.

Lemma 3.1. *Let $k \geq 1$ and $\bar{m}, \bar{\ell} \in \mathbb{Z}_{\geq 0}^k$. If $\bar{J} \in \Omega_1^k$ has pairwise disjoint sets and $|J_i| \geq m_i$ for $i \in [k]$, then*

$$\mathbb{P}\left(h_{n,1}(i) \leq \ell_i, \, d_n(i) \geq m_i; \, i \in [k] \,\middle|\, \bar{\mathcal{S}}_{n,1} = \bar{J}\right)$$

$$= \prod_{i=1}^{k} \mathbb{P}\left(h_{n,1}(i) \leq \ell_i, \, d_n(i) \geq m_i \,\middle|\, \mathcal{S}_{n,1}(i) = J_i\right); \tag{3.5}$$

if $\bar{J} \in \mathcal{A}_{\bar{m}}$ then

$$\mathbb{P}\left(d_n(i) \geq m_i; \, i \in [k] \,\middle|\, \bar{\mathcal{S}}_{n,1} = \bar{J}\right) = 2^{-\sum_i m_i}. \tag{3.6}$$

Finally, if $\delta \in (0,2)$ satisfies $m_i < (2 - \delta) \ln n$ for $i \in [k]$, then $\mathcal{B}_\delta \subset \mathcal{A}_{\bar{m}}$.

Proof: Let $k \geq 1$, $\bar{m} \in \mathbb{Z}_{\geq 0}^k$ and let $\bar{J} \in \Omega_1^k$ satisfy $|J_i| \geq m_i$ for $i \in [k]$. For each $i \in [k]$, write $J_i = \{j_{i,1}, j_{i,2}, \ldots, j_{i,|J_i|}\}$ with $j_{i,1} > j_{i,2} > \cdots > j_{i,|J_i|}$ and let $D_i = \{j_{i,1}, j_{i,2}, \ldots, j_{i,m_i}\}$.

On the event that $\bar{\mathcal{S}}_{n,1} = \bar{J}$, for each $i \in [k]$,

$$h_{n,1}(i) = \sum_{\ell=1}^{|J_i|} h_{i,j_\ell},$$

and $d_n(i) \geq m_i$ if and only if $h_{i,j_\ell} = 0$ for all $j_\ell \in D_i$. Consequently, for $\bar{\ell} \in \mathbb{Z}_{\geq 0}^k$, the event $\{h_{n,1}(i) \leq \ell_i, d_n(i) \geq m_i, i \in [k]\}$ depends solely on the variables $(h_{i,j_{i,\ell}})_{j_{i,\ell} \in D_i}$ for $i \in [k]$. If $\bar{J}$ has pairwise disjoint sets, then so has $\bar{D}$ pairwise disjoint sets and (3.5) follows.

Now, at each step of the coalescent process, exactly one vertex increases its degree. Observe that if $\bar{J} \in \mathcal{A}_{\bar{m}}$ then $\mathbb{P}\left(d_n(i) \geq m_i, i \in [k] \mid \bar{\mathcal{S}}_{n,1} = \bar{J}\right) > 0$ and the indices in $\cup_{i \in [k]} D_i$ corresponds to steps at which precisely one of the vertices in $[k]$ increases its degree. It follows that $\bar{D}$ must have pairwise disjoint sets and so, conditional on $\bar{\mathcal{S}}_{n,1} = \bar{J}$, the variables $(h_{i,j_{i,\ell}})_{j_\ell \in D_i}$, $i \in [k]$ are independent Bernoulli variables with mean $1/2$; establishing (3.6).

Finally, if $\delta \in (0,2)$ satisfies $m_i < (2 - \delta) \ln n$ for $i \in [k]$, then $\bar{J} \in \mathcal{B}_\delta$ has pairwise disjoint sets and $|J_i| \geq m_i$ for each $i \in [k]$. Hence, $\mathbb{P}\left(d_n(i) \geq m_i \mid \mathcal{S}_{n,1}(i) = J_i\right) > 0$, and deterministically $h_n(i) \leq n$ for all $i \in [k]$; we then infer from (3.5) with $\bar{\ell} = (n, \ldots, n)$ that $\bar{J} \in \mathcal{A}_{\bar{m}}$. □

The next lemma states that the distribution of $\bar{\mathcal{S}}_{n,1}$ is asymptotically approximated by k copies of independent truncated selection sets; we delay its proof to Section 3.3.

Lemma 3.2. *Let $k \geq 2$. Let $\bar{\mathcal{R}}_n := (\mathcal{R}_n(i), i \in [k])$ be k independent copies of $\mathcal{S}_{n,1}(1)$. For any $\delta \in (0,2)$, we have*

$$\mathbb{P}\left(\bar{\mathcal{S}}_{n,1} \in \mathcal{B}_\delta\right) = 1 + O(\ln^{-2} n), \tag{3.7}$$

and, uniformly for $\bar{J} \in \mathcal{B}_\delta$,

$$\mathbb{P}\left(\bar{\mathcal{S}}_{n,1} = \bar{J}\right) = (1 + O(\ln^{-1} n))\mathbb{P}\left(\bar{\mathcal{R}}_n = \bar{J}\right). \tag{3.8}$$

Before proceeding to the proofs of Lemmas 2.7, 2.8 and 3.2 we recall a result from Addario-Berry and Eslava (2018) which establishes asymptotic independence of the degrees of any finite number of vertices in T_n'.

Proposition 3.3 (Proposition 4.2 in Addario-Berry and Eslava (2018)). *Fix $c \in (0,2)$ and $k \in \mathbb{Z}_+$. There exists $\beta = \beta(c,k) > 0$ such that uniformly over positive integers $m_1, \ldots, m_k < c \ln n$,*

$$\mathbb{P}\left(d_n(i) \geq m_i, i \in [k]\right) = 2^{-\sum_i m_i}(1 + o(n^{-\beta})).$$

3.1. *Proof of Lemma 2.7.* Fix $k \in \mathbb{Z}_+$, $\bar{a} \in [0,2)^k$ and $\bar{b} \in \mathbb{Z}^k$. Let $\bar{x} \in \mathbb{R}^k$, $m_i := \lfloor a_i \ln n \rfloor + b_i$ and $\ell_i := (1 - a_i/2) \ln n + x_i \sqrt{(1 - a_i/4) \ln n}$; let $c \in (0,2)$ satisfy $m_i < c \ln n$ for all $i \in [k]$ and n be large enough.

Let us write $m = m_1$. We infer from (3.1) that, conditional on $\{d_n(1) \geq m, S_{n,1}(1) \geq m\}$, we have $h_{n,1}(1) \stackrel{\mathcal{L}}{=} \text{Bin}(\hat{S}_n, 1/2)$; where $\hat{S}_n$ has the law of $S_{n,1}(1) - m$ conditional on $S_{n,1}(1) \geq m$. In particular,

$$\mathbb{P}\left(h_{n,1}(1) \leq \ell_1 \mid d_n(1) \geq m, S_{n,1}(1) \geq m\right)$$
$$= \mathbb{P}\left(\text{Bin}(S_{n,1}(1) - m, 1/2) \leq \ell_1 \mid S_{n,1}(1) \geq m\right);$$

from which we obtain

$$\mathbb{P}\left(h_{n,1}(1) \leq \ell_1, d_n(1) \geq m\right)$$
$$=\mathbb{P}\left(\text{Bin}(S_{n,1}(1) - m, 1/2) \leq \ell_1 \,|\, S_{n,1}(1) \geq m\right) \mathbb{P}\left(d_n(1) \geq m, S_{n,1}(1) \geq m\right)$$
$$+ \mathbb{P}\left(h_{n,1}(1) \leq \ell_1, d_n(1) \geq m, S_{n,1}(1) < m\right).$$

Now, recall that $d_n(1) \geq m$ only if $S_n(1) \geq m$ and observe that the event $\{d_n(1) \geq m\}$ implies that throughout the complete coalescent process there are at least m fair, independent coin flips that land heads which are also independent from the selection sets. Since both $S_n(1)$ and $S_{n,1}(1)$ concentrate around $2 \ln n$ we have

$$\lim_{n \to \infty} \frac{\mathbb{P}\left(d_n(1) \geq m, S_{n,1}(1) \geq m\right)}{\mathbb{P}\left(d_n(1) \geq m\right)} = \lim_{n \to \infty} \frac{2^{-m}\mathbb{P}\left(S_{n,1}(1) \geq m\right)}{2^{-m}\mathbb{P}\left(S_n(1) \geq m\right)} = 1,$$

and

$$\lim_{n \to \infty} \frac{\mathbb{P}\left(d_n(1) \geq m, S_{n,1}(1) < m\right)}{\mathbb{P}\left(d_n(1) \geq m\right)} \leq \lim_{n \to \infty} \frac{2^{-m}\mathbb{P}\left(S_{n,1}(1) < m\right)}{2^{-m}\mathbb{P}\left(S_n(1) \geq m\right)} = 0.$$

Putting together all the estimates above, we obtain as $n \to \infty$

$$\mathbb{P}\left(h_{n,1}(1) \leq \ell_1 \,|\, d_n(1) \geq m\right) \to \mathbb{P}\left(\text{Bin}(S_{n,1}(1) - m, 1/2) \leq \ell_1 \,|\, S_{n,1}(1) \geq m\right); \qquad (3.9)$$

which establishes, by Lemma 2.6, the desired convergence if $k = 1$ and more generally, by the exchangeability of the vertices, for any $i \in [k]$,

$$\lim_{n \to \infty} \mathbb{P}\left(h_{n,1}(i) \leq \ell_i \,|\, d_n(i) \geq m_i\right) = \Phi(x_i). \qquad (3.10)$$

For $k \geq 2$, assume that

$$\mathbb{P}\left(h_{n,1}(i) \leq \ell_i, d_n(i) \geq m_i, \, i \in [k]\right)$$
$$=\prod_{i=1}^{k} \mathbb{P}\left(h_{n,1}(i) \leq \ell_i, d_n(i) \geq m_i\right) + o\left(2^{-\sum_i m_i}\right). \qquad (3.11)$$

By Proposition 3.3, we have

$$\prod_{i=1}^{k} \mathbb{P}\left(d_n(i) \geq m_i\right) = (1 + o(1))2^{-\sum_i m_i} = \mathbb{P}\left(d_n(i) \geq m_i, \, i \in [k]\right). \qquad (3.12)$$

Putting together (3.10)–(3.12) implies, for arbitrary $\bar{x} \in \mathbb{R}^k$,

$$\mathbb{P}\left(h_{n,1}(i) \leq \ell_i, d_n(i) \geq m_i, \, i \in [k]\right)$$
$$=(1 + o(1))\mathbb{P}\left(d_n(i) \geq m_i, \, i \in [k]\right) \prod_{i=1}^{k} \mathbb{P}\left(h_{n,1}(i) \leq \ell_i \,|\, d_n(i) \geq m_i\right)$$
$$=(1 + o(1))\mathbb{P}\left(d_n(i) \geq m_i, \, i \in [k]\right) \prod_{i=1}^{k} \Phi(x_i);$$

this yields the desired convergence in law for $k \geq 2$.

It then remains to prove (3.11). Lemma 3.1 and the choice of c gives $\mathcal{B}_{2-c} \subset \mathcal{A}_{\bar{m}}$, together with Lemma 3.2 we have

$$\mathbb{P}\left(\bar{\mathcal{S}}_{n,1} \in \mathcal{A}_{\bar{m}} \setminus \mathcal{B}_{n,2-c}\right) \leq \mathbb{P}\left(\bar{\mathcal{S}}_{n,1} \in \Omega_1^k \setminus \mathcal{B}_{n,2-c}\right) = o(1). \qquad (3.13)$$

Now, we decompose the event in the left-hand side of (3.11) and use (3.5), (3.6) and (3.13) to neglect the contribution of $\bar{J} \notin \mathcal{B}_{2-c}$,

$$\mathbb{P}\left(h_{n,1}(i) \leq \ell_i, d_n(i) \geq m_i, i \in [k]\right)$$

$$= \sum_{\bar{J} \in \mathcal{B}_{n,2-c}} \mathbb{P}\left(h_{n,1}(i) \leq \ell_i, d_n(i) \geq m_i, i \in [k] \mid \bar{\mathcal{S}}_{n,1} = \bar{J}\right) \mathbb{P}\left(\bar{\mathcal{S}}_{n,1} = \bar{J}\right) + o\left(2^{-\sum_i m_i}\right)$$

$$= \sum_{\bar{J} \in \mathcal{B}_{n,2-c}} \prod_{i=1}^{k} \mathbb{P}\left(h_{n,1}(i) \leq \ell_i, d_n(i) \geq m_i \mid \bar{\mathcal{S}}_{n,1} = \bar{J}\right) \mathbb{P}\left(\bar{\mathcal{S}}_{n,1} = \bar{J}\right) + o\left(2^{-\sum_i m_i}\right)$$

$$= \sum_{\bar{J} \in \Omega_1^k} \prod_{i=1}^{k} \mathbb{P}\left(h_{n,1}(i) \leq \ell_i, d_n(i) \geq m_i \mid \bar{\mathcal{S}}_{n,1} = \bar{J}\right) \mathbb{P}\left(\bar{\mathcal{S}}_{n,1} = \bar{J}\right) + o\left(2^{-\sum_i m_i}\right);$$

establishing (3.11).

3.2. *Proof of Lemma 2.8.* Let $k \geq 2$, $c \in (0,2)$ and $\varepsilon > 0$. Let $m_i = m_i(n) < c \ln n$ for all $i \in [k]$. By Proposition 3.3 and the exchangeability of the vertices, it suffices to show that

$$\mathbb{P}\left(h_{n,2}(1) \geq \varepsilon \sqrt{\ln n}, \, d_n(i) \geq m_i, \, i \in [k]\right) = o\left(2^{-\sum_i m_i}\right).$$

Let us assume that uniformly over $\bar{J} \in \mathcal{B}_{n,2-c}$,

$$\mathbb{P}\left(h_{n,2}(1) \geq \varepsilon \sqrt{\ln n}, \, d_n(i) \geq m_i, \, i \in [k] \mid \bar{\mathcal{S}}_{n,1} = \bar{J}\right) = o\left(2^{-\sum_i m_i}\right). \tag{3.14}$$

Then using (3.14), (3.6) and (3.7), we get

$$\mathbb{P}\left(h_{n,2}(1) \geq \varepsilon \sqrt{\ln n}, \, d_n(i) \geq m_i, \, i \in [k]\right)$$

$$= \sum_{\bar{J} \in \mathcal{A}_{\bar{m}}} \mathbb{P}\left(h_{n,2}(1) \geq \varepsilon \sqrt{\ln n}, \, d_n(i) \geq m_i, \, i \in [k] \mid \bar{\mathcal{S}}_{n,1} = \bar{J}\right) \mathbb{P}\left(\bar{\mathcal{S}}_{n,1} = \bar{J}\right)$$

$$\leq \mathbb{P}\left(\bar{\mathcal{S}}_{n,1} \in \mathcal{B}_{2-c}\right) o\left(2^{-\sum_i m_i}\right) + 2^{-\sum_i m_i} \mathbb{P}\left(\bar{\mathcal{S}}_{n,1} \in \mathcal{A}_{\bar{m}} \setminus \mathcal{B}_{2-c}\right) = o\left(2^{-\sum_i m_i}\right).$$

Now, to prove (3.14), let $\bar{J} \in \mathcal{B}_{2-c}$ and simplifying notation write $[\ln^2 n] = [\lfloor \ln^2 n \rfloor]$. Observe that $|J_i| \geq m_i$ for $i \in [k]$ and so the selection times $\mathcal{S}_n(1) \cap [\ln^2 n]$ are independent of the event $\{d_n(i) \geq m_i, \, i \in [k]\} \cap \{\bar{\mathcal{S}}_{n,1} = \bar{J}\}$. Moreover, $|\mathcal{S}_n(1) \cap [\ln^2 n]| \overset{\mathcal{L}}{=} S_{\lfloor \ln^2 n \rfloor + 1}$ which has mean $(1 + o(1)) 2 \ln \ln n$. Using Berstein's inequalities (e.g., see Boucheron et al. (2013, Theorem 2.10)) we get

$$\mathbb{P}\left(|\mathcal{S}_n(1) \cap [\ln^2 n]| \geq \varepsilon \sqrt{\ln n} \,\Big|\, d_n(i) \geq m_i, \, i \in [k], \, \bar{\mathcal{S}}_{n,1} = \bar{J}\right)$$

$$= \mathbb{P}\left(S_{\lfloor \ln^2 n \rfloor + 1}(1) \geq \varepsilon \sqrt{\ln n}\right) = o(1).$$

It is straightforward that $h_{n,2}(1)$ is stochastically dominated by $|\mathcal{S}_n(1) \cap [\ln^2 n]|$, so the estimate above together with (3.6) (recall that $\mathcal{B}_{2-c} \subset \mathcal{A}_{\bar{m}}$) establishes (3.14).

3.3. *Proof of Lemma 3.2.* Let $k \geq 2$ and

$$\tau_k := \max\{1 < j \leq n : s_{j,i} = s_{j,i'} = 1 \text{ for distinct } i, i' \in [k]\}.$$

Then,

$$\mathbb{P}((\mathcal{S}_{n,1}(i), \, i \in [k]) \text{ are pairwise disjoint}) = \mathbb{P}\left(\tau_k \leq \ln^2 n\right)$$

$$\geq 1 - 2k^2 \ln^{-2} n; \tag{3.15}$$

the equality holds by definition of τ_k and the condition on $\bar{\mathcal{S}}_{n,1}$ having pairwise disjoint sets, while the inequality follows from Addario-Berry and Eslava (2018, Lemma 4.6).

Let $\delta \in (0,2)$, by Berstein's inequality (e.g., see Boucheron et al. (2013, Theorem 2.10)) there is $\beta = \beta(\delta) > 0$ such that for any $i \in [k]$

$$\mathbb{P}\left(|S_{n,1}(i) - 2\ln n| \geq \delta \ln n\right) = o(n^{-\beta}); \tag{3.16}$$

it follows that a union bound, together with (3.15) and (3.16) yields (3.7).

For the proof of (3.8) we will use the following factors. For $\ln^2 n < m \leq n$ let

$$p_{m,0} := \frac{(m-k)(m-k-1)}{m(m-1)}, \qquad\qquad q_{m,0} := \left(1 - \frac{2}{m}\right)^k, \tag{3.17}$$

$$p_{m,1} := \frac{2(m-k)}{m(m-1)}, \qquad\qquad q_{m,1} := \frac{2}{m}\left(1 - \frac{2}{m}\right)^{k-1}. \tag{3.18}$$

There exists a constant $C = C(k) > 0$ such that for m large enough, we have for $\sigma \in \{0,1\}$,

$$\left(1 - \frac{C}{m^2}\mathbf{1}_{[\sigma=0]}\right)q_{m,\sigma} < p_{m,\sigma} < q_{m,\sigma}\left(1 + \frac{C}{m}\mathbf{1}_{[\sigma=1]}\right). \tag{3.19}$$

Indeed, write $q_{m,0} = 1 - \frac{2k}{m} + \frac{2k(k-1)}{m^2} + cm^{-3}$ and $q_{m,1} = \frac{2}{m}\left(1 - \frac{2(k-1)}{m} + c'm^{-2}\right)$, it is straightforward that $\max\{|c|, |c'|\} \leq 2^k$. We then have

$$q_{m,0} - p_{m,0} = \frac{k(k-1)(m-2)}{m^2(m-1)} + cm^{-3},$$

$$p_{m,1} - q_{m,1} = \frac{2}{m}\left(\frac{(k-1)(m-2)}{m(m-1)} - c'm^{-2}\right);$$

which implies that $q_{m,0} - p_{m,0} = \Theta(m^{-2})$ and $q_{m,1} - p_{m,1} = \Theta(m^{-2})$. Similarly, since $\lim_{m\to\infty} q_{m,0} = 1$ and $\lim_{m\to\infty} mq_{m,1} = 2$, there is $C = C(k) > 0$ such that for m large enough, $0 < q_{m,0} - p_{m,0} \leq Cq_{m,0}m^{-2}$ and $0 < p_{m,1} - q_{m,1} \leq Cq_{m,1}m^{-1}$.

In what follows fix $\bar{J} \in \mathcal{B}_\delta$ and let $\bar{\mathcal{R}}_n$ be k i.i.d. copies of $\mathcal{S}_{n,1}(1)$. Set $r_{m,i} = \mathbf{1}_{[m \in \mathcal{R}_n(i)]}$ and $j_{m,i} = \mathbf{1}_{[m \in J_i]}$ for $\ln^2 n < m \leq n$ and $i \in [k]$. Let $\sigma_m = \sum_{i\in[k]} j_{m,i}$ and observe that $\sigma_m \in \{0,1\}$. Recall that $\mathcal{S}_{n,1}(i) = \{\ln^2 n < m \leq n : s_{m,i} = 1\}$ for $i \in [k]$ and let $A_m := \{s_{m,i} = j_{m,i}, i \in [k]\}$ for $\ln^2 n < m \leq n$. Then

$$\mathbb{P}\left(\bar{\mathcal{S}}_{n,1} = \bar{J}\right) = \mathbb{P}(A_n)\prod_{m=\lfloor \ln^2 n\rfloor+1}^{n-1}\mathbb{P}\left(A_m \mid A_l, m < l \leq n\right) = \prod_{m=\lfloor \ln^2\rfloor+1n}^{n} p_{m,\sigma_m};$$

to see this, use that $\bar{J}$ has pairwise disjoint sets to infer that $\#\{T_m(i); i \in [k]\} = k$ for all $m > \ln^2 n$; that is, at no point we have two trees $T_m(i)$ and $T_m(i')$ selected to be merged with $i, i' \in [k]$ and $m > \ln^2 n$. Then, for $\sigma \in \{0,1\}$, $p_{m,\sigma}$ is the probability of selecting σ trees from $\{T_m(i); i \in [k]\}$ when selecting two distinct trees uniformly at random from the forest F_m.

Next, observe that, for n large enough,

$$\prod_{m=\lfloor \ln^2 n\rfloor+1}^{n}\left(1 - \frac{C}{m^2}\right) > 1 - \sum_{m=\lfloor \ln^2 n\rfloor+1}^{n}\frac{2C}{m^2} > 1 - 2C\int_{\ln^2 n}^{\infty} x^{-2}dx = 1 - 2C\ln^{-2} n.$$

On the other hand, we have that $\sum \sigma_m = \sum_{i\in[k]}|J_i| \leq (2+\delta)k\ln n$ and, for n large enough,

$$\prod_{m:\sigma_m=1}\left(1 + \frac{C}{m}\right) \leq \exp\left(\frac{(2+\delta)Ck\ln n}{\ln^2 n}\right) < 1 + \frac{(2+\delta)2Ck}{\ln n};$$

putting together these estimates we obtain

$$\prod_{m=\lfloor \ln^2 n \rfloor+1}^{n} p_{m,\sigma_m} = (1 + O(\ln^{-1} n)) \prod_{m=\lfloor \ln^2 n \rfloor+1}^{n} q_{m,\sigma_m}.$$

The independence of the sets in $\bar{\mathcal{R}}_n$ gives $\mathbb{P}\left(\bar{\mathcal{R}}_n = \bar{J}\right) = \prod_{m=\lfloor \ln^2 n \rfloor+1}^{n} q_{m,\sigma_m}$ and so $\mathbb{P}\left(\bar{\mathcal{S}}_{n,1} = \bar{J}\right) = (1 + O(\ln^{-1} n))\mathbb{P}\left(\bar{\mathcal{R}}_n = \bar{J}\right)$; establishing (3.8).

4. Proof of Proposition 2.2

Recall that we write $d_n(v) = \deg_{T'_n}(v)$ and $h_n(v) = h_{T'_n}(v)$ and, from (2.2), $\mu_n = (1-(\log e)/2)\ln n$ and $\sigma_n = (1 - (\log e)/4)\ln n$. A simpler version of the next lemma appeared in Addario-Berry and Eslava (2018, Lemma 5.1).

Lemma 4.1. *Let $k \in \mathbb{Z}_+$. Let $(c_i)_{i\in[k]}$ be a non-decreasing sequence of integers and let $(B_i)_{i\in[k]}$ be sets in $\mathcal{B}_\mathbb{R}$ such that $B_i \cap B_j = \emptyset$ whenever $c_i = c_j$ and $i \neq j$. Set $k' = \min\{i \geq 0 : c_{i+1} = c_k\}$ and $m_i := \lfloor \log n \rfloor + c_i$ for $i \in [k]$, and*

$$\mathcal{D}_{\bar{m}} := \{d_n(i) = m_i, d_n(j) \geq m_j \, 1 \leq i \leq k' < j \leq k\},$$

$$\mathcal{H}_{\bar{B}} := \left\{ \frac{h_n(i) - \mu_n \ln n}{\sqrt{\sigma_n^2 \ln n}} \in B_i, \, i \in [k] \right\}.$$

Then

$$\mathbb{P}\left(\mathcal{D}_{\bar{m}}, \mathcal{H}_{\bar{B}}\right) = (1 + o(1))2^{-k'-\sum_i m_i} \prod_{i=1}^{k} \Phi(B_i).$$

Proof: By the inclusion-exclusion principle,

$$\mathbb{P}\left(d_n(i) = m_i, \, i \in [k']\right) = \sum_{\ell=0}^{k'} \sum_{\substack{S \subset [k'] \\ |S|=\ell}} (-1)^\ell \mathbb{P}\left(d_n(i) \geq m_i + \mathbf{1}_{[i\in S]}, \, i \in [k']\right). \tag{4.1}$$

By intersecting the event $\mathcal{H}_{\bar{B}} \cap \{d_n(j) \geq m_j, \, k' < j \leq k\}$ with $\{d_n(i) = m_i, \, i \in [k']\}$ and each of the events $\{d_n(i) \geq m_i + \mathbf{1}_{[i\in S]}, \, i \in [k']\}$ from (4.1) we obtain

$$\mathbb{P}\left(\mathcal{D}_{\bar{m}}, \mathcal{H}_{\bar{B}}\right) = \sum_{\ell=0}^{k'} \sum_{\substack{S \subset [k'] \\ |S|=\ell}} (-1)^\ell \mathbb{P}\left(\mathcal{H}_{\bar{B}}, d_n(i) \geq m_i + \mathbf{1}_{[i\in S]}, \, i \in [k]\right). \tag{4.2}$$

By Proposition 3.3, for each $S \subset [k']$ with $|S| = \ell$ we have

$$\mathbb{P}\left(d_n(i) \geq m_i + \mathbf{1}_{[i\in S]}, \, i \in [k]\right) = 2^{-\ell-\sum_i m_i}(1 + o(1)).$$

Conditioning on $\{d_n(i) \geq m_i + \mathbf{1}_{[i\in S]}, \, i \in [k]\}$, by Theorem 2.1 (with $a_i = \log e$ and $b_i = c_i + \mathbf{1}_{[i\in S]}$), we get

$$\mathbb{P}\left(\mathcal{H}_{\bar{B}}, d_n(i) \geq m_i + \mathbf{1}_{[i\in S]}, \, i \in [k]\right) = (1 + o(1))2^{-\ell-\sum_i m_i} \prod_{i=1}^{k} \Phi(B_i).$$

Finally, we use the above estimate for each of the summands of the right-hand side of (4.2) to obtain the desired expression; namely,

$$\mathbb{P}\left(\mathcal{D}_{\bar{m}},\, \mathcal{H}_{\bar{B}}\right) = (1 + o(1))2^{-\sum_i m_i} \prod_{i=1}^{k} \Phi(B_i) \sum_{\ell=0}^{k'} \sum_{\substack{S \subset [k'] \\ |S|=\ell}} (-1)^{\ell} 2^{-\ell}$$

$$= (1 + o(1))2^{-k' - \sum_i m_i} \prod_{i=1}^{k} \Phi(B_i). \qquad \square$$

Proof of Proposition 2.2: Fix $c \in (0, 2)$ and $K, M \in \mathbb{Z}_+$. Let $(a_k)_{k \in [K]}$ be non-negative integers such that $\sum_{k \in [K]} a_k = M$. Let $(j_k)_{k \in [K]}$ be a non-decreasing sequence of integers with $K' = \min\{k \geq 0 : j_{k+1} = j_K\}$ such that $0 \leq j_1 + \log n$ and $j_K + \log n < c \ln n$. Let $(B_k)_{k \in [K]}$ be sets in $\mathcal{B}_{\mathbb{R}}$ satisfying $B_k \cap B_\ell = \emptyset$ whenever $j_k = j_\ell$ and $k \neq \ell$.

We define $m_i \in \mathbb{Z}_+$ and $A_i \subset \mathbb{R}$ as follows. For each $k \in [K]$, if $\sum_{\ell=1}^{k-1} a_\ell < i \leq \sum_{\ell=1}^{k} a_\ell$ then set $m_i = \lfloor \log n \rfloor + j_k$ and let $A_i = B_{j_k}$. Let $M' = \sum_{k \in [K']} a_k$. With foresight, we verify that $\prod_{i=1}^{M} \Phi(A_i) = \prod_{k=1}^{K} \Phi(B_k)^{a_k}$ and similarly,

$$M \log n - M' - \sum_{i=1}^{M} m_i = \sum_{k=1}^{K'}(-j_k - 1 + \varepsilon_n)a_k + \sum_{k=K'+1}^{K}(-j_K + \varepsilon_n)a_k.$$

Now, consider the sets

$$\mathcal{D}_{\bar{m}} := \{d_n(i) = m_i,\, d_n(j) \geq m_j,\, 1 \leq i \leq M' < j \leq M\},$$

$$\mathcal{H}_{\bar{A}} := \left\{ \frac{h_n(i) - \mu_n \ln n}{\sqrt{\sigma_n^2 \ln n}} \in A_i,\, i \in [M] \right\}.$$

Recall that $X_i^{(n)}(B), X_{\geq i}^{(n)}(B),\, i \in \mathbb{Z}, B \in \mathcal{B}_{\mathbb{R}}$ are sums of indicator functions; by the definition of T_n' (as a relabelling of T_n) and the exchangeability of its vertex set, we have

$$\mathbb{E}\left[\prod_{k=1}^{K'} \left(X_{j_k}^{(n)}(B_k)\right)_{a_k} \prod_{k=K'+1}^{K} \left(X_{\geq j}^{(n)}(B_k)\right)_{a_k} \right] = (n)_M \mathbb{P}\left(\mathcal{D}_{\bar{m}},\, \mathcal{H}_{\bar{A}}\right).$$

Finally, using the above estimates and since $(n)_M = n^M(1 + o(n^{-1}))$, Lemma 4.1 implies

$$\mathbb{E}\left[\prod_{k=1}^{K'} \left(X_{j_k}^{(n)}(B_k)\right)_{a_k} \prod_{k=K'+1}^{K} \left(X_{\geq j_k}^{(n)}(B_k)\right)_{a_k} \right]$$

$$= (1 + o(1)) \left(2^{M \log n - M' - \sum_i m_i}\right) \prod_{i=1}^{M} \Phi(A_i)$$

$$= (1 + o(1)) \prod_{k=1}^{K'} \left(2^{-j_k - 1 + \varepsilon_n} \Phi(B_k)\right)^{a_k} \prod_{k=K'+1}^{K} \left(2^{-j' - \varepsilon_n} \Phi(B_k)\right)^{a_k},$$

as desired. $\qquad \square$

5. Proof of Theorem 1.2 and Corollary 1.3

Throughout the section we let $\varepsilon \in [0, 1)$ and consider an increasing subsequence $(n_\ell)_{\ell \geq 1}$ with $\log n_\ell - \lfloor \log n_\ell \rfloor \to \varepsilon$ as $\ell \to \infty$. Recall also that $\mu_{n_\ell} = (1 - (\log e)/2) \ln n_\ell$ and $\sigma_{n_\ell}^2 = (1 - (\log e)/4) \ln n_\ell$ as in (2.2). To simplify notation, let $\lambda_j := 2^{-j-1+\varepsilon}$ for each $j \in \mathbb{Z}$.

Proof of Theorem 1.2: We will prove that $\mathcal{MP}^{(n_\ell)}$ converges weakly to $\mathcal{MP}^\varepsilon$, as $\ell \to \infty$, in $\mathcal{M}^{\#}_{\mathbb{Z}^* \times \mathbb{R}}$. The weak convergence is equivalent to the statement in Theorem 1.2 since we may list the vertices in T_n in decreasing order of degree as $v^{(i)}$ and let $\mathcal{P} = \{P_i, i \geq 1\}$ as defined in (1.3); then

$$\mathcal{MP}^\varepsilon = \sum_{i \geq 1} \delta_{(\lfloor P_i \rfloor + \varepsilon, \xi_{P_i})}, \quad \text{and} \quad \mathcal{MP}^{(n)} = \sum_{i=1}^{n} \delta_{(\deg_{T_n}(v^{(i)}) - \lfloor \log n \rfloor, (h(v^{(i)}) - \mu_n)/\sigma_n)}.$$

Consider the semiring of bounded sets that generates $\mathbb{Z}^* \times \mathbb{R}$ defined by

$$\mathcal{A} := \{\{i\} \times (a, b], i \in \mathbb{Z}, a, b \in \mathbb{R}\} \cup \{[i, \infty] \times (a, b], i \in \mathbb{Z}, a, b \in \mathbb{R}\}. \tag{5.1}$$

It suffices to prove convergence of the counting measures of finite collections of disjoint sets in $\mathcal{A}$; see Daley and Vere-Jones (2003, Proposition 9.2.III and Theorem 11.1.VII).

Consider any non-decreasing sequence of integers $(j_k)_{k \in [K]}$ with $0 \leq K' = \min\{k : j_{k+1} = j_K\}$ and let $(B_k)_{k \in K}$ with $B_k = (a_k, b_k]$ satisfying $B_k \cap B_\ell = \emptyset$ whenever $j_k = j_\ell$ and $k \neq \ell$. Observe that the variables

$$(\mathcal{P}^\varepsilon(\{j_1\} \times B_1), \ldots, \mathcal{P}^\varepsilon(\{j_{K'}\} \times B_{K'}) \quad \text{and}$$
$$\mathcal{P}^\varepsilon([j_{K'+1}, \infty] \times B_{K'+1}), \ldots, \mathcal{P}^\varepsilon([j_K, \infty] \times B_K)), \tag{5.2}$$

are independent with distributions

$$\mathcal{P}^\varepsilon(\{i\}, B) \stackrel{\mathcal{L}}{=} \text{Poi}(2^{-i-1+\varepsilon}\Phi(B)) \qquad \text{and} \qquad \mathcal{P}^\varepsilon([i, \infty], B) \stackrel{\mathcal{L}}{=} \text{Poi}(2^{-i+\varepsilon}\Phi(B)).$$

Recall that, if $X \stackrel{\mathcal{L}}{=} \text{Poi}(\lambda)$, then $\mathbb{E}[(X)_a] = \lambda^a$ for all integers $a \geq 0$. Setting $c := 3/2$, we infer that for n large enough $0 \leq j_1 + \log n < j_K + \log n < c \ln n$. It follows from Proposition 2.2 and the method of moments (e.g., see Janson et al. (2000, Section 6.1)) that

$$(X^{(n_\ell)}_{j_1}(B_1), \ldots, X^{(n_\ell)}_{j_{K'}}(B_{K'}), X^{(n_\ell)}_{\geq j_{K'+1}}(B_{K'+1}), \ldots, X^{(n_\ell)}_{\geq j_K}(B_K))$$

converge in distribution to the variables in (5.2); completing the proof of the weak convergence of $\mathcal{MP}^{(n_\ell)}$ to $\mathcal{MP}^\varepsilon$ in $\mathcal{M}^{\#}_{\mathbb{Z}^* \times \mathbb{R}}$. $\qquad \square$

Lemma 5.1. *Let $\varepsilon > 0$. The random variable $\mathbf{M}_\varepsilon$ from (1.4) is well defined.*

Proof: Let $\varepsilon > 0$. For each $m \in \mathbb{Z}$, let $p_m = (1 - e^{-\lambda_m})e^{-\lambda_m}$. Observe that $2\lambda_m = \lambda_{m-1}$ and so

$$\sum_{m \in \mathbb{Z}} p_m = \sum_{m \in \mathbb{Z}} e^{-\lambda_m} - e^{-\lambda_{m-1}} = \lim_{m \to \infty} e^{-\lambda_m} - \lim_{m \to -\infty} e^{-\lambda_{m-1}} = 1.$$

Let $(Y_m, m \in \mathbb{Z})$ be a collection of independent random variables where each Y_m has the distribution of a Poisson random variable with mean λ_m conditioned to being positive. Let L_ε be a random variable, independent of $(Y_m, m \in \mathbb{Z})$ such that $\mathbb{P}(L_\varepsilon = m) = p_m$, for $m \in \mathbb{Z}$.

We will show that letting $\mathbf{M}_\varepsilon = Y_{L_\varepsilon}$ recovers the probability mass function in (1.4). To see this, observe that

$$\mathbb{P}(\mathbf{M}_\varepsilon = k) = \sum_{m \in \mathbb{Z}} \mathbb{P}(L_\varepsilon = m)\mathbb{P}(Y_m = k) = \sum_{m \in \mathbb{Z}} (1 - e^{-\lambda_m})e^{-\lambda_m} \cdot \frac{e^{-\lambda_m}\lambda_m^k}{k!(1 - e^{-\lambda_m})},$$

from which the lemma follows. $\qquad \square$

Proof of Corollary 1.3: Let $k \geq 1$. Since $\mathcal{P}$ is independent of $(N_i)_{i \geq 1}$ it follows from Theorem 1.2 that

$$\left((h_{T_{n_\ell}}(v^{(i)}) - \mu_{n_\ell})/\sigma_{n_\ell}, i \in [k]\right) \stackrel{\mathcal{L}}{\longrightarrow} (N_i, i \in [k]);$$

on the other hand, conditional on $|\mathcal{M}_{n_\ell}| = k$ we have $\mathcal{M}_{n_\ell} = \{v^{(1)}, \ldots, v^{(k)}\}$ and, Theorem 1.2 gives

$$\left((h_{T_{n_\ell}}(v_i) - \mu_{n\ell})/\sigma_{n\ell}, \, v_i \in \mathcal{M}_{n_\ell} \right) \overset{\mathcal{L}}{=} \left((h_{T_{n_\ell}}(v^{(i)}) - \mu_{n\ell})/\sigma_{n\ell}, \, i \in [k] \right).$$

So it suffices to prove[3] that $|\mathcal{M}_{n_\ell}| \overset{\mathcal{L}}{\longrightarrow} \mathcal{M}_\varepsilon$ to establish Corollary 1.3.

Let $j \in \mathbb{Z}$. Define $X_j^{(n)} := X_j^{(n)}(\mathbb{R}) \overset{\mathcal{L}}{=} \mathrm{Poi}(\lambda_j)$ and $X_{\geq j+1}^{(n)} := X_{\geq j+1}^{(n)}(\mathbb{R}) \overset{\mathcal{L}}{=} \mathrm{Poi}(\lambda_j)$. For any $k \in \mathbb{Z}_+$, as $n_\ell \to \infty$,

$$\mathbb{P}\left(\Delta_{n_\ell} = \lfloor \log n_\ell \rfloor + j, |\mathcal{M}_{n_\ell}| = k \right) = \mathbb{P}\left(X_j^{(n)} = k, \, X_{\geq j+1}^{(n)} = 0 \right)$$
$$\to \mathbb{P}\left(X_j = k, \, X_{\geq j+1} = 0 \right).$$

It follows from the independence of $X_j^{(n)}$ and $X_{\geq j+1}^{(n)}$ that

$$\lim_{J \to \infty} \lim_{n_\ell \to \infty} \sum_{j=-J}^{J} \mathbb{P}\left(\Delta_{n_\ell} = \lfloor \log n_\ell \rfloor + j, |\mathcal{M}_{n_\ell}| = k \right) = \sum_{j \in \mathbb{Z}} e^{-2\lambda_j} \frac{\lambda_j^k}{k!}. \tag{5.3}$$

It remains to show that the tails are negligible. That is,

$$\lim_{J \to \infty} \lim_{n_\ell \to \infty} \mathbb{P}\left(\Delta_{n_\ell} < \lfloor \log n_\ell \rfloor - J \right) + \mathbb{P}\left(\Delta_{n_\ell} > \lfloor \log n_\ell \rfloor + J \right)$$
$$= \lim_{J \to \infty} \lim_{n_\ell \to \infty} \mathbb{P}\left(X_{\geq J}^{(n_\ell)} = 0 \right) + \mathbb{P}\left(X_{\geq J+1}^{(n_\ell)} > 1 \right)$$
$$= \lim_{J \to \infty} e^{-\lambda_{-J+1}} + 1 - e^{-\lambda_J} = 0. \tag{5.4}$$

Indeed, $\liminf_{n_\ell \to \infty} \mathbb{P}(|\mathcal{M}_{n_\ell}| = k)$ is lower bounded by (5.3) and, similarly, the sum of (5.3) and (5.4) is an upper bound for $\limsup_{n_\ell \to \infty} \mathbb{P}(|\mathcal{M}_{n_\ell}| = k)$, completing the proof. $\qquad\square$

Acknowledgements

References

Addario-Berry, L. Partition functions of discrete coalescents: from Cayley's formula to Frieze's $\zeta(3)$ limit theorem. In *XI Symposium on Probability and Stochastic Processes*, volume 69 of *Progr. Probab.*, pp. 1–45. Birkhäuser/Springer, Cham (2015). MR3558134.

Addario-Berry, L. and Eslava, L. High degrees in random recursive trees. *Random Structures Algorithms*, **52** (4), 560–575 (2018). MR3809688.

Banerjee, S. and Bhamidi, S. Persistence of hubs in growing random networks. *Probab. Theory Related Fields*, **180** (3-4), 891–953 (2021). MR4288334.

Berestycki, N. *Recent progress in coalescent theory*, volume 16 of *Ensaios Matemáticos*. Sociedade Brasileira de Matemática, Rio de Janeiro (2009). ISBN 978-85-85818-40-1. MR2574323.

Bergeron, F., Flajolet, P., and Salvy, B. Varieties of increasing trees. In *CAAP '92 (Rennes, 1992)*, volume 581 of *Lecture Notes in Comput. Sci.*, pp. 24–48. Springer, Berlin (1992). MR1251994.

Bhamidi, S. Universal techniques to analyze preferential attachment tree and networks: Global and local analysis (2007). Unpublished.

---

[3]This convergence was already implicit from the convergence of $\mathcal{P}^{(n_\ell)} \overset{\mathcal{L}}{\longrightarrow} \mathcal{P}^\varepsilon$ in Addario-Berry and Eslava (2018).

Billingsley, P. *Convergence of probability measures.* Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons, Inc., New York, second edition (1999). ISBN 0-471-19745-9. MR1700749.

Boucheron, S., Lugosi, G., and Massart, P. *Concentration Inequalities. A Nonasymptotic Theory of Independence.* Oxford University Press, Oxford (2013). ISBN 978-0-19-953525-5. MR3185193.

Daley, D. J. and Vere-Jones, D. *An introduction to the theory of point processes. Vol. I. Elementary theory and methods.* Probability and its Applications (New York). Springer-Verlag, New York, second edition (2003). ISBN 0-387-95541-0. MR1950431.

Devroye, L. Applications of the theory of records in the study of random trees. *Acta Inform.*, **26** (1-2), 123–130 (1988). MR969872.

Devroye, L. and Lu, J. The strong convergence of maximal degrees in uniform random recursive trees and dags. *Random Structures Algorithms*, **7** (1), 1–14 (1995). MR1346281.

Drmota, M. *Random trees. An interplay between combinatorics and probability.* SpringerWien-NewYork, Vienna (2009). ISBN 978-3-211-75355-2. MR2484382.

Durrett, R. *Probability: theory and examples.* Duxbury Press, Belmont, CA, second edition (1996). ISBN 0-534-24318-5. MR1609153.

Eslava, L. A non-increasing tree growth process for recursive trees and applications. *Combin. Probab. Comput.*, **30** (1), 79–104 (2021). MR4205660.

Goh, W. and Schmutz, E. Limit distribution for the maximum degree of a random recursive tree. *J. Comput. Appl. Math.*, **142** (1), 61–82 (2002). Special issue: Probabilistic methods in combinatorics and combinatorial optimization. MR1910519.

Janson, S., Ł uczak, T., and Rucinski, A. *Random graphs.* Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley-Interscience, New York (2000). ISBN 0-471-17541-2. MR1782847.

Mahmoud, H. M. Limiting distributions for path lengths in recursive trees. *Probab. Engrg. Inform. Sci.*, **5** (1), 53–59 (1991). MR1183165.

Móri, T. F. The maximum degree of the Barabási-Albert random tree. *Combin. Probab. Comput.*, **14** (3), 339–348 (2005). MR2138118.

Na, H. S. and Rapoport, A. Distribution of nodes of a tree by degree. *Math. Biosci.*, **6**, 313–329 (1970). MR278985.

Pittel, B. Note on the heights of random recursive trees and random m-ary search trees. *Random Structures Algorithms*, **5** (2), 337–347 (1994). MR1262983.