

## TREES GROWN UNDER YOUNG-AGE PREFERENTIAL ATTACHMENT

MERRITT R. LYON,<sup>\*\*</sup> AND  
HOSAM M. MAHMOUD,<sup>\*\*\*</sup> *The George Washington University*

### Abstract

We introduce a class of non-uniform random recursive trees grown with an attachment preference for young age. Via the Chen–Stein method of Poisson approximation, we find that the outdegree of a node is characterized in the limit by ‘perturbed’ Poisson laws, and the perturbation diminishes as the node index increases. As the perturbation is attenuated, a pure Poisson limit ultimately emerges in later phases. Moreover, we derive asymptotics for the proportion of leaves and show that the limiting fraction is less than one half. Finally, we study the insertion depth in a random tree in this class. For the insertion depth, we find the exact probability distribution, involving Stirling numbers, and consequently we find the exact and asymptotic mean and variance. Under appropriate normalization, we derive a concentration law and a limiting normal distribution. Some of these results contrast with their counterparts in the uniform attachment model, and some are similar.

*Keywords:* Random tree; recursive tree; outdegree; insertion depth; tree leaves; preferential attachment; Poisson approximation; Chen–Stein method; phase transition; small world; Stirling numbers

2010 Mathematics Subject Classification: Primary 05C05; 05C82; 90B15  
Secondary 60C05; 60F25; 60F05

### 1. Introduction

The classic (uniform) random recursive tree is a natural stochastic growth model, which has been proposed for numerous applications in such areas as chain letters [9] and philology [14]. These trees grow according to a uniform attachment scheme in which each existing node has an equal chance of recruiting a new entrant. For a detailed discussion of properties, see [7], [8], and [16].

Certain applications have encouraged new paths of research. It was observed that uniform attachment does not capture particular clustering features in real-world networks and does not give rise to the power laws and scale-free behavior that are characteristic of real applications. Real-world networks have features described by the Matthew principle, according to which ‘the rich get richer’ or ‘success breeds success’. Many of these features are generated by a degree-preferential attachment model, in which fat nodes of higher degree are given preference at recruiting new entrants. More precisely, a node recruits according to a probability proportional to its outdegree [17]. The model introduced by Barabási and Albert in [3] and discussed in [1] is slightly different, where the recruiting is done with probability proportional

---

Received 20 September 2019; revision received 4 May 2020.

\* Postal address: Department of Statistics, The George Washington University, Washington, DC 20052, USA.

\*\* Email address: [mlyon37@gwu.edu](mailto:mlyon37@gwu.edu)

\*\*\* Email address: [hosam@gwu.edu](mailto:hosam@gwu.edu)

to one plus the outdegree. The same Barabási–Albert scheme was rigorously developed for the first time in [4]. The random graphs in these sources are not limited to trees. For a broader discussion, see [2].

Until now, the term *preferential attachment* has been reserved for this degree-preference model. However, other applications can conceivably require us to consider an alternative preferential attachment mechanism. For example, in the bonding of atoms to form compounds, unsaturated atoms have higher affinity to other unsaturated atoms. Less saturated atoms are hungrier for bonding. This constitutes a preferential attachment scheme inversely proportional to the valence of an atom in a chemical compound.

Other preferential attachment schemes may depend on the dynamics of the tree growth. In a social network, affinity may be proportional to node age, whereby older nodes represent popularity and experience in social success. Or as an alternative model, affinity may be proportional to the youthfulness of a node, which models the eagerness of newcomers to expand their social network. Trees grown under old-age preference have been considered in the recent book [11]. Here we consider preferential attachment dominated by youth.

## 2. Young-age attachment trees

We study a random tree model that assumes that younger nodes have stronger affinity. This assumption is appropriate for networks in which new entrants tend to recruit more members. For example, frequently clubs offer new members a special deal – possibly discount memberships – that they can share with friends for a specified time. Or in the case of modeling disease transmission, for diseases with varying intensity of communicability, recently infected persons will be more likely to infect others.

In this model, a recursive tree grows in the following fashion. The tree starts at time 0 with one node; let us label it 1. At each subsequent point  $t \geq 1$  in discrete time, one of the nodes in the existing tree attracts (recruits) the  $n$ th ( $(t + 1)$ st) entrant in the system, and a node labeled  $n$  is attached to the recruiter, considered the parent. Thus  $t - i + 1$  is the amount of time node  $i$  spent living in the tree; we take it to be the age of node  $i$  by time  $t$ .

The tree grows in such a way that younger nodes have a higher probability of attracting new nodes. Concretely, the probability that the  $i$ th node ( $i = 1, \dots, n - 1$ ) is chosen as the parent of node  $n$  is proportional to the node index, which implies younger nodes have greater affinity. Letting  $\mathcal{A}_{i,n}$  be the event that the  $n$ th node attaches to the  $i$ th node, the proposed young-age attachment model is governed by the probability

$$\mathbb{P}(\mathcal{A}_{i,n}) = \frac{2i}{n(n-1)} \quad \text{for } 1 \leq i \leq n-1. \quad (2.1)$$

Figure 1 illustrates the six young-age preferential trees of size 4. The numbers in the top row are their probabilities.

## 3. Scope and organization

The manuscript is organized as follows. Distributional properties of the outdegree of node  $i = 1, 2, \dots, n$  are derived in Section 4. The mean and variance are established, and Poisson limit results are found via Chen–Stein Poisson approximation methods. Tree leaves are examined in Section 5, and limit results are obtained for the fraction of leaves in a given tree as the size of the tree grows. We find that the young-age probability model results in a fraction of leaves less than 1/2. Section 6 is devoted to the investigation of various properties of the insertion depth of node  $n$ , where it is found that the exact distribution involves Stirling numbers

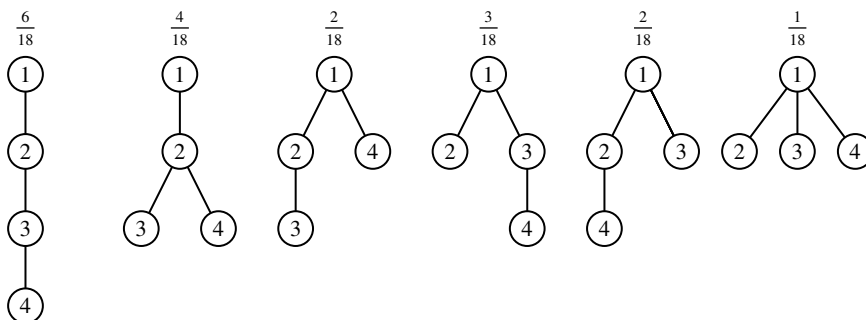


FIGURE 1: Young-age attachment recursive trees of size 4 and their probabilities.

of the first kind. Upon appropriate normalization, concentration laws and Gaussian limits are established. In the concluding remarks of Section 7, the young-age model is contrasted with the uniform case and areas of potential future research are discussed.

### 4. Outdegrees

Let  $\Delta_{i,n}$  be the *outdegree* of node  $i$  in a young-age attachment tree of size  $n$  – a count of the children of node  $i$  right after the node  $n$  joins the tree. If we take  $\mathbb{I}_{\mathcal{B}}$  to be the indicator of event  $\mathcal{B}$ , we can express the random variable  $\Delta_{i,n}$  as a sum of indicators for the events  $\mathcal{A}_{i,j}$  with  $j > i$ :

$$\Delta_{i,n} = \sum_{j=i+1}^n \mathbb{I}_{\mathcal{A}_{i,j}}.$$

#### 4.1. Exact moment generating function

The events  $\mathcal{A}_{i,j}$  are independent. Consequently, the moment generating function of  $\Delta_{i,n}$  is the product of the individual Bernoulli moment generating functions:

$$\phi_{\Delta_{i,n}}(t) = \prod_{j=i+1}^n \left( 1 + \frac{2i}{j(j-1)}(e^t - 1) \right). \tag{4.1}$$

4.1.1. *Mean and variance.* Using the moment generating function, we have the expectation and variance. In the presentation, we need the  $n$ th harmonic number of the second order,  $H_n^{(2)} = \sum_{k=1}^n 1/k^2$ .

**Proposition 4.1.** *Let  $\Delta_{i,n}$  be the outdegree of node  $i$  right after node  $n$  joins the young-age attachment tree. We have*

$$\begin{aligned} \mathbb{E}[\Delta_{i,n}] &= 2 \left( 1 - \frac{i}{n} \right), \\ \text{Var} [\Delta_{i,n}] &= 2 \left( 1 - \frac{i}{n} \right) - 4i^2 \left( 2H_n^{(2)} - 2H_i^{(2)} + \frac{2}{n} - \frac{2}{i} + \frac{1}{i^2} - \frac{1}{n^2} \right). \end{aligned}$$

*Proof.* Take the first derivative of the moment generating function with respect to  $t$ , and evaluate it at 0 using the chain rule, to get

$$\phi'_{\Delta_{i,n}}(0) = \mathbb{E}[\Delta_{i,n}] = \sum_{k=i+1}^n \frac{2i}{k(k-1)}.$$

Decomposing the summand into partial fractions, the sum telescopes:

$$\mathbb{E}[\Delta_{i,n}] = 2i \sum_{j=i+1}^n \left( \frac{1}{j-1} - \frac{1}{j} \right) = 2i \left( \frac{1}{i} - \frac{1}{n} \right) = 2 \left( 1 - \frac{i}{n} \right). \tag{4.2}$$

The variance of  $\Delta_{i,n}$  follows after computing the second derivative of the moment generating function evaluated at  $t = 0$ . To expedite the computation, write the derivative of the moment generating function using the chain rule

$$\begin{aligned} \frac{d}{dt} \phi_{\Delta_{i,n}}(t) &= \sum_{k=i+1}^n \frac{2ie^t}{k(k-1)} \times \frac{\phi_{\Delta_{i,n}}(t)}{\left(1 + \frac{2i}{k(k-1)}(e^t - 1)\right)} \\ &= 2ie^t \phi_{\Delta_{i,n}}(t) \sum_{k=i+1}^n \frac{1}{k(k-1)} \times \frac{1}{\left(1 + \frac{2i}{k(k-1)}(e^t - 1)\right)}. \end{aligned}$$

The second derivative at 0 is

$$\begin{aligned} \frac{d^2}{dt^2} \phi_{\Delta_{i,n}}(0) &= \mathbb{E}[\Delta_{i,n}^2] \\ &= 2i(e^t \phi'_{\Delta_{i,n}}(t) + e^t \phi_{\Delta_{i,n}}(t)) \Big|_{t=0} \sum_{k=i+1}^n \frac{1}{k(k-1)} \\ &\quad + 2ie^t \phi_{\Delta_{i,n}}(t) \sum_{k=i+1}^n \frac{1}{k(k-1)} \times \frac{-2ie^t/(k(k-1))}{\left(1 + \frac{2i}{k(k-1)}(e^t - 1)\right)^2} \Big|_{t=0} \\ &= 2i(\mathbb{E}[\Delta_{i,n}] + 1) \times \frac{1}{2i} \mathbb{E}[\Delta_{i,n}] - 4i^2 \sum_{k=i+1}^n \frac{1}{k^2(k-1)^2} \\ &= \mathbb{E}^2[\Delta_{i,n}] + \mathbb{E}[\Delta_{i,n}] - 4i^2 \sum_{k=i+1}^n \frac{1}{k^2(k-1)^2}. \end{aligned}$$

Using the form (4.2), we have

$$\begin{aligned} \text{Var} [\Delta_{i,n}] &= \mathbb{E}[\Delta_{i,n}^2] - \mathbb{E}^2[\Delta_{i,n}] \\ &= \mathbb{E}[\Delta_{i,n}] - 4i^2 \sum_{k=i+1}^n \frac{1}{k^2(k-1)^2} \\ &= 2 \left( 1 - \frac{i}{n} \right) - 4i^2 \sum_{k=i+1}^n \frac{1}{k^2(k-1)^2}. \end{aligned}$$

The variance as stated in Proposition 4.1 follows after expanding the summand in partial fractions and recognizing the harmonic numbers in the sums. □

4.1.2. *Interpretation.* We immediately see one way in which the behavior in young-age preferential attachment trees contrasts with the uniform attachment model. In young-age preferential attachment trees, for fixed  $i$ , the expected outdegree of node  $i$  remains small and bounded by 2, whereas in the uniform attachment tree the expected outdegree of node  $i$  (fixed) by the time node  $n$  joins is asymptotic to  $\ln n$ , and thus grows without a bound.

In the young-age preferential attachment tree, we see ‘phases’ depending on the relation of  $i$  to  $n$ . For ‘small’  $i$ , where  $i = o(n)$ , such as the case  $i = 7$  or  $i = \lceil 3 \ln n + 6 \rceil$ , we have  $\mathbb{E}[\Delta_{i,n}] \rightarrow 2$ . For  $i \sim \alpha n$  with  $0 < \alpha \leq 1$ , such as the example  $i = \lceil \frac{5}{7}n + 2\sqrt{n} + 12 \ln n + \pi^2 \rceil$ , we have  $\mathbb{E}[\Delta_{i,n}] \rightarrow 2(1 - \alpha)$ . Note that for the very late arriving nodes, corresponding to  $i \sim n$ , such as the case  $i = \lfloor n - 7n^{3/4} + \sqrt{n} \rfloor$ , we have  $\mathbb{E}[\Delta_{i,n}] \rightarrow 0$ .

One can notice a delicate balance. It is true that if  $n$  is large, the relative affinity of a young node with large index  $i$  is much higher than the old early nodes. However, the early nodes have stayed much longer in the tree and had repeated opportunities to recruit, particularly when they themselves were young and had high affinity. So, after all, it is not surprising that the early nodes have slightly higher expected outdegrees.

The variance  $\text{Var} [\Delta_{i,n}]$  is quite close to the expected value  $\mathbb{E}[\Delta_{i,n}]$ . Again, we have different phases depending on the relation between  $i$  and  $n$ . There are two phases for ‘small’  $i$  with  $i = o(n)$ . When  $i$  is fixed, we have

$$\text{Var} [\Delta_{i,n}] \rightarrow 8i + 8i^2 \left( \frac{\pi^2}{6} - H_i^{(2)} \right) - 2.$$

When  $i = o(n)$  and  $i \rightarrow \infty$ , we have

$$\text{Var} [\Delta_{i,n}] \rightarrow 2.$$

For  $i \sim \alpha n$ , with  $0 < \alpha \leq 1$ , we get  $\text{Var} [\Delta_{i,n}] \rightarrow 2(1 - \alpha)$ . Note again that for  $i \sim n$ , we have  $\text{Var} [\Delta_{i,n}] \rightarrow 0$ .

### 4.2. Limit distribution

The moment generating function (4.1) allows us to directly derive the main limit result of this section. However, we first outline some necessary background material concerning probability distance measures and Chen–Stein approximation.

4.2.1. *Preliminaries.* For any two random variables  $X$  and  $Y$ , respectively, with induced measures  $\mathbb{P}$  and  $\mathbb{Q}$  on the real line, we define the total variation distance

$$d_{\text{TV}}(X, Y) := \sup_{A \in \mathcal{B}} |\mathbb{P}(A) - \mathbb{Q}(A)|,$$

where  $A$  is an element of the Borel sigma algebra on  $\mathbb{R}$ .

It is well known that for discrete random variables, convergence in total variation distance is equivalent to convergence in distribution.

**Lemma 4.1.** ([10]) *For a sequence  $X_n$  of discrete random variables and a discrete random variable  $X$ , we have*

$$d_{\text{TV}}(X_n, X) \rightarrow 0 \iff X_n \xrightarrow{d} X.$$

This lays out a strategy for the proof. We aim to show that the total variation distance between  $X_n$  and a limiting random variable converges to 0, implying convergence in distribution.

A useful tool is a bound on the total variation distance between Poisson random variables. In what follows we denote a Poisson random variable with parameter  $\theta$  by  $\text{Poi}(\theta)$ .

**Lemma 4.2.** *If  $v > \mu$  and  $X$  and  $Y$  are respectively distributed like  $\text{Poi}(v)$  and  $\text{Poi}(\mu)$ , then we have*

$$d_{\text{TV}}(X, Y) \leq v - \mu.$$

*Proof.* We use  $\stackrel{d}{=}$  for equality in distribution. Let  $Y^*$  and  $Z$  be independent random variables with  $Y^* \stackrel{d}{=} \text{Poi}(\mu)$ ,  $Z \stackrel{d}{=} \text{Poi}(v - \mu)$  and define  $X^* = Y^* + Z \stackrel{d}{=} \text{Poi}(v)$ . By [10, Lemma 7.1] we have

$$d_{\text{TV}}(X^*, Y^*) \leq \mathbb{P}(X^* \neq Y^*) = \mathbb{P}(Z \neq 0) = 1 - \mathbb{P}(Z = 0) = 1 - e^{-(v-\mu)} \leq v - \mu.$$

Note that  $(X^*, Y^*)$  is a coupling of  $(X, Y)$ , so  $d_{\text{TV}}(X, Y) = d_{\text{TV}}(X^*, Y^*) \leq v - \mu$ . □

The main Chen–Stein result we use applies to approximating sums of independent Bernoulli random variables with a Poisson random variable [15].

**Proposition 4.2.** *Let  $X_1, \dots, X_n$  be independent Bernoulli random variables with  $\mathbb{P}(X_i = 1) = p_i$ . If  $W_n = \sum_{i=1}^n X_i$  and  $\lambda_n = \mathbb{E}[W_n] = \sum_{i=1}^n p_i$ , then we have*

$$d_{\text{TV}}(W_n, \text{Poi}(\lambda_n)) \leq \min\left\{1, \frac{1}{\lambda_n}\right\} \sum_{i=1}^n p_i^2 \leq \min\{1, \lambda_n\} \max_{1 \leq i \leq n} p_i.$$

*Proof.* See Theorem 4.6 of [15] □

4.2.2. *Limit laws.* The index  $i$  and the age  $n$  can be tied together in many complicated ways. For instance, one might think of a bizarre sequence in which

$$i = i(n) = \begin{cases} 2 & \text{if } n \text{ is even,} \\ \lfloor \sqrt{n} \rfloor & \text{if } n \text{ is odd.} \end{cases}$$

Such a sequence is not likely to appear in any application and does not lead to convergence. Like sequences are not within the scope of this investigation. We focus on relations between  $i$  and  $n$  with a systematic unifying theme for the growth of both. Namely, our investigation covers convergence relations that arise when  $i = i(n) = \alpha n + o(n)$ , for  $\alpha \in [0, 1]$ . So  $i$  has an infinite limit, except for the case  $\alpha = 0$ , where we require the  $o(n)$  term to have a non-oscillating leading term, such as, for example,  $i = 7$  or  $i = \lceil 3 \ln n + 6 \rceil$ .

**Theorem 4.1.** *As  $n \rightarrow \infty$ ,  $\Delta_{i,n}$  converges in distribution. Consider a node with index  $i = i(n) = \alpha n + o(n)$ , for  $0 \leq \alpha \leq 1$ . When  $\lim_{n \rightarrow \infty} i(n)$  exists, we have*

$$\Delta_{i,n} \xrightarrow{d} \Delta_i = \text{Poi}(2(1 - \alpha)) + Y_{\lim_{n \rightarrow \infty} i(n)},$$

where  $Y_{\lim_{n \rightarrow \infty} i(n)}$  is a zero-mean integer-valued random variable negatively correlated with the  $\text{Poi}(2(1 - \alpha))$  component.

For fixed  $i$ , with  $\alpha = 0$ , we have

$$\mathbb{E}[\Delta_i] = 2,$$

$$\text{Var} [\Delta_i] = 2 - \frac{4}{3} \pi^2 i^2 + 8i^2 H_i^{(2)} + 8i - 4,$$

$$0 < d_{\text{TV}}(\Delta_i, \text{Poi}(2)) \leq \frac{2}{i + 1},$$

whereas, for any growing  $i = i(n) \rightarrow \infty$ , we have  $Y_\infty = 0$ , almost surely.

*Proof.* For the first statement of the theorem, we work with a logarithmic transformation of  $\phi_{\Delta_{i,n}}(t)$ :

$$\ln \phi_{\Delta_{i,n}}(t) = \sum_{j=i+1}^n \ln \left( 1 + \frac{2i}{j(j-1)}(e^t - 1) \right).$$

We seek a Taylor series expansion of  $\ln(1+x)$  around zero for each summand. To make use of the expansion, we need

$$-1 < \frac{2i}{j(j-1)}(e^t - 1) < 1 \quad \text{for } j \geq 1.$$

These inequalities restrict the values of  $t$  for which the expansion is well-defined. The upper bounds require

$$t < \ln \left( 1 + \frac{j(j-1)}{2i} \right)$$

for  $j \geq 1$ , and the upper bounds are nested as  $j$  increases. So all the inequalities will be true when

$$t < \ln \left( 1 + \frac{(i+1)i}{2i} \right) = \ln \left( 1 + \frac{i+1}{2} \right),$$

and since  $i \geq 1$ , the expansion is well-defined for all  $i, j$ , when  $t < \ln 2$ .

With  $1 \leq i < j \leq n$ , for negative  $t$  the lower bound is automatically satisfied. In summary, Taylor series expansions are valid for all  $i$  and  $j$ , when  $t \in (-\infty, \ln 2)$ .

The next step is to show that  $\lim_n \phi_{\Delta_{i,n}}(t)$  is finite for all  $t \in (-\infty, \ln 2)$ . Trivially, we have  $\lim_{n \rightarrow \infty} \phi_{\Delta_{i,n}}(0) = 1$ . Next consider  $t \in (0, \ln 2)$ . Take a first-order Taylor series expansion around zero using the remainder form and apply the expression for the mean  $\mathbb{E}[\Delta_{i,n}]$ , to get

$$\begin{aligned} \ln \phi_{\Delta_{i,n}}(t) &= \sum_{j=i+1}^n \frac{2i}{j(j-1)}(e^t - 1) - \frac{1}{2} \sum_{j=i+1}^n \left( \frac{1}{1+x_{i,j}} \right)^2 \left( \frac{2i}{j(j-1)}(e^t - 1) \right)^2, \\ &= 2 \left( 1 - \frac{i}{n} \right) (e^t - 1) - 2i^2 (e^t - 1)^2 \sum_{j=i+1}^n \left( \frac{1}{1+x_{i,j}} \right)^2 \left( \frac{1}{j^2(j-1)^2} \right), \end{aligned}$$

for some

$$x_{i,j} \in \left( 0, \frac{2i}{j(j-1)}(e^t - 1) \right),$$

for all  $j \geq 1$ .

Evidently the series is convergent; let us define

$$h_{\lim_{n \rightarrow \infty} i(n)}(t) = \lim_{n \rightarrow \infty} -2 \sum_{j=i(n)+1}^n \left( \frac{1}{1+x_{i,j}} \right)^2 \left( \frac{i(n)}{j(j-1)}(e^t - 1) \right)^2,$$

and note that  $|h_{\lim_{n \rightarrow \infty} i(n)}(t)| < \infty$ .

Since

$$2 \left( 1 - \frac{i}{n} \right) (e^t - 1) \rightarrow 2(1 - \alpha)(e^t - 1),$$

we have

$$\ln \phi_{\Delta_{i,n}}(t) \rightarrow 2(1 - \alpha)(e^t - 1) + h_{\lim_{n \rightarrow \infty} i(n)}(t) < \infty.$$

An identical argument holds for the negative part of the interval and thus

$$\phi_{\Delta_{i,n}}(t) \rightarrow e^{2(1-\alpha)(e^t-1)} e^{h_{\lim_{n \rightarrow \infty} i(n)}(t)} \quad \text{for } t \in (-\infty, \ln 2).$$

As the convergence of  $\phi_{\Delta_{i,n}}(t)$  occurs over a finite open interval containing 0, by the Curtiss theorem [6], the limit function  $e^{2(1-\alpha)(e^t-1)} e^{h_{\lim_{n \rightarrow \infty} i(n)}(t)}$  is the moment generating function of *some* random variable, which we call  $\Delta_i$ .

From the structure of the limit moment generating function, the limiting random variable  $\Delta_i$  is ‘close’ to a  $\text{Poi}(2(1 - \alpha))$  random variable along with a perturbation random variable  $Y_{\lim_{n \rightarrow \infty} i(n)}$  with moment generating function  $e^{h_{\lim_{n \rightarrow \infty} i(n)}(t)}$ .

*Fixed i.* Let  $P_{i,n}$  be a sequence of totally independent Poisson random variables, with

$$P_{i,n} \stackrel{d}{=} \text{Poi}\left(2\left(1 - \frac{i}{n}\right)\right).$$

These will be the sequence of approximating random variables that we want to be ‘close’ to  $\Delta_{i,n}$ .

The triangle inequality holds for the total variation distance:

$$d_{\text{TV}}(\Delta_i, \text{Poi}(2)) \leq d_{\text{TV}}(\Delta_i, \Delta_{i,n}) + d_{\text{TV}}(\Delta_{i,n}, \text{Poi}(2)).$$

Using the triangle inequality a second time, we have an inequality valid for all  $n$ , which is namely

$$d_{\text{TV}}(\Delta_i, \text{Poi}(2)) \leq d_{\text{TV}}(\Delta_i, \Delta_{i,n}) + d_{\text{TV}}(\Delta_{i,n}, P_{i,n}) + d_{\text{TV}}(P_{i,n}, \text{Poi}(2)).$$

We now consider each term in order. Since  $\Delta_{i,n} \xrightarrow{d} \Delta_i$ , by Lemma 4.1,  $d_{\text{TV}}(\Delta_i, \Delta_{i,n}) \rightarrow 0$ . With  $\Delta_{i,n}$  being a sum of independent indicators, Proposition 4.2 is applicable, with  $\lambda_n = 2(1 - i/n)$ . From the upper inequality of Proposition 4.2, for large  $n$  we get

$$d_{\text{TV}}(\Delta_{i,n}, P_{i,n}) \leq \min\left\{1, 2\left(1 - \frac{i}{n}\right)\right\} \max_{i+1 \leq j \leq n} \frac{2i}{j(j-1)} = 1 \times \frac{2i}{i(i+1)} = \frac{2}{i+1}.$$

Finally, by Lemma 4.2, we have

$$d_{\text{TV}}(P_{i,n}, \text{Poi}(2)) \leq 2 - 2\left(1 - \frac{i}{n}\right) = \frac{2i}{n}.$$

Taken together, as  $n \rightarrow \infty$ , we have

$$0 < \lim_{n \rightarrow \infty} d_{\text{TV}}(\Delta_i, \text{Poi}(2)) \leq \lim_{n \rightarrow \infty} d_{\text{TV}}(\Delta_i, \Delta_{i,n}) + \lim_{n \rightarrow \infty} \frac{2}{i+1} + \lim_{n \rightarrow \infty} \frac{2i}{n}.$$

For fixed  $i$ , this becomes

$$0 < d_{\text{TV}}(\Delta_i, \text{Poi}(2)) \leq 0 + \lim_{n \rightarrow \infty} \frac{2}{i+1} + 0 = \frac{2}{i+1},$$

as claimed.



The mean and variance of the limiting distribution follows from uniform integrability arguments. Except in pathological cases, the limit of the individual random variable moments will pass through to the limit distribution. Furthermore, similar arguments show that the  $Y_{\lim_{i \rightarrow \infty} i(n)}$  component is negatively correlated with the Poisson piece.

Growing  $i$ . Consider  $i = i(n) = \alpha n + o(n)$ , with  $\alpha \in (0, 1]$ . By the triangle inequality, we have

$$d_{TV}(\Delta_{i,n}, \text{Poi}(2(1 - \alpha))) \leq d_{TV}(\Delta_{i,n}, P_{i,n}) + d_{TV}(P_{i,n}, \text{Poi}(2(1 - \alpha))).$$

By Proposition 4.2 and Lemma 4.2, we obtain

$$d_{TV}(\Delta_{i,n}, \text{Poi}(2(1 - \alpha))) \leq \frac{2}{i + 1} + 2 \left| \alpha - \frac{i}{n} \right|.$$

As  $n \rightarrow \infty$ , we have the convergence  $2/(i + 1) \rightarrow 0$ , since  $i \rightarrow \infty$  and  $i/n \rightarrow \alpha$ . So we have

$$d_{TV}(\Delta_{i,n}, \text{Poi}(2(1 - \alpha))) \rightarrow 0.$$

Thus we have  $\Delta_{i,n} \xrightarrow{d} \text{Poi}(2(1 - \alpha))$  and  $Y_\infty = 0$ . □

**Remark 4.1.** The limit law in Theorem 4.1 combines all the phases discussed in the interpretations in Section 4.1.2 and better explains how they merge at the ‘seam lines’. One can say that

$$\Delta_{i,n} \xrightarrow{d} \text{Poi}(2(1 - \alpha)) + Y_{\lim_{n \rightarrow \infty} i(n)}.$$

In the very early phase (fixed  $i$ ),  $\alpha = 0$ , and the role of  $Y_{\lim_{n \rightarrow \infty} i(n)} = Y_i$  is pronounced. The assertion that, for fixed  $i$ , we have

$$0 < d_{TV}(\Delta_i, \text{Poi}(2)) \leq \frac{2}{i + 1}$$

is actually a qualifying description of the random variable  $Y_i$ .

Then, for intermediate  $i$ , growing to infinity, but remaining  $o(n)$ ,  $\alpha$  is still 0, and the effect of  $Y_{\lim_{n \rightarrow \infty} i(n)} = Y_\infty = 0$  disappears. An instance of this case is  $i = \lceil 3 \ln n + 6 \rceil$ .

In the linear phase  $i \sim \alpha n$ , with  $0 < \alpha < 1$ , the Poisson parameter is attenuated. An instance of this case is  $i = \lceil \frac{5}{7}n + 2\sqrt{n} + 12 \ln n + \pi^2 \rceil$ . Ultimately the Poisson parameter becomes 0, when  $i \sim n$ . Here the limit is a  $\text{Poi}(0)$  random variable, which is identically 0. An instance of this case is  $i = \lfloor n - 7n^{3/4} + \sqrt{n} \rfloor$ .

**Remark 4.2.** It is assumed in Theorem 4.1 that  $\lim_{n \rightarrow \infty} i(n)$  exists. If this limit does not exist, as in the example discussed at the beginning of this subsection, the interpretation is that  $\Delta_{i,n}$  does not converge in distribution at all to any limit. Rather, it oscillates.

### 5. Leaves

Let  $\mathbb{I}_{i,n}$  be an indicator equal to 1 when node  $i$  is a leaf in a tree of size  $n$  and 0 otherwise. Define  $L_n$  to be the number of leaves in a tree of size  $n$ , which can be expressed as the sum of the individual leaf indicators:

$$L_n = \sum_{i=1}^n \mathbb{I}_{i,n}.$$

Under the young-age attachment model, the root is never a leaf in a tree with size  $n > 1$ ; node  $n$  is always a leaf in any size tree. Since node attachments are independent, we can directly calculate the mean and variance of the leaf indicators.

**Lemma 5.1.**

$$\mathbb{E}[\mathbb{L}_{i,n}] = \prod_{j=i+1}^n \left(1 - \frac{2i}{j(j-1)}\right),$$

and

$$\text{Var} [\mathbb{L}_{i,n}] = \prod_{j=i+1}^n \left(1 - \frac{2i}{j(j-1)}\right) \left[1 - \prod_{j=i+1}^n \left(1 - \frac{2i}{j(j-1)}\right)\right].$$

*Proof.* For node  $i$  to be a leaf in a tree of size  $n > i$ , it must be passed over by every node  $j > i$ , which each occurs with probability  $1 - 2i/(j(j-1))$ . These attachments are independent, whence the joint probability is the product of the individual probabilities:

$$\mathbb{E}[\mathbb{L}_{i,n}] = \mathbb{P}(\mathbb{L}_{i,n} = 1) = \prod_{j=i+1}^n \left(1 - \frac{2i}{j(j-1)}\right).$$

The variance follows immediately since the  $\mathbb{L}_{i,n}$  are Bernoulli random variables. □

We have a complete characterization of the *marginal* distributions of the leaf indicators in a tree of size  $n > 1$ :

$$\begin{aligned} \mathbb{L}_{1,n} &= 0 \quad \text{with probability 1,} \\ \mathbb{L}_{i,n} &= \text{Bernoulli} \left( \prod_{j=i+1}^n \left(1 - \frac{2i}{j(j-1)}\right) \right) \quad \text{for } 1 < i < n, \\ \mathbb{L}_{n,n} &= 1 \quad \text{with probability 1.} \end{aligned}$$

**5.1. Joint distribution**

While node attachments are independent, the resulting  $\mathbb{L}_{i,n}$  are dependent. Nevertheless, the structure of the leaf indicators allows us to derive useful covariance results, which we use to study the asymptotic behavior. Throughout this section we state the results in a more general form since they hold for any sequence of attachment distributions.

Recall that  $\mathcal{A}_{i,n}$  denotes the event that the  $n$ th node attaches to the  $i$ th node. Assume now that for node  $n \geq 2$ , we have a general attachment model:

$$\mathbb{P}(\mathcal{A}_{i,n}) = f_n(i) \quad \text{for } 1 \leq i \leq n - 1,$$

with  $f_n(i) \geq 0$  and  $\sum_{i=1}^{n-1} f_n(i) = 1$ .

In principle this allows for arbitrary individual attachment probabilities for each node at each time of entry. We can directly calculate the cross-expectation of the  $\mathbb{L}_{i,n}$ .

**Lemma 5.2.** *With  $i < j$ , we have*

$$\mathbb{E}[\mathbb{L}_{i,n}\mathbb{L}_{j,n}] = \prod_{p=i+1}^j (1 - f_p(i)) \prod_{p=j+1}^n (1 - f_p(i) - f_p(j)).$$

*Proof.* Assume without loss of generality that  $i < j$ . The product  $\mathbb{L}_{i,n}\mathbb{L}_{j,n}$  is equal to one when node  $i$  is passed over by the nodes  $p = i + 1$  to  $j$ , and then *both* nodes  $i$  and  $j$  are passed over for nodes  $p = j + 1$  to  $n$ . The first sequence of events each happen with probability  $1 - f_p(i)$  and the latter sequence of events each occur with probability  $1 - f_p(i) - f_p(j)$ . All of these node attachments are independent, so the joint probability is the product of probabilities for the individual events.  $\square$

The cross-expectation allows us to find an upper bound for the covariance of the leaf indicators.

**Theorem 5.1.** *For  $i \neq j$ , we have*

$$\text{Cov} [\mathbb{L}_{i,n}, \mathbb{L}_{j,n}] \leq 0.$$

*Proof.* The rightmost expression for the cross-expectation in Lemma 5.2 is

$$\prod_{p=j+1}^n (1 - f_p(i) - f_p(j)).$$

We note for each  $p$ :

$$1 - f_p(i) - f_p(j) \leq 1 - f_p(i) - f_p(j) + f_p(i)f_p(j) = (1 - f_p(i))(1 - f_p(j)).$$

Thus we have

$$\prod_{p=j+1}^n (1 - f_p(i) - f_p(j)) \leq \prod_{p=j+1}^n (1 - f_p(i))(1 - f_p(j)).$$

Substituting this inequality into the cross-expectation of Lemma 5.2 and collecting the  $i$  and  $j$  terms, we obtain

$$\begin{aligned} \mathbb{E}[\mathbb{L}_{i,n}\mathbb{L}_{j,n}] &\leq \prod_{p=i+1}^j (1 - f_p(i)) \prod_{p=j+1}^n (1 - f_p(i))(1 - f_p(j)) \\ &= \prod_{p=i+1}^n (1 - f_p(i)) \prod_{p=j+1}^n (1 - f_p(j)) \\ &= \mathbb{E}[\mathbb{L}_{i,n}] \mathbb{E}[\mathbb{L}_{j,n}], \end{aligned}$$

which implies that the covariance is non-positive.  $\square$

**5.2. Mean and variance**

Specifying the young-age preferential model, we use the cross-expectation in Lemma 5.2 to express the covariance between leaves with  $i < j$ .

**Corollary 5.1.**

$$\begin{aligned} \text{Cov} [\mathbb{L}_{i,n}, \mathbb{L}_{j,n}] &= \prod_{p=i+1}^j \left(1 - \frac{2i}{p(p-1)}\right) \prod_{p=j+1}^n \left(1 - \frac{2(i+j)}{p(p-1)}\right) \\ &\quad - \prod_{p=i+1}^n \left(1 - \frac{2i}{p(p-1)}\right) \prod_{p=j+1}^n \left(1 - \frac{2j}{p(p-1)}\right). \end{aligned}$$

With the expression for the expectation of  $\mathbb{L}_{i,n}$  in hand, we can calculate the mean and variance for  $L_n$ , when  $n > 1$ :

$$\begin{aligned} \mathbb{E}[L_n] &= \sum_{i=1}^n \mathbb{E}[\mathbb{L}_{i,n}] = \sum_{i=2}^{n-1} \prod_{j=i+1}^n \left(1 - \frac{2i}{j(j-1)}\right) + 1, \\ \text{Var}[L_n] &= \sum_{i=2}^{n-1} \text{Var}[\mathbb{L}_{i,n}] + 2 \sum_{i=2}^{n-2} \sum_{j=i+1}^{n-1} \text{Cov}[\mathbb{L}_{i,n}, \mathbb{L}_{j,n}] \\ &= \sum_{i=2}^{n-1} \prod_{j=i+1}^n \left(1 - \frac{2i}{j(j-1)}\right) \left[1 - \prod_{j=i+1}^n \left(1 - \frac{2i}{j(j-1)}\right)\right] \\ &\quad + 2 \sum_{i=2}^{n-2} \sum_{j=i+1}^{n-1} \text{Cov}[\mathbb{L}_{i,n}, \mathbb{L}_{j,n}], \end{aligned}$$

where  $\text{Cov}[\mathbb{L}_{i,n}, \mathbb{L}_{j,n}]$  is the covariance in Corollary 5.1.

### 5.3. Fraction of leaves

One of our main interests is in the fraction of leaves in a tree of size  $n$  as  $n \rightarrow \infty$ . We first find the limit of the expected fraction of leaves.

**Lemma 5.3.**

$$\mathbb{E}\left[\frac{L_n}{n}\right] \rightarrow \frac{1}{2} \left(1 - \frac{1}{e^2}\right) = 0.432332 \dots$$

*Proof.* For  $i > 1$ , let

$$\ell_{i,n} = \ln(\mathbb{E}[\mathbb{L}_{i,n}]) = \ln\left(\prod_{j=i+1}^n \left(1 - \frac{2i}{j(j-1)}\right)\right).$$

Using an exact first-order Taylor series expansion of  $\ln(1 - x)$  around zero, we have

$$\begin{aligned} \ell_{i,n} &= \sum_{j=i+1}^n \ln\left(1 - \frac{2i}{j(j-1)}\right) \\ &= - \sum_{j=i+1}^n \frac{2i}{j(j-1)} - \sum_{j=i+1}^n \frac{1}{(1 - x_{i,j})^2} \left(\frac{2i}{j(j-1)}\right)^2 \\ &= -2\left(1 - \frac{i}{n}\right) - \sum_{j=i+1}^n \frac{1}{(1 - x_{i,j})^2} \left(\frac{2i}{j(j-1)}\right)^2, \end{aligned}$$

where each

$$x_{i,j} \in \left(0, \frac{2i}{j(j-1)}\right).$$

This Taylor series expansion makes it possible to get useful bounds for each  $\ell_{i,n}$  term valid for all  $n > 0$ . Note that the second term of  $\ell_{i,n}$  is negative, so we immediately have the upper bound

$$\ell_{i,n} \leq -2\left(1 - \frac{i}{n}\right).$$

For the lower bound, examine each term in the second summation, and observe that over the range of  $j$  in the sum we have

$$\frac{1}{(1 - x_{i,j})^2} \left( \frac{2i}{j(j-1)} \right)^2 \leq \frac{1}{\left(1 - \frac{2i}{i(i+1)}\right)^2} \left( \frac{2i}{ji} \right)^2.$$

With this bound being valid for  $i > 1$ , we establish

$$\frac{1}{(1 - x_{i,j})^2} \left( \frac{2i}{j(j-1)} \right)^2 \leq \frac{1}{\left(1 - \frac{2}{(2+1)}\right)^2} \left( \frac{2}{j} \right)^2 \leq \frac{36}{i^2}.$$

We can now use these bounds to sandwich the fraction of leaves:

$$\frac{1}{n} \sum_{i=2}^n e^{-2(1-i/n)-36/i^2} \leq \frac{\mathbb{E}[L_n]}{n} = \frac{1}{n} \sum_{i=2}^n e^{\ell_{i,n}} \leq \frac{1}{n} \sum_{i=1}^n e^{-2(1-i/n)}.$$

The upper bound can be simplified by summing it as a geometric series, which gives

$$\begin{aligned} \frac{\mathbb{E}[L_n]}{n} &\leq \frac{e^{-2}}{n} \sum_{i=1}^n e^{2i/n} \\ &= \frac{e^{-2}}{n} \left( \frac{e^{2(n+1)/n} - 1}{e^{2/n} - 1} \right) \\ &= \frac{e^{-2}}{n} \left( \frac{e^{2(n+1)/n} - 1}{(1 + 2/n + O(n^{-2})) - 1} \right) \\ &\rightarrow \frac{1}{2} \left( 1 - \frac{1}{e^2} \right). \end{aligned}$$

The extraction of the asymptotics from the lower bound is a little more involved. First expand the term  $\exp(-36/i^2)$  in series with an explicit first term and an error of the exact order  $1/i^2$ . This means that there is a starting index value  $i_0$  and a positive constant  $K$  such that the error is at most  $K/i^2$ , for all  $i \geq i_0$ . We can then split the series at  $i_0$ , getting

$$\begin{aligned} \frac{\mathbb{E}[L_n]}{n} &\geq \frac{1}{n} \sum_{i=2}^n e^{-2(1-i/n)-36/i^2} \\ &\geq \frac{1}{n} \left( \sum_{i=2}^{i_0-1} e^{-2(1-i/n)} \left( 1 - \frac{K_i}{i^2} \right) + \sum_{i=i_0}^n e^{-2(1-i/n)} \left( 1 - \frac{K}{i^2} \right) \right) \\ &= O\left(\frac{1}{n}\right) + \frac{e^{-2}}{n} \left( \frac{e^{2(n+1)/n} - e^{2i_0/n}}{e^{2/n} - 1} \right) \\ &= O\left(\frac{1}{n}\right) + \frac{e^{-2}}{n} \left( \frac{e^{2(n+1)/n} - e^{2i_0/n}}{(1 + 2/n + O(n^{-2})) - 1} \right) \\ &= O\left(\frac{1}{n}\right) + e^{-2} \left( \frac{e^{2(n+1)/n} - e^{2i_0/n}}{2 + O(n^{-1})} \right) \\ &\rightarrow \frac{1}{2} \left( 1 - \frac{1}{e^2} \right). \end{aligned}$$

Given the convergence of the upper and lower bounds, the sandwich theorem ascertains that

$$\frac{\mathbb{E}[L_n]}{n} \rightarrow \frac{1}{2} \left( 1 - \frac{1}{e^2} \right). \quad \square$$

We can strengthen this expected fraction result into convergence in  $\mathcal{L}^2$ , and thus in probability, too. We need the following result from the theory of consistent estimators in statistics.

**Lemma 5.4.** *Let  $\{X_n\}$  be a sequence of random variables with  $\mathbb{E}[X_n^2] < \infty$ , for all  $n \geq 1$ . If  $\mathbb{E}[X_n] \rightarrow \mu \in \mathbb{R}$  and  $\text{Var}[X_n] \rightarrow 0$ , then  $X_n \xrightarrow{\mathcal{L}^2} \mu \in \mathbb{R}$ .*

*Proof.* See Theorem 8.2 of [12, pages 54–55]. □

We also use the fact that for random variables with bounded variance and non-positive correlations, the variability of the average goes to zero.

**Lemma 5.5.** *Let  $\{B_n\}$  be a sequence of random variables with non-positive correlations and  $\text{Var}[B_n] \leq M$ , for some  $M > 0$ , then*

$$\text{Var} \left[ \sum_{i=1}^n \frac{B_i}{n} \right] \rightarrow 0.$$

*Proof.* For each  $B_n$  we have  $\text{Var}[B_n] \leq M$  and  $\text{Cov}[B_i, B_j] \leq 0$  for  $i \neq j$ , so we have

$$\text{Var} \left[ \sum_{i=1}^n \frac{B_i}{n} \right] = \frac{1}{n^2} \left( \sum_{i=1}^n \text{Var}[B_i] + \sum_{i \neq j} \text{Cov}[B_i, B_j] \right) \leq \frac{M}{n}$$

and

$$\lim_{n \rightarrow \infty} \text{Var} \left[ \sum_{i=1}^n \frac{B_i}{n} \right] = 0. \quad \square$$

Finally, this gives us our main result.

**Theorem 5.2.**

$$\frac{L_n}{n} \xrightarrow{\mathcal{L}^2} \frac{1}{2} \left( 1 - \frac{1}{e^2} \right) = 0.432332 \dots$$

*Proof.* The result follows immediately by noting the non-positive covariance between the  $\mathbb{L}_{i,n}$ , as in Theorem 5.1, and applying the previous two lemmas to  $L_n/n$ . □

From the  $\mathcal{L}^2$  convergence we get the following.

**Corollary 5.2.**

$$\frac{L_n}{n} \xrightarrow{a.s.} \frac{1}{2} \left( 1 - \frac{1}{e^2} \right).$$

*Proof.* In-probability convergence is immediate since  $\mathcal{L}^p$  convergence for any  $p > 0$  implies in-probability. The almost-sure result follows from the fact that in-probability convergence is equivalent to almost-sure convergence for countable probability spaces (see [5, Chapter 2, Section 2] for a discussion). □

### 6. Insertion depth

In this section we derive results about random insertion depth under young-age preferential attachment. We let  $D_n$  denote the depth of the node labeled  $n$  in a young-age attachment tree. We first find the exact moment generating function, from which the exact distribution and a miscellany of asymptotic results follow.

Let  $T_n$  be the tree existing right after node  $n$  enters. We condition on  $T_{n-1}$ , for some  $t \in \mathbb{R}$ , and use the young-age probability model (2.1) to write

$$\mathbb{E}[e^{D_n t} | T_{n-1}] = \sum_{k=1}^{n-1} e^{(D_k+1)t} \frac{2k}{n(n-1)}.$$

Taking expectation of both sides and using the law of iterated expectations, we arrive at a recursive equation for the moment generating function of  $D_n$ :

$$\phi_{D_n}(t) = \sum_{k=1}^{n-1} \mathbb{E}[e^{(D_k+1)t}] \frac{2k}{n(n-1)} = 2e^t \sum_{k=1}^{n-1} \phi_{D_k}(t) \frac{k}{n(n-1)}.$$

Differencing, we obtain

$$n(n-1)\phi_{D_n}(t) - (n-1)(n-2)\phi_{D_{n-1}}(t) = 2e^t(n-1)\phi_{D_{n-1}}(t).$$

After refactoring, we have

$$\phi_{D_n}(t) = \left( \frac{n-2+2e^t}{n} \right) \phi_{D_{n-1}}(t).$$

Iterating this relation back, we get

$$\phi_{D_n}(t) = \left( \prod_{j=1}^{n-2} \frac{j+2e^t}{j+2} \right) \phi_{D_2}(t).$$

We have  $\phi_{D_2}(t) = e^t$  and  $\phi_{D_1}(t) = 1$  as boundary conditions. It follows that

$$\phi_{D_n}(t) = \frac{\prod_{j=1}^{n-2} (j+2e^t)}{n!} 2e^t = \frac{1}{n!} \prod_{j=0}^{n-2} (j+2e^t) \quad \text{for } n \geq 2.$$

We express the the next theorem in terms of the unsigned Stirling numbers of the first kind, which are coefficients in the polynomial representing the rising factorial:

$$x(x+1) \cdots (x+n-1) = \sum_{k=0}^n \left[ \begin{matrix} n \\ k \end{matrix} \right] x^k,$$

**Theorem 6.1.** *Let  $D_n$  be the depth of the  $n$ th node in a young-age attachment tree. The exact distribution of this insertion depth is*

$$\mathbb{P}(D_n = k) = \frac{2^k}{n!} \left[ \begin{matrix} n-1 \\ k \end{matrix} \right], \quad k = 1, \dots, n-1.$$

*Proof.* Using the moment generating function presented right before the theorem, we can derive the exact probability mass function of  $D_n$ . Setting  $t = \ln u$  in the moment generating function switches the view to the probability generating function  $\zeta_n(u)$ . Namely, it is

$$\zeta_n(u) = \phi_n(\ln u) = \frac{2u(2u + 1) \cdots (2u + n - 2)}{n!}.$$

Hence we have

$$\zeta_n(u) = \frac{1}{n!} \sum_{k=0}^{n-1} \binom{n-1}{k} (2u)^k.$$

The exact distribution of  $D_n$  is obtained by extracting coefficients. □

The exact distribution of the  $n$ th insertion depth in young-age recursive trees is the same as its counterpart in random binary search trees. Hence a number of results follow immediately, which we list without proof. The interested reader can examine Chapter 2 in [13]. In the following proposition, all asymptotics are taken as  $n \rightarrow \infty$ .

**Proposition 6.1.** *Let  $D_n$  be the depth of insertion of the  $n$ th node in a young-age preferential attachment tree. We have*

$$\begin{aligned} \mathbb{E}[D_n] &= 2(H_n - 1) \sim 2 \ln n, \\ \text{Var}[D_n] &= 2(H_n + 1) - 4H_n^{(2)} \sim 2 \ln n, \\ \frac{D_n}{\ln n} &\xrightarrow{\mathcal{L}^2} 2, \quad \frac{D_n}{\ln n} \xrightarrow{a.s.} 2, \\ \frac{D_n - 2 \ln n}{\sqrt{2 \ln n}} &\xrightarrow{d} N(0, 1). \end{aligned}$$

### 7. Conclusion

The non-uniform young-age attachment model introduced has both similarities to and differences from the standard uniform model. As seen in Proposition 6.1, the average and variance of the insertion depth in these trees is *double* that of the uniform case – and both random trees live in the ‘small world’. Like the uniform model, the insertion depth in the young-age model is also asymptotically normal (Proposition 6.1).

In the case of tree leaves and outdegrees, the young-age model differs substantially from the uniform model. The asymptotic fraction of leaves in young-age attachment trees is significantly less than 1/2 (the asymptotic fraction in uniform recursive trees), while in the limit the outdegrees are Poisson or ‘near-Poisson’. These facts imply that the nodes in the young-age model have a better chance of recruiting than the uniform case, but they tend to have ‘sparse’ connections with fewer outnode links.

The young-age model generates some natural directions for future research. Does the behavior carry over to probability models driven by different powers of the node index? Such an attachment mechanism could be used to model networks with varying degrees of age-based affinity. For example, one can work with a general class of growth models in which

$$\mathbb{P}(\mathcal{A}_{i,n}) \propto i^\alpha \quad \text{for } 1 \leq i \leq n - 1, \quad \alpha \in \mathbb{R}.$$



The growth index  $\alpha$  would be a parameter of interest and could be estimated from empirical data under different sampling schemes. Preliminary work suggests that, as  $\alpha$  grows, asymptotic normality remains for the insertion depth albeit with a different character of the variance. For large  $\alpha$ , the growth model approaches the complete linear chaining of the nodes.

### Acknowledgement

We thank the referees for thoughtful reports and constructive criticism.

### References

- [1] ALBERT, R. AND BARABÁSI, A. (2002). Statistical mechanics of complex networks. *Rev. Modern Phys.* **41**, 47–97.
- [2] BARABÁSI, A. (2016). *Network Science*. Cambridge University Press, Cambridge.
- [3] BARABÁSI, A. AND ALBERT, R. (1999). Emergence of scaling in random networks. *Science* **286**, 509–512.
- [4] BOLLOBÁS, B., RIORDAN, O. AND SPENCER, J. (2001). The degree sequence of a scale-free random graph process. *Random Structures Algorithms* **18**, 279–290.
- [5] CHOW, Y. AND TEICHER, M. (1997). *Probability Theory: Independence, Interchangeability, Martingales*. Springer, New York.
- [6] CURTISS, J. (1942). A note on the theory of moment generating functions. *Ann. Math. Statist.* **13**, 430–433.
- [7] DRMOTA, M. (2009). *Random Trees: An Interplay Between Combinatorics and Probability*. Springer, New York.
- [8] FRIEZE, A. AND KAROŃSKI, M. (2016). *Introduction to Random Graphs*, 2nd edn. Cambridge University Press, Cambridge.
- [9] GASTWIRTH, J. AND BHATTACHARYA, P. (1984). Two probability models of pyramid or chain letter schemes demonstrating that their promotional claims are unreliable. *Operat. Res.* **32**, 527–536.
- [10] GUT, A. (2013). *Probability: A Graduate Course*, 2nd edn. Springer, New York.
- [11] HOFRI, M. AND MAHMOUD, H. (2018). *Algorithmics of Nonuniformity: Tools and Paradigms*. CRC Press, Boca Raton.
- [12] LEHMANN, E. AND CASELLA, G. (1998). *Theory of Point Estimation*, 2nd edn. Springer, New York.
- [13] MAHMOUD, H. (1992). *Evolution of Random Search Trees*. Wiley, New York.
- [14] NAJOCK, D. AND HEYDE, C. (1982). On the number of terminal vertices in certain random trees with an application to stemma construction in philology. *J. Appl. Prob.* **19**, 675–680.
- [15] ROSS, N. (2011). Fundamentals of Stein’s method. *Prob. Surv.* **8**, 210–293.
- [16] SMYTHE, R. AND MAHMOUD, H. (1995). A survey of recursive trees. *Theory Prob. Math. Statist.* **51**, 1–27.
- [17] SMYTHE, R., MAHMOUD, H. AND SZYMAŃSKI, J. (1993). On the structure of plane-oriented recursive trees and their branches. *Random Structures Algorithms* **4**, 151–176.