Network Science (2025), **13**, e5, 1–20
doi:10.1017/nws.2025.3

CAMBRIDGE UNIVERSITY PRESS

RESEARCH ARTICLE

Exponentially preferential trees

Rafik Aguech[1], Hosam Mahmoud[2], Hanene Mohamed[3] and Zhou Yang[2]

[1]Laboratory AGA, University of Monastir, Monastir, Tunisia, [2]Department of Statistics, The George Washington University, Washington, USA, and [3]Modal'X, UPL, Université Paris Nanterre, Nanterre, France
Corresponding author: Hosam Mahmoud; Email: hosam@gwu.edu

Abstract

We introduce the exponentially preferential recursive tree and study some properties related to the degree profile of nodes in the tree. The definition of the tree involves a radix $a > 0$. In a tree of size n (nodes), the nodes are labeled with the numbers $1, 2, \ldots, n$. The node labeled i attracts the future entrant $n+1$ with probability proportional to a^i.
We dedicate an early section for algorithms to generate and visualize the trees in different regimes. We study the asymptotic distribution of the outdegree of node i, as $n \to \infty$, and find three regimes according to whether $0 < a < 1$ (subcritical regime), $a = 1$ (critical regime), or $a > 1$ (supercritical regime). Within any regime, there are also phases depending on a delicate interplay between i and n, ramifying the asymptotic distribution within the regime into "early," "intermediate" and "late" phases. In certain phases of certain regimes, we find asymptotic Gaussian laws. In certain phases of some other regimes, small oscillations in the asymototic laws are detected by the Poisson approximation techniques.

Keywords: Random structure; recursive tree; affinity; Gaussian law; Poisson approximation

AMS subject classifications: 60F05; 60G20; 60F05

1. Scope

The recursive tree is a classic hierarchical structure. Several models of randomness are used in a variety of applications. Dozens of research papers have been devoted to the uniform model alone, many of them are surveyed in Smythe and Mahmoud (1995). Today, the uniform recursive tree is a standard entry in books on random structures (Drmota, 2009; Hofri and Mahmoud, 2018; Frieze and Karoński, 2016). The uniform recursive tree is used as a model in many applications. Some of the classic applications are in pyramid schemes (Gastwirth and Bhattacharya, 1984) and Philology (Najock and Heyde, 1982).

Driven by other applications, interest was developed into nonuniform models, wherein the attachment of new nodes is done preferentially according to some criterion. The earliest of these preferential models is a probability scheme in which nodes of higher outdegrees are favored (Szymański, 1987). Other preferential ideas are based on node fitness (Dereich and Ortgiese, 2014), old age of nodes (Hofri and Mahmoud, 2018), node Youthfulness (Lyon and Mahmoud, 2020), and power-weights on the nodes (Lyon and Mahmoud, 2022).

1.1 A proposed preferential model

In the present investigation, we propose a new preferential model parameterized by a real positive radix, which we call a. In this exponential model, the affinity of a node is the radix raised to the node label. A precise definition is given in Subsection 1.3.

For instance, if the radix is less than 1, nodes that appear first (older in the tree) have an attraction power that is larger than newer nodes. One sees such a phenomenon in the growth of networks, where older nodes have a bigger chance of growth than younger ones. For instance, in a graph representing the growth of technology companies, nodes representing giants like Microsoft© and Apple© have a higher chance of attracting new subscribers than a node representing a small start-up company.

We assume that the reader is familiar with the jargon of trees, such as "node," "vertex," "edge," "root," "ancestors," "descendants," "children," "parents," "recruiting," "affinity," etc.

1.2 The building algorithm

A recursive tree is grown by an attachment algorithm that operates in the following way. Initially (at time $n = 0$), there is a node labeled 1. At each subsequent epoch $n \geq 1$ of discrete time, a node labeled $n + 1$ joins the tree by choosing one of the existing nodes as a parent and attaching itself to it via a new edge. The parent selection is determined according to some probability model on the set $\{1, \dots, n\}$. Note that the labels on any root-to-leaf path are in an increasing sequence. For this reason, some authors call these structures "increasing trees" (Bergeron et al., 1992).

These trees have been studied under several probability models, notably including the natural uniform model and preferential models based on favoring certain nodes according to some criterion. The first preferential criterion in the literature is to select a node with probability proportional to 1 plus its outdegree (Szymański, 1987; Mahmoud et al., 1993). This model gained popularity, as it offers scalability properties and power laws that are met in certain trees in nature (Barabási and Albert, 1999). For over a decade, the terminology "preferential attachment" stood solely for preference by node outdegrees.

More recently, authors broke away from this narrower definition of "preference" to tree models with other types of preference (Hofri and Mahmoud, 2018; Lyon and Mahmoud, 2020; Lyon and Mahmoud, 2022).

1.3 Exponentially preferential trees

In this investigation, we consider a new exponentially preferential attachment algorithm, wherein node i at time $n - 1$ recruits with probability proportional to a^i, for some positive constant a. Specifically, if we define $A_{n,i}$ as the event that node i recruits (the node labeled $n + 1$) when the tree has n nodes in it (that is, when the tree is of age $n - 1$), we would have

$$\mathbb{P}(A_{n,i}) = \frac{a^i}{\sum_{j=1}^{n} a^j} = \begin{cases} \dfrac{1}{n}, & \text{if } a = 1; \\[2ex] \dfrac{(a-1)\, a^{i-1}}{a^n - 1}, & \text{otherwise.} \end{cases} \tag{1}$$

In the sequel, we observe a trichotomy of the positive real line into three regimes for a, and in each regime we have a different behavior. We call the regime $0 < a < 1$ *subcritical*, the regime $a = 1$ *critical*, and the regime $a > 1$ *supercritical*.

We call a tree grown according to this distribution for the choice of parent an *exponentially preferential tree with radix a*. Since this is the only kind we study in this manuscript, we refer to it simply as the "tree" most of the time. When $a = 1$, we have the special case of uniform recruiting, which is extensively studied (Bergeron et al., 1992; Drmota, 2009; Hofri and Mahmoud, 2018; Frieze and Karoński, 2016; Smythe and Mahmoud, 1995).

Figure 1 displays the six exponentially preferential trees of size 4 with radix $a = 1/2$. The numbers above the trees are their probabilities. Note the high probability assigned to the bushiest tree at the far right. In the uniform model, this tree only has probability $1/6$. To discern the these

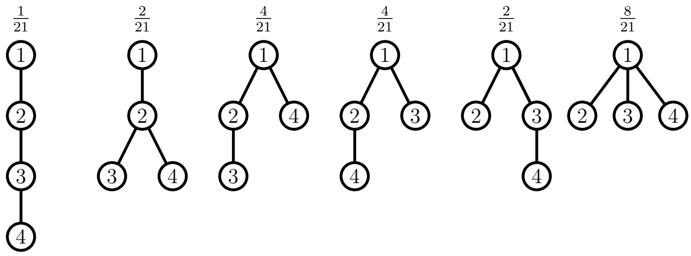
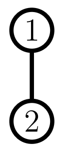
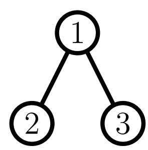


Figure 1. The exponentially preferential attachment recursive trees of size 4 with radix $a = 1/2$ and their probabilities.

probabilities, we illustrate the computation for one tree. We choose the fourth one from the left (with probability 4/21), as it has multiple nodes recruiting and a node (the root) recruiting twice. Initially, we have a root node labeled with 1. With probability 1, this root node recruits the node labeled with 2. So, the tree



appears with probability 1. The nodes 1 and 2 are now in a competition to attract node 3, with probabilities $\frac{(1/2)^1}{(1/2)^1+(1/2)^2} = 2/3$ and $\frac{(1/2)^2}{(1/2)^1+(1/2)^2} = 1/3$. The tree



emerges with probability $1 \times 2/3$. The nodes labeled with 1, 2, and 3 are now in competition to attract the node labeled with 4, with respective probabilities, $\frac{(1/2)^1}{(1/2)^1+(1/2)^2+(1/2)^3} = 4/7$, $\frac{(1/2)^2}{(1/2)^1+(1/2)^2+(1/2)^3} = 2/7$, and $\frac{(1/2)^3}{(1/2)^1+(1/2)^2+(1/2)^3} = 1/7$. Whence, if the node labeled with 2 is the one that recruits, we get the third tree on the right in Figure 1 with probability $1 \times 2/3 \times 2/7 = 4/21$.

This example illustrates the dynamic nature of the attraction probability at node i. While a^i is a fixed number, the scaling used is changing with time.

2. Generation and visualization

Before we present any theoretical results, it may help the reader grasp the gist of the varied behavior of the random exponentially preferential trees in the three regimes with diagrams.

To produce images, we first need a generating algorithm to provide the data. We present one such algorithm that sequentially cranks out the edges that join the tree. The edges appear in the form $(n+1, r)$, where r is the recruiter, when the tree size is n. For instance, the pair $(78, 50)$ stands for an edge joining the vertex labeled 78 to a tree of size 77, in which node 50 is the recruiter.

As an illustration of the action of the algorithm, we trace through the evolution of the third tree on the right in figure 1. The algorithm builds the tree edge by edge, first constructing the edge

(2,1), then adding the edge (3,2) and finally completing the description of this tree by adding the edge (4,2) to the list of edges.

Once the tree description is obtained in the form of a list of edges, we can visualize the tree by a drawing obtained with the aid of a tree-graphing package.

The algorithm assumes it can access the function

$$F(s,i) \leftarrow \sum_{r=1}^{i} \frac{(a-1)\,a^{r-1}}{a^s-1} = \frac{a^i-1}{a^s-1},$$

which accumulates the probabilities $\mathbb{P}(A_{s,1}) + \cdots + \mathbb{P}(A_{s,i})$ for the purpose of generating the recruiting index. With $F(s,i)$ having a closed form, it can be evaluated in $O(1)$ time.

The building algorithm assumes the existence of the primitive function **random**, which generates a random number uniformly distributed over the interval $(0, 1)$.

The core of the algorithm repeats the calculation of an index when the tree is of size " *size*," for $size = 1, \ldots, n-1$. At each size between 1 and $n-1$, a random variable U distributed uniformly between 0 and 1 is generated. If the value of U falls between $F(size, r-1)$ and $F(size, r)$, we take the recruiter to be r. This recruiter is receiving the node $size + 1$, and we store the pair (edge) $(size + 1, r)$ in the array R of recruiters. At the end of the execution of the algorithm, the array R holds a complete description of a tree of size n.

Here is a possible version in pseudo code:

```
for size from 1 to n − 1 do
        U ← random
        r ← 1
        while U > F(size, r) and r < size do
            r ← r + 1
        R[size] ← (size + 1, r)
```

The inner **while** loop may run an order of *size* in the worst case, and the outer **for** loop is driving *size* through $n-1$ iterations. The overall execution performs in $O(n^2)$ time. By this algorithm, we obtained three trees of size $n = 100$ each, under the settings $a = 1/2$, $a = 1$, $a = 2$, respectively. The data (edges) were then fed into the tree-drawing package "Pyvis," which produced the three images in Figure 2. The root of each tree is shown as a red star. The figure shows a random tree in the subcritical regime with radix $a = 1/2$ (top left), a random uniform (standard) recursive tree, with radix $a = 1$ (top right), and a random tree in the supercritical regime with radix $a = 2$ (bottom).

In Figure 2, we chose a drawing style to fill the space, rather than one going down vertically (as in the more traditional vertical drawing as in Figure 1). The vertical drawing would use the space sparsely.

The reader will notice right away that in the subcritical tree (the one at the top left in Figure 2), the nodes cluster near the root, making it a shrubby structure. In the uniform tree (the one at the top right in Figure 2), the nodes are all over the place, whereas in the supercritical tree (the one at the bottom in Figure 2), many nodes drag the tree toward higher altitudes, making it a tall and scrawny tree with short branches sprouting out of a main thin trunk.

2.1 Notation

The indicator of event $\mathcal{E}$ is $\mathbb{I}_{\mathcal{E}}$. The following presentation involves $H_p^{(r)} = \sum_{k=1}^{p} 1/k^r$, the pth harmonic number of order r.[1] The harmonic numbers of the first two orders have well-known asymptotic equivalents (as $n \to \infty$):[2]

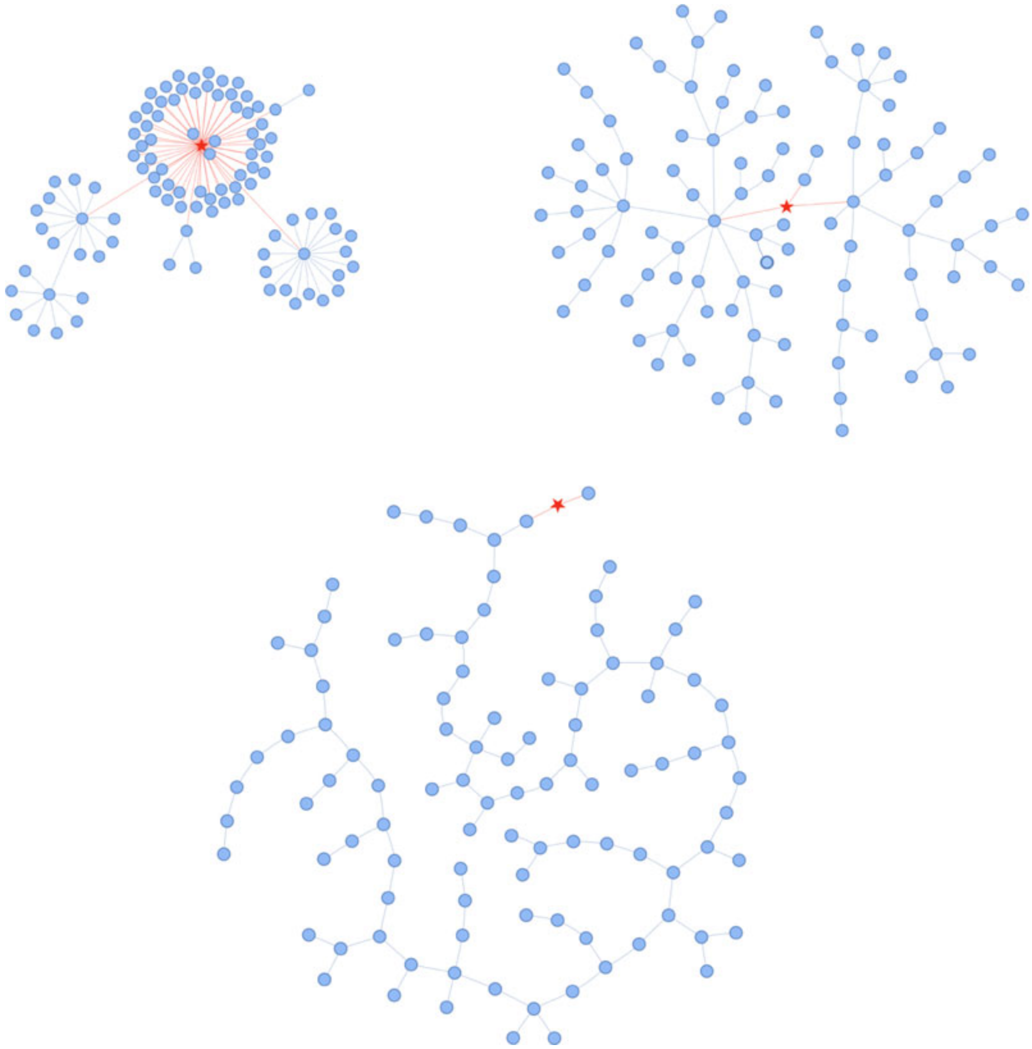


Figure 2. Randomly generated trees of size 100: subcritical (top left) with radix 1/2, uniform (top right) with radix $a = 1$, supercritical (bottom) with radix 2.

$$H_n = \ln n + \gamma + O\left(\frac{1}{n}\right); \tag{2}$$

$$H_n^{(2)} = \frac{\pi^2}{6} + O\left(\frac{1}{n}\right). \tag{3}$$

Ceils and floors appear in the calculations. The floor of a real number x can be represented by removing the floor and compensating for the fractional part of x, sometimes denoted by $\{x\}$. That is, we have

$$\lfloor x \rfloor = x - \{x\}.$$

For $i \geq 1$, and $a > 1$, in the sequel, we use the numbers

$$c_i^{(1)}(a) = \sum_{k=i}^{\infty} \frac{1}{a^k - 1},$$
$$c_i^{(2)}(a) = \sum_{k=i}^{\infty} \frac{1}{(a^k - 1)^2}.$$

We denote the normally distributed random variable with mean μ and variance $\sigma^2 > 0$ by $\mathcal{N}(\mu, \sigma^2)$ and denote the Poisson random variable with parameter $\lambda > 0$ by Poi (λ). The notation $\mathcal{L}(X)$ stands for the law (probability distribution) of a random variable X.

The total variation distance (d_{TV}) between the laws of the nonnegative integer-valued random variables X and Y is defined as

$$d_{TV}\big(\mathcal{L}(X), \mathcal{L}(Y)\big) = \frac{1}{2} \sum_{j=0}^{\infty} \big|\mathbb{P}(X = j) - \mathbb{P}(Y = j)\big|.$$

As some authors do, we simplify the notation of the total variation distance to $d_{TV}(X, Y)$, but it should be understood that it is the distance between the laws of these variables.

The following theorem by Barbour and Holst (Barbour and Hall, 1984) is beneficial in obtaining Poisson approximations for node degrees.

Theorem 2.1. *Let $X_1, \ldots, X_n$ be independent Bernoulli random variables, such that $\mathbb{P}(X_k = 1) = p_k$, for $i = 1, \ldots, n$, and let $S_n = \sum_{k=1}^{n} X_k$ be the sum of these variables. Define*

$$\lambda_{n,1} = \sum_{k=1}^{n} \mathbb{E}[X_k] = \sum_{k=1}^{n} p_k, \qquad \textit{and} \qquad \lambda_{n,2} = \sum_{k=1}^{n} p_k^2.$$

We have

$$d_{TV}\big(S_n,\ \text{Poi}\ (\lambda_{n,1})\ \big) \le (1 - e^{-\lambda_{n,1}})\, \frac{\lambda_{n,2}}{\lambda_{n,1}}.$$

The theorem is one of several versions fitting in the machinery of Poisson approximation (and the more general framework of Chen−Stein methods) (Barbour et al., 1992).

3. Node outdegrees

Let $\Delta_{n,i}$ be the outdegree of node i in a tree of size n. It is related to the degree of node i. Except for the root, any node degree is 1 plus its outdegree. As the root is the only node that does not have a parent, its degree is the same as its outdegree.

Remark 3.1. *In a tree of size n, the sum of the outdegrees is $n - 1$.*

The outdegree of node i increases, when it recruits, and we have the representation

$$\Delta_{n,i} = \sum_{k=i}^{n-1} \mathbb{I}_{A_{k,i}}. \tag{4}$$

3.1 Mean

Take expectations of (4). We then get, for $n \ge 2$, a mean value for the ith degree in the form

$$\mathbb{E}[\Delta_{n,i}] = \sum_{k=i}^{n-1} \mathbb{P}(\mathbb{I}_{A_{k,i}} = 1) = \begin{cases} \displaystyle\sum_{k=i}^{n-1} \frac{1}{k}, & \text{if } a = 1; \\ \displaystyle\sum_{k=i}^{n-1} \frac{(a-1)\, a^{i-1}}{a^k - 1}, & \text{otherwise.} \end{cases} \tag{5}$$

Proposition 3.1. *Let $\Delta_{n,i}$ be the outdegree of node i in an exponentially preferential tree of size n and radix $0 < a < 1$. As $n \to \infty$, we have*

$$\mathbb{E}[\Delta_{n,i}] = \begin{cases} (1-a)\, a^{i-1}\, (n-i) + O(a^i), & \text{if } a < 1; \\ \ln n - \ln i + O\big(\frac{1}{i}\big), & \text{if } a = 1. \end{cases}$$

Otherwise, a is greater than 1, and the case is ramified according to the relationship between i and n. As $n \to \infty$, we have the phases:

$$\mathbb{E}[\Delta_{n,i}] = \begin{cases} (a-1)\, a^{i-1} c_i^{(1)}(a) + O\big(\frac{1}{a^{n-i}}\big), & i \text{ fixed}; \\ 1 - \frac{1}{a^{n-i}} + O\big(\frac{1}{a^i}\big), & n \ge i \to \infty. \end{cases}$$

Proof. We need to cover three different regimes:

- **The subcritical regime** ($0 < a < 1$): Start with the lower display in (5) in the form

$$\mathbb{E}[\Delta_{n,i}] = (1-a)\, a^{i-1} \sum_{k=i}^{n-1} \frac{1}{1-a^k}.$$

  A Taylor series expansion of the summand yields:

$$\begin{aligned} \mathbb{E}[\Delta_{n,i}] &= (1-a)\, a^{i-1} \sum_{k=i}^{n-1} \big(1 + O(a^k)\big) \\ &= (1-a)\, a^{i-1} \Big(n - i + O\Big(\sum_{k=i}^{n-1} a^k\Big)\Big) \\ &= (1-a)\, a^{i-1} (n-i) + O(a^i). \end{aligned}$$

- **The critical regime** ($a = 1$): Here we use the upper display in (5), yielding

$$\mathbb{E}[\Delta_{n,i}] = \sum_{k=i}^{n-1} \frac{1}{k} = H_{n-1} - H_{i-1}.$$

  Using the asymptotic equivalent in (2), as both n and i approach ∞, we get

$$\mathbb{E}[\Delta_{n,i}] = \ln n - \ln i + O\Big(\frac{1}{i}\Big),$$

  a known result (Javanian and Asl, 2003; Mahmoud, 2019).

- **The supercritical regime** $a > 1$: Again, we start with the lower display in (5). We have two phases in the lives of the various nodes:

(i) The index i is a constant. In this phase, we have

$$\mathbb{E}[\Delta_{n,i}] = (a-1)\, a^{i-1} \sum_{k=i}^{n-1} \frac{1}{a^k - 1} = (a-1)\, a^{i-1}\, c_i^{(1)}(a) + O\Big(\frac{1}{a^{n-i}}\Big).$$

(ii) The index i is increasing with n. We can approximate the series in the lower display in (5) with a geometric series. To this end, we quantify the absolute error $\mathcal{E}_{n,i}$ from the bound

$$\mathcal{E}_{n,i} := \left| \sum_{k=i}^{n-1} \frac{1}{a^k - 1} - \sum_{k=i}^{n-1} \frac{1}{a^k} \right| = \sum_{k=i}^{n-1} \frac{1}{a^k - 1} - \sum_{k=i}^{n-1} \frac{1}{a^k} = \sum_{k=i}^{n-1} \frac{1}{a^k(a^k - 1)}.$$

From a Taylor series expansion of the summand, we get

$$\begin{aligned}
\mathcal{E}_{n,i} &:= \sum_{k=i}^{n-1} \frac{1}{a^{2k}} \Big(1 + O\Big(\frac{1}{a^k}\Big)\Big) \\
&= \frac{1}{a^{2i}} \Big(\frac{1 - 1/a^{2(n-i)}}{1 - 1/a^2}\Big) + O\Big(\sum_{k=i}^{n-1} \frac{1}{a^{3k}}\Big) \\
&= \frac{1}{a^{2(i-1)}} \Big(\frac{1 - 1/a^{2(n-i)}}{a^2 - 1}\Big) + O\Big(\frac{1}{a^{3(i-1)}} \Big(\frac{(1 - 1/a^{3(n-i)})}{a^3 - 1}\Big)\Big) \\
&= O\Big(\frac{1}{a^{2i}}\Big).
\end{aligned}$$

In this phase, we have

$$\begin{aligned}
\mathbb{E}[\Delta_{n,i}] &= (a-1)\, a^{i-1} \Big(\Big(\sum_{k=i}^{n-1} \frac{1}{a^k}\Big) + \mathcal{E}_{n,i}\Big) \\
&= (a-1)\, a^{i-1} \Big(\Big(\sum_{k=i}^{n-1} \frac{1}{a^k}\Big) + O\Big(\frac{1}{a^{2i}}\Big)\Big) \\
&= (a-1)\, a^{i-1}\, \frac{1 - (1/a)^{n-i}}{a^i (1 - 1/a)} + O\Big(\frac{1}{a^i}\Big) \\
&= 1 - \frac{1}{a^{n-i}} + O\Big(\frac{1}{a^i}\Big).
\end{aligned}$$

□

3.2 Variance

In (4), the indicators $\mathbb{I}_{A_{n,i}}$, for $n \ge 1$, are independent. It follows that

$$\mathbb{V}\mathrm{ar}[\Delta_{n,i}] = \mathbb{V}\mathrm{ar}\Big[\sum_{k=i}^{n-1} \mathbb{I}_{A_{k,i}}\Big] = \sum_{k=i}^{n-1} \mathbb{V}\mathrm{ar}[\mathbb{I}_{A_{k,i}}].$$

That is, we have

$$\mathbb{V}\mathrm{ar}[\Delta_{n,i}] = (a-1)\, a^{i-1} \sum_{k=i}^{n-1} \frac{1}{a^k - 1} - (a-1)^2\, a^{2i-2} \sum_{k=i}^{n-1} \Big(\frac{1}{a^k - 1}\Big)^2. \quad (6)$$

This combinatorial form is reducible in the uniform case ($a = 1$), where we get

$$\mathbb{V}\mathrm{ar}[\Delta_{n,i}] = \sum_{k=i}^{n-1} \frac{1}{k}\Big(1 - \frac{1}{k}\Big) = H_{n-1} - H_{i-1} - \big(H_{n-1}^{(2)} - H_{i-1}^{(2)}\big). \quad (7)$$

Again we have an asymptotic trichotomy.

Proposition 3.2. *Let $\Delta_{n,i}$ be the outegree of node i in an exponential preferential tree of size n and radix $0 < a < 1$. We then have*

$$\mathbb{V}\mathrm{ar}[\Delta_{n,i}] = \begin{cases} (1-a)\, a^{i-1}\big(1-(1-a)\, a^{i-1}\big)(n-i) + O(a^i), & \textit{if } a < 1; \\ \ln n - \ln i + O(\frac{1}{i}), & \textit{if } a = 1. \end{cases}$$

Otherwise, a is greater than 1, and the case is ramified according to the relationship between i and n. As $n \to \infty$, we have the phases:

$$\mathbb{V}\mathrm{ar}[\Delta_{n,i}] = \begin{cases} (a-1)\, a^{i-1} c_i^{(1)}(a) & \\ \qquad - (a-1)^2\, a^{2i-2} c_i^{(2)}(a) + O\big(\frac{1}{a^{n-i}}\big), & i \textit{ fixed}; \\ \frac{2}{a+1} - \frac{1}{a^{n-i}} + \frac{a-1}{(a+1)a^{2n-2i}} + O\big(\frac{1}{a^i}\big), & n \geq i \to \infty. \end{cases}$$

Proof. We need to cover three different regimes:

- **The subcritical regime ($0 < a < 1$):** We first write the exact variance in (6) in the form

$$\mathbb{V}\mathrm{ar}[\Delta_{n,i}] = (1-a)\, a^{i-1} \sum_{k=i}^{n-1} \frac{1}{1-a^k} - (1-a)^2\, a^{2i-2} \sum_{k=i}^{n-1} \Big(\frac{1}{1-a^k}\Big)^2.$$

  As we did in the proof of the mean, a Taylor series expansion for each series yields

$$\begin{aligned} \mathbb{V}\mathrm{ar}[\Delta_{n,i}] &= \big((1-a)\, a^{i-1}(n-i) + O(a^i)\big) \\ &\qquad - \big((1-a)^2\, a^{2i-2}(n-i) + O(a^{2i})\big) \\ &= (1-a)\, a^{i-1}\big(1-(1-a)\, a^{i-1}\big)(n-i) + O(a^i). \end{aligned}$$

- **The critical regime ($a = 1$):** Asymptotically, as both n and $i \leq n$ approach ∞, from (7) and the asymptotics in (2)–(3), we get

$$\begin{aligned} \mathbb{V}\mathrm{ar}[\Delta_{n,i}] &= \Big(\Big(\ln n + \gamma + O\Big(\frac{1}{n}\Big)\Big) - \Big(\ln i + \gamma + O\Big(\frac{1}{i}\Big)\Big) \\ &\qquad - \Big(\Big(\frac{\pi^2}{6} + O\Big(\frac{1}{n}\Big)\Big) - \Big(\frac{\pi^2}{6} + O\Big(\frac{1}{i}\Big)\Big)\Big) \\ &= \ln n - \ln i + O\Big(\frac{1}{i}\Big), \end{aligned}$$

  a known result (Javanian and Asl, 2003; Mahmoud, 2019).
- **The supercritical regime ($a > 1$):** In (6), we have phases:

  (i) The index i is a constant. In this case, we have

$$\begin{aligned} \mathbb{V}\mathrm{ar}[\Delta_{n,i}] &= \Big((a-1)\, a^{i-1}\, c_i^{(1)}(a) + O\Big(\frac{1}{a^{n-i}}\Big)\Big) \\ &\qquad - \Big((a-1)^2\, a^{2i-2}\, c_i^{(2)}(a) + O\Big(\frac{1}{a^{2n-2i}}\Big)\Big) \\ &= (a-1)\, a^{i-1}\, c_i^{(1)}(a) - (a-1)^2\, a^{2i-2}\, c_i^{(2)}(a) + O\Big(\frac{1}{a^{n-i}}\Big). \end{aligned}$$

(ii) The index i is increasing with n. We resort again to the approximation of the two sums by geometric series. In this phase, we have

$$\begin{aligned}\mathbb{V}\mathrm{ar}[\Delta_{n,i}] &= (a-1)\,a^{i-1}\Big(\Big(\sum_{k=i}^{n-1}\frac{1}{a^k}\Big)+O\Big(\frac{1}{a^{2i}}\Big)\Big)\\ &\qquad -(a-1)^2\,a^{2i-2}\Big(\sum_{k=i}^{n-1}\frac{1}{a^{2k}}\Big)+O\Big(\frac{1}{a^{4i}}\Big)\Big)\\ &= (a-1)\,a^{i-1}\,\frac{1-(1/a)^{n-i}}{a^i(1-1/a)}\\ &\qquad -(a-1)^2\,a^{2i-2}\,\frac{1-(1/a^2)^{n-i}}{a^{2i}(1-1/a^2)}+O\Big(\frac{1}{a^i}\Big)\\ &= \Big(1-\frac{1}{a^{n-i}}\Big)-\frac{a-1}{a+1}\Big(1-\frac{1}{a^{2n-2i}}\Big)+O\Big(\frac{1}{a^i}\Big)\\ &= \frac{2}{a+1}-\frac{1}{a^{n-i}}+\frac{a-1}{(a+1)a^{2n-2i}}+O\Big(\frac{1}{a^i}\Big).\end{aligned}$$

□

Corollary 3.1. *In the subcritical regime* ($0 < a < 1$)*, when* $(n-i)a^i \to \infty$, *we have*

$$\frac{\Delta_{n,i}}{(n-i)\,a^i} \xrightarrow{\mathbb{P}} \frac{1-a}{a},$$

and in the critical regime ($a = 1$)*, when* $n/i \to \infty$, *we have*

$$\frac{\Delta_{n,i}}{\ln(n/i)} \xrightarrow{\mathbb{P}} 1,$$

3.3 Distributions

In view of the trichotomy, we observed in the mean and variance, it should be anticipated that the asymptotic distribution of the outdegree would have three regimes, too, according as where a is on the real line.

Theorem 3.1. *Let* $\Delta_{n,i}$ *be the outegree of node i in an exponentially preferential tree of size n and radix* $a > 0$. *We then have:*

(i) Let $g(n)$ *be a positive integer-valued function increasing to infinity, such that* $g(n) = o(\ln n)$. *In the subcritical regime* $(0 < a < 1)$, *we have phases:*

(a) In the early phase $(1 \le i \le \lfloor \log_{\frac{1}{a}} n\rfloor - g(n))$,[3] *as* $n \to \infty$, we have[4]

$$\frac{\Delta_{n,i} - \frac{1-a}{a}\,a^i\,n}{\sqrt{a^i\,n}} \xrightarrow{\mathcal{L}} \mathcal{N}\Big(0, \frac{1-a}{a}\Big).$$

(b) In the intermediate phase ($i = \lfloor \log_{\frac{1}{a}} n\rfloor + c$, *and* $c \in \mathbb{Z}$), *we have*

$$d_{TV}\left(\Delta_{n,i},\ \mathrm{Poi}\left(\frac{1-a}{a}\,a^{c-\{\log_{\frac{1}{a}} n\}}\right)\right) \to 0.$$

(c) Let $h(n)$ *be a positive integer-valued function increasing to infinity, that grows faster than a constant, but remains at most* $n - \lfloor \log_{\frac{1}{a}} n\rfloor$. *In the late phase* ($i = \lfloor \log_{\frac{1}{a}} n\rfloor + h(n)$), *we have*

$$\Delta_{n,i} \xrightarrow{\mathbb{P}} 0.$$

(ii) *In the critical regime* ($a = 1$), *we have phases:*

(a) *In the early phase* $n/i \to \infty$, *we have*

$$\frac{\Delta_{n,i} - \ln(n/i)}{\sqrt{\ln(n/i)}} \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1).$$

(b) *In the intermediate phase* $i \sim cn$ *(and* $c \in (0, 1)$*), we have*

$$\Delta_{n,i} \xrightarrow{\mathbb{P}} \text{Poi}\left(\ln \frac{1}{c}\right).$$

(c) *In the late phase* $i \sim n$, *we have*

$$\Delta_{n,i} \xrightarrow{\mathbb{P}} 0.$$

(iii) *In the supercritical regime* ($a > 1$)*, we have phases:*

(a) *In the early phase* (i *fixed*)*, as* $n \to \infty$, *we have*

$$\Delta_{n,i} \xrightarrow{a.s.} \Delta_i^*,$$

and, for $k \geq 1$, *the limiting random variable has the distribution*

$$\mathbb{P}(\Delta_i^* = k) = (a-1)^k a^{(i-1)k} \lim_{n\to\infty} \left(\Big(\prod_{\ell=i}^{n-1} \frac{1}{a^\ell - 1} \Big) \sum_{i \leq j_1 < j_2 < \cdots < j_k \leq n-1} \right.$$
$$\left. \times \prod_{\substack{i \leq m \leq n-1 \\ m \notin \{j_1, \ldots, j_k\}}} \Big((a^m - 1) - (a-1)a^{i-1} \Big) \right).$$

(b) *Let* $b(n)$ *be a function growing to infinity, in such a way that* $n - b(n)$ *also grows to infinity. In the intermediate phase* ($1 \leq i = n - b(n)$)*, we have*[5]

$$\mathbb{P}(\Delta_{n,n-b(n)} = k) = (a-1)^k a^{(n-b(n)-1)k} \Big(\prod_{\ell=i}^{n-1} \frac{1}{a^\ell - 1} \Big)$$
$$\times \sum_{n-b(n) \leq j_1 < j_2 < \cdots < j_k \leq n-1} \prod_{\substack{i \leq m \leq n-1 \\ m \notin \{j_1, \ldots, j_k\}}} \Big((a^m - 1) - (a-1)a^{n-b(n)-1} \Big).$$

(c) *In the late phase* ($1 \leq i = n - c$, *with* $c \in \mathbb{N}$)*, we have* $\Delta_{n,n-c} \xrightarrow{a.s.} \Delta_c^\star$, *and* $\Delta_c^\star$ *has the distribution*

$$\mathbb{P}(\Delta_c^\star = k) = \frac{(a-1)^k}{a^{c(c+1)/2}} \sum_{1 \leq r_1 < r_2 < \cdots < r_k \leq c} \prod_{\substack{1 \leq s \leq c \\ s \notin \{r_1, \ldots, r_k\}}} (a^s - a + 1).$$

Proof. We need to cover three different regimes:

(i) **The subcritical regime:** Within the regime $0 < a < 1$, we recognize phases:

(a) **The early phase** ($1 \leq i = \lfloor \log_{\frac{1}{a}} n \rfloor - g(n)$): Let

$$s_{n,i}^2 = \mathbb{V}\text{ar}[\Delta_{n,i}] = \sum_{k=i}^{n-1} \mathbb{V}\text{ar}[\mathbb{I}_{A_{k,i}}].$$

From the calculation of the variance in this regime (cf. Proposition 3.2), we have $s_{n,i}^2 \sim (1-a)\, a^{i-1}(n-i)$, as $a^i(n-i) \to \infty$. The case is amenable to normality via Lindeberg's central limit theorem, if $n-i \to \infty$, in which case we have a sum of a large number of independent indicators (Bernoulli random variables).

For i to be in this phase, we must have $\log_{\frac{1}{a}} \left(a^i\,(n-i)\right) \to \infty$, as $n \to \infty$. That is to say,

$$-i + \log_{\frac{1}{a}} \left(n\left(1 - \frac{i}{n}\right)\right) = -i + \log_{\frac{1}{a}} n + \log_{\frac{1}{a}} \left(1 - \frac{i}{n}\right) = \left(\log_{\frac{1}{a}} n\right) - i + O(i/n)$$

must increase to ∞. If $\lfloor \log_{\frac{1}{a}} n \rfloor - i \to \infty$, such an asymptotic relation holds, when i increases up to $\lfloor \log_{\frac{1}{a}} n \rfloor - g(n)$, for any positive integer function $g(n)$ that is $o(\ln n)$. Fix $\varepsilon > 0$, and define Lindeberg's quantity

$$L_{n,i}(\varepsilon) = \frac{1}{s_{n,i}^2} \sum_{k=i}^{n-1} \int_{\left|\,\mathbb{I}_{A_{k,i}} - \mathbb{E}[\mathbb{I}_{A_{k,i}}]\,\right| > \varepsilon s_{n,i}} \mathbb{I}_{A_{k,i}}^2 \, d\mathbb{P},$$

where $\mathbb{P}$ is the underlying probability measure. The indicators are Bernoulli random variables bounded by 1. Hence, we have

$$\left|\,\mathbb{I}_{A_{k,i}} - \mathbb{E}[\mathbb{I}_{A_{k,i}}]\,\right| \le \mathbb{I}_{A_{k,i}} + \mathbb{E}[\mathbb{I}_{A_{k,i}}] \le 2,$$

whereas $\varepsilon s_{n,i}$ grows to infinity, no matter how small ε is, or what the value of i is within the specified phase. In other words, the sets in the integration are all empty for large enough n (greater than some $n_0 = n_0(\varepsilon, i)$). We can now read the Lindeberg quantity as

$$L_{n,i}(\varepsilon) = \frac{1}{s_{n,i}^2} \sum_{k=i}^{n} \int_{\phi} \mathbb{I}_{A_{k,}i}^2 \, d\mathbb{P} = 0.$$

We have verified that, within the phase $i = \lfloor \log_{\frac{1}{a}} n \rfloor - g(n)$, the quantity $L_{n,i}(\varepsilon) \to 0$, for all $\varepsilon > 0$. Thus, we have the Gaussian law

$$\frac{\Delta_{n,i} - (1-a)\, a^{i-1}\,(n-i) + O(a^i)}{\sqrt{(1-a)\, a^{i-1}\left(1-(1-a)\, a^{i-1}\right)(n-i) + O(a^i)}} \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1).$$

Toward simpler appearance, we use Slutsky's theorem (Karr, 1993), pp. 146–147 to remove some factors:

$$\frac{\Delta_{n,i} - \frac{1-a}{a}\, a^i\, n}{\sqrt{a^i n}} \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \frac{1-a}{a}\right).$$

(b) **The intermediate phase** ($i = \left\lfloor \log_{\frac{1}{a}} n \right\rfloor + c$):

Let $i = \left\lfloor \log_{\frac{1}{a}} n \right\rfloor + c = \log_{\frac{1}{a}} n - r_n + c$, where $r_n = \{\log_{\frac{1}{a}} n\} \in [0,\ 1)$, and $c \in \mathbb{Z}$. In this phase, we have

$$a^i(n-i) = a^{\log_{\frac{1}{a}} n - r_n + c}\left(n - \log_{\frac{1}{a}} n + r_n - c\right) \sim \frac{a^{c-r_n}}{n} \times n = a^{c-r_n}.$$

In the notation of Theorem 2.1, we have

$$\lambda_{n,1} = \sum_{k=i}^{n} \mathbb{P}(\mathbb{I}_{A_{k,i}} = 1) = \sum_{k=i}^{n} \frac{(1-a)a^{i-1}}{1-a^k} \sim (1-a)a^{i-1}(n-i) \sim \frac{1-a}{a}\, a^{c-r_n},$$

and

$$\begin{aligned}\lambda_{n,2} &= \sum_{k=i}^{n-1} \mathbb{P}^2(\mathbb{I}_{A_{k,i}} = 1)\\ &= \sum_{k=i}^{n-i} \frac{(1-a)^2 a^{2i-2}}{(1-a^k)^2}\\ &\sim (1-a)^2 a^{2i-2}\big(n - i + O(a^i)\big)\\ &\sim \Big(\frac{1-a}{a}\Big)^2 a^{2(\log_{\frac{1}{a}} n - r_n + c)} \times n\\ &\sim \Big(\frac{1-a}{a}\Big)^2 \frac{a^{2c-2r_n}}{n^2} \times n\\ &\to 0.\end{aligned}$$

By that theorem, we conclude

$$d_{TV}\left(\Delta_{n,i},\ \mathrm{Poi}\left(\frac{1-a}{a}\, a^{c-r_n}\right)\right) \to 0.$$

(c) **The late phase** ($i = \lfloor \log_{\frac{1}{a}} n\rfloor + h(n) \le n$, **with** $h(n) \to \infty$): In this late phase, $\mathbb{V}\mathrm{ar}[\Delta_{n,i}] \sim (1-a)a^{i-1}(n-i) \to 0$. As is well known, convergence of the variance of a sequence of random variables to 0 implies that the sequence converges weakly to a constant, a consequence of Chebyshev's inequality. The limiting constant must be the constant obtained from the L_1 converges $\Delta_{n,i} \xrightarrow{\mathcal{L}} 0$.

(ii) **The critical regime:** In this uniform attachment case, the distribution as stated is known. We refer the reader to two different proofs in Javanian and Asl (2003); Mahmoud (2019).

(iii) **The supercritical regime:** In this regime we, work from the exact distribution to produce local limit theorems in the different phases. For $\Delta_{n,i}$ to be equal to k, node i must recruit k times and fail to recruit $n-i-k$ times. We can partition the event $\Delta_{n,i} = k$ into *disjoint* sets according to the the size of the tree at the times of recruiting. Suppose the k successes in recruiting occur when the tree sizes are $i \le j_1 < j_2 < \ldots j_k \le n-1$. The probability of this event is

$$\Big(\prod_{\ell\in\{j_1,\ldots,j_k\}} \mathbb{P}(\mathbb{I}_{A_{\ell,i}} = 1)\Big)\Big(\prod_{\substack{i\le m\le n-1\\ m\notin\{j_1,\ldots,j_k\}}} \mathbb{P}(\mathbb{I}_{A_{m,i}} = 0)\Big).$$

Using the probabilities in the lower display in (1), we obtain

$$\begin{aligned}\mathbb{P}(\Delta_{n,i} = k) &= \sum_{i\le j_1<j_2<\cdots j_k\le n-1} \Big(\prod_{\ell\in\{j_1,\ldots,j_k\}} \mathbb{P}(\mathbb{I}_{A_{\ell,i}} = 1)\Big)\\ &\qquad\qquad \times \Big(\prod_{\substack{i\le m\le n-1\\ m\notin\{j_1,\ldots,j_k\}}} \mathbb{P}(\mathbb{I}_{A_{m,i}} = 0)\Big)\\ &= \sum_{i\le j_1<j_2<\cdots<j_k\le n-1} \Big(\prod_{\ell\in\{j_1,\ldots,j_k\}} \frac{(a-1)a^{i-1}}{a^\ell - 1}\Big)\\ &\qquad\qquad \times \prod_{\substack{i\le m\le n-1\\ m\notin\{j_1,\ldots,j_k\}}} \Big(1 - \frac{(a-1)a^{i-1}}{a^m - 1}\Big) \qquad (8)\end{aligned}$$

$$= (a-1)^k a^{(i-1)k} \Big(\prod_{\ell=i}^{n-1} \frac{1}{a^\ell - 1} \Big) \sum_{i \le j_1 < j_2 < \cdots < j_k \le n-1} \times \prod_{\substack{i \le m \le n-1 \\ m \notin \{j_1, \ldots, j_k\}}} \Big((a^m - 1) - (a-1) a^{i-1} \Big). \tag{9}$$

According to Kolmogorov's criterion (Theorem 22.3 in Billingsley (2012)), for independent zero-mean random variables $X_1, X_2, X_3, \ldots$, when $\sum_{k=1}^{\infty} \mathbb{V}\text{ar}[X_k] < \infty$, the sum $\sum_{k=1}^{n} X_k$ converges almost surely to a limit.
In the supercritical regime, when $n - i \to \infty$, we have

$$\sum_{k=i}^{n} \mathbb{V}\text{ar}\big[\mathbb{I}_{A_{k,i}} - \mathbb{E}[\mathbb{I}_{A_{k,i}}]\big] \le \sum_{k=1}^{\infty} \frac{(a-1)\, a^{i-1}}{a^k - 1} \Big(1 - \frac{(a-1)\, a^{i-1}}{a^k - 1} \Big) < \infty.$$

In view of Kolmogorov's criterion, when $n - i = g(n) \to \infty$,

$$\Delta_{n,i} - \mathbb{E}[\Delta_{n,i}] = \sum_{k=i}^{n} \mathbb{I}_{A_{k,i}} - \sum_{k=i}^{n} \mathbb{E}[\mathbb{I}_{A_{k,i}}]$$

converges to a limit.
Having shown that in the supercritical phase we have $\sum_{k=i}^{n} \mathbb{E}[\mathbb{I}_{A_{k,i}}] = \mathbb{E}[\Delta_{n,i}]$ is convergent (cf. Proposition 3.1), we see right away that, if $n - i \to \infty$, we would have $\Delta_{n,i} = \sum_{k=i}^{n} \mathbb{I}_{A_{k,i}}$ converging to a limit.

(a) **The early phase (i fixed):** Certainly, in this phase $n - i \to \infty$, as $n \to \infty$. By Kolmogorov's criterion, $\Delta_{n,i}$ converges to a limit, which we call $\Delta_i^\star$. We can determine the distribution of $\Delta_i^\star := \lim_{n\to\infty} \Delta_{n,i}$ by the following argument.
Since $\Delta_{n,i}$ converges almost surly, it also converges in distribution. The limit of the latter probabilities exists (and must be the distribution of the almost-sure limit, too). Indeed, $\Delta_{n,i}$ converges almost surely to a limit Δ_i^* with a distribution determined as the limit of the probabilities in (9).

(b) **The intermediate phase (i grows faster than a constant, but slower than $n - c$, for any $c \in \mathbb{N}$).** In this phase, i is $n - d(n)$, for an integer-valued function $d(n) \to \infty$, in such a way that $n - d(n)$ also tends to infinity. In this case, for any fixed k, (8) takes the form:[6]

$$\mathbb{P}(\Delta_{n,n-d(n)} = k) = \sum_{n-d(n) \le j_1 < j_2 < \cdots < j_k \le n-1} \Big(\prod_{\ell \in \{j_1, \ldots, j_k\}} \frac{(a-1) a^{n-d(n)-1}}{a^\ell - 1} \Big) \times \prod_{\substack{n-d(n) \le m \le n-1 \\ m \notin \{j_1, \ldots, j_k\}}} \Big(1 - \frac{(a-1) a^{n-d(n)-1}}{a^m - 1} \Big)$$

$$= \frac{(a-1)^k a^{k(n-d(n)-1)}}{\prod_{r=n-d(n)}^{n-1} (a^r - 1)} \sum_{n-d(n) \le j_1 < j_2 < \cdots < j_k \le n-1} \times \prod_{\substack{n-d(n) \le m \le n-1 \\ m \notin \{j_1, \ldots, j_k\}}} \big(a^m - 1 - (a-1) a^{n-d(n)-1} \big).$$

(c) **The late phase ($1 \le i = n - c$, and $c \in \mathbb{N}$):** Starting with the probabilities in the form (8), we write an asymptotic equivalent.

Note that while the indices in the sum are large numbers there, is only a finite number of them (c of them to be exact). It is therefore legitimate to take the asymptotic terms individually:

$$\begin{aligned}\mathbb{P}(\Delta_{n,n-c}=k) &\sim \sum_{n-c\le j_1<j_2<\cdots<j_k\le n-1}\Big(\prod_{\ell\in\{j_1,\ldots,j_k\}}\frac{a-1}{a^{\ell-(n-c)+1}}\Big)\\ &\qquad\times\prod_{\substack{n-c\le m\le n-1\\ m\notin\{j_1,\ldots,j_k\}}}\Big(1-\frac{a-1}{a^{m-(n-c)+1}}\Big)\\ &\sim\frac{(a-1)^k}{\prod_{r=n-c}^{n-1}a^{r-(n-c)+1}}\\ &\qquad\times\sum_{n-c\le j_1<j_2<\cdots<j_k<n-1}\ \prod_{\substack{n-c\le m\le n-1\\ m\notin\{j_1,\ldots,j_k\}}}(a^{m-(n-c)+1}-a+1)\\ &\to\frac{(a-1)^k}{aa^2\ldots a^c}\sum_{1\le r_1<r_2<\cdots<r_k\le c}\ \prod_{\substack{1\le s\le c\\ s\notin\{r_1,\ldots,r_k\}}}(a^s-a+1).\end{aligned}$$

□

3.4 Illustrative examples

In any of the three regimes of a, there is an intriguing interplay between n and i. We only discuss interpretations and examples from the subcritical and supercritical regimes, since the critical phase is well studied, and illustrative examples of it can be found elsewhere. We refer a reader interested in a discussion of the critical regime to Javanian and Asl (2003); Mahmoud (2019).

3.5 The subcritical regime

When $a<1$, the term a^{i-1} in $\mathbb{E}[\Delta_{n,i}]$ decreases exponentially fast in i. Take $a=1/2$, for instance. With this radix, for i in the subcritical phase, we have the convergence

$$\frac{\Delta_{n,i}-\frac{n}{2^i}}{\sqrt{\frac{n}{2^i}}}\xrightarrow{\mathcal{L}}\mathcal{N}(0,1).$$

A Gaussian law holds so long as i is well below $\lfloor\log_{\frac{1}{a}} n\rfloor$ (differing by an increasing function from that critical level). When i approaches the critical phase, Poisson approximations kick in to replace the normal distribution. When i is tied to $\lfloor\log_{\frac{1}{a}} n\rfloor$ by a constant, it is related to $\log_{\frac{1}{a}} n$ via corrections obtained by removing the floors. These corrections are oscillating functions in n and are uniformly dense on the real line (Kuipers and Niederreiter, 1974). So, there is not really convergence to a Poisson limit, but rather approximations to a family of Poisson distributions, with parameters lying in the range $[\frac{1}{2^c},\frac{2}{2^c})$.

Table 1 shows the behavior in the subcritical regime for $a=1/2$ and some selected phases. Note the fourth and fifth entries (from the top) which lie in the intermediate phase, where there is no limit per se, but rather good approximations by various Poisson distributions. For instance, in the phase $i=\lceil\log_2 n\rceil-5$, Poi (18.4136) is a good approximation for the distribution of $\Delta_{20000,10}$ at $n=20000$, Poi (18.4182) is a good approximation for the distribution of $\Delta_{20001,10}$ at $n=20001$, and Poi (18.4228) is a good approximation for the distribution of $\Delta_{20002,10}$ at $n=20002$.

For the last two entries in Table 1, the variance is diminishing at a fast rate, and the sum of the variances converges. By Kolmogorov's theorem, we have an almost-sure convergence in both cases.

Table 1. The asymptotic mean, variance, and distribution of the outdegree of an exponentially preferential tree with radix 1/2 in some selected phases

i	Mean	Variance	Distribution
1	$\frac{n}{2}$	$\frac{n}{4}$	$\frac{\Delta_{n,1}-\frac{1}{2}n}{\sqrt{n}} \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \frac{1}{4}\right)$
5	$\frac{n}{32}$	$\frac{n}{32}$	$\frac{\Delta_{n,5}-\frac{1}{32}n}{\sqrt{n}} \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \frac{1}{32}\right)$
$\lfloor \ln\ln n \rfloor - 3$	$\frac{8n}{2^{\lfloor \ln\ln n \rfloor}}$	$\frac{8n}{2^{\lfloor \ln\ln n \rfloor}}$	$\frac{\Delta_{n,\lfloor \ln\ln n \rfloor-3}-\frac{8n}{2^{\lfloor \ln\ln n \rfloor}}}{\sqrt{\frac{8n}{2^{\lfloor \ln\ln n \rfloor}}}} \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1)$
$\lfloor \log_2 n \rfloor - 5$	$32 \times 2^{\{\log_2 n\}}$	$32 \times 2^{\{\log_2 n\}}$	$d_{TV}\left(\Delta_{n,\lfloor \log_2 n \rfloor-5}, \text{ Poi}\left(32 \times 2^{\{\log_2 n\}}\right)\right) \to 0$
$\lfloor \log_2 n \rfloor + 5$	$\frac{2^{\{\log_2 n\}}}{32}$	$\frac{2^{\{\log_2 n\}}}{32}$	$d_{TV}\left(\Delta_{n,\lfloor \log_2 n \rfloor+5}, \text{ Poi}\left(\frac{2^{\{\log_2 n\}}}{32}\right)\right) \to 0$
$\lceil \sqrt{n}+\pi \rceil$	$\frac{n}{2^{\lceil \sqrt{n}+\pi \rceil}}$	$\frac{n}{2^{\lceil \sqrt{n}+\pi \rceil}}$	$\Delta_{n,\lceil \sqrt{n}+\pi \rceil} \xrightarrow{a.s.} 0$
$n-5$	$\frac{32}{2^n}$	$\frac{32}{2^n}$	$\Delta_{n,n-5} \xrightarrow{a.s.} 0$

Table 2. The limiting value of the outdegree of the first few entries in an exponentially preferential tree with radix 2

i	$\lim_{n\to\infty} \mathbb{E}[\Delta_{n,i}]$
1	1.607
2	1.213
3	1.093
4	1.044
5	1.021
6	1.010
7	1.005
8	1.002
9	1.001
10	1.000
11	1.000
12	1.000

3.6 The supercritical regime

As an instance, take $a = 2$. Over an extended period of time, the average root outdegree converges: $\mathbb{E}[\Delta_{n,1}] \to c_1^{(1)}(2) \approx 1.607$. Table 2 shows the asymptotic outdegree for the first 12 entries in the tree, approximated to three decimal places.

A late entry in the tree, such as $i = n - \lfloor n^{1/4} \rfloor$, has an average outdegree

$$\mathbb{E}[\Delta_{n,i}] = 1 + O\left(\frac{1}{2n^{1/4}}\right)$$

A much later entrant, such as a node with with the index $n - 4$ has an average outdegree

$$\mathbb{E}[\Delta_{n,i}] = 1 - \frac{1}{2^4} + O\left(\frac{1}{2^n}\right) = \frac{15}{16} + O\left(\frac{1}{2^n}\right).$$

We discern a "thinning" occurring in the tree. At a node with a high index i, the subtree of descendants is tapered almost into a path.

At $a=2$, cancellations occur in the formula in the expression in Theorem 3.1 (iiia), greatly simplifying the calculation for the root outdegree (also its degree, the case $i=1$) and giving transparency into the limiting distribution:

$$\begin{aligned}\mathbb{P}(\Delta_{n,1}=k) &\to \mathbb{P}(\Delta_i^*=k)\\ &= \lim_{n\to\infty}\Big(\Big(\prod_{\ell=1}^{n-1}\frac{1}{2^\ell-1}\Big)\sum_{1\le j_1<j_2<\cdots j_k\le n-1}\ \prod_{m\notin\{j_1,\ldots,j_k\}}(2^m-2)\Big).\end{aligned}$$

Note that if $j_1\neq 1$, the product retains $m=1$ as one of its indices and $2^m-2=0$ for this value of m, annihilating the entire part of the expression with $j_1=1$ in the sum, simplifying it further to

$$\begin{aligned}\mathbb{P}(\Delta_1^*=k) &= \lim_{n\to\infty}\Big(\Big(\prod_{\ell=1}^{n-1}\frac{1}{2^\ell-1}\Big)\sum_{1<j_2<\cdots<j_k\le n-1}\ \prod_{m\notin\{1,j_2,\ldots,j_k\}}2(2^{m-1}-1)\Big)\\ &= \lim_{n\to\infty}\Big(\Big(\prod_{\ell=1}^{n-1}\frac{1}{2^\ell-1}\Big)\sum_{1<j_2<\cdots<j_k\le n-1}\frac{2^{n-k-1}\prod_{m=2}^{n-1}(2^{m-1}-1)}{(2^{j_2-1}-1)\cdots(2^{j_k-1}-1)}\Big)\\ &= \lim_{n\to\infty}\Big(\frac{2^{n-k-1}}{2^{n-1}-1}\sum_{1<j_2<\cdots<j_k\le n-1}\frac{1}{(2^{j_2-1}-1)\cdots(2^{j_k-1}-1)}\Big)\\ &= \frac{1}{2^k}\sum_{1<j_2<\cdots<j_k\le\infty}\frac{1}{(2^{j_2-1}-1)\cdots(2^{j_k-1}-1)}.\end{aligned}$$

The first few values in the sequence $\mathbb{P}(\Delta_1^*=k)$ are[7]

$$\begin{aligned}\mathbb{P}(\Delta_1^*=1) &= \frac{1}{2};\\ \mathbb{P}(\Delta_1^*=2) &= \frac{1}{4}\sum_{\ell=2}^{\infty}\frac{1}{2^\ell-1}\approx 0.1516737881\ldots\\ \mathbb{P}(\Delta_1^*=3) &= \frac{1}{8}\sum_{\ell=2}^{\infty}\sum_{m=\ell+1}^{\infty}\frac{1}{(2^\ell-1)(2^m-1)}\approx 0.01442126698\ldots.\end{aligned}$$

The first three values alone contain about 0.6661 of the mass of the limiting distribution. The root comes early and stays the longest in the tree. So, it has repeated chances for recruiting. However, the probabilities are very quickly diminished by the appearance of nodes of higher indices, with higher chances of recruiting. While the limit distribution of the root outdegree (which is also the degree) is supported on $\mathbb{N}$, there is a high probability of remaining small (confined to the values 1,2,3).

The probability formula in the intermediate phase of the supercritical regime is unwieldy (Theorem 3.1 (iiib)), yet it can be used to tell us something about the asymptotic structure of the tree. With $a=2$ and $i=n-\lfloor\ln n\rfloor$, we can compute the probability of the intermediate node i in the following way. Here, $b(n)$ is $\lfloor\ln n\rfloor$. For $k=0$, the set $\{j_1,\ldots,j_0\}$ is empty, and the only product that stands is the one on the full set $\{n-h(n),\ldots,n-1\}$. We have

$$\mathbb{P}(\Delta_{n,n-\lfloor\ln n\rfloor}=0)=\frac{1}{\prod_{r=n-\lfloor\ln n\rfloor}^{n-1}(2^r-1)}\prod_{m=n-\lfloor\ln n\rfloor}^{n-1}\big(2^m-1-2^{n-\lfloor\ln n\rfloor-1}\big).$$

At $n = 1000$, this probability is about 0.2933.

The very late nodes have the lion's share. Consider $a = 2$, and the tree when the size is some large n. The outdegree of node n is 0, as it has not recruited yet (which is consistent with zero mean and zero variance as given by the exact and asymptotic formulas). Being of the second to highest index in the tree, node $n - 1$, has a chance of recruiting node n and has a chance of missing. The probability of node $n - 1$ recruiting node n is $2^{n-1}(2 - 1)/(2^n - 1) \to 1/2$. Thus, the outdegree of the penultimate node is asymptotically distributed like a Bernoulli(1/2) random variable.

Node $n - 2$ has two chances at recruiting by time n and has an asymptotic distribution on $\{0, 1, 2\}$ with mean 1/4, and so on. According to Theorem 3.1 (iiic), we have

$$\mathbb{P}(\Delta^\star_{n,n-2} = k) \to \mathbb{P}(\Delta^\star_2 = k) \to \frac{1}{2^3} \sum_{1 \le r_1 < r_2 < \cdots < r_k \le 2} \prod_{\substack{1 \le s \le 2 \\ s \notin \{r_1,\ldots,r_k\}}} (2^s - 1),$$

for $k = 0, 1, 2$. For $k = 0$, the set $r_1, r_2, \ldots, r_0$ is empty, and we compute

$$\mathbb{P}(\Delta^\star_2 = 0) \to \frac{1}{8} \prod_{\substack{1 \le s \le 2 \\ s \notin \phi}} (2^s - 1) = \frac{1 \times 3}{8} = \frac{3}{8}.$$

Further, we have

$$\mathbb{P}(\Delta^\star_2 = 1) \to \frac{1}{8} \sum_{1 \le k_1 \le 2} \prod_{\substack{1 \le s \le 2 \\ s \notin \{k_1\}}} (2^s - 1) = \frac{3 + 1}{8} = \frac{4}{8}.$$

We can find $\mathbb{P}(\Delta^\star_2 = 2)$ from $1 - \mathbb{P}(\Delta^\star_2 = 0) - \mathbb{P}(\Delta^\star_2 = 1) = 1/8$. In summary, we have

$$\Delta^\star_2 = \begin{cases} 0, & \text{with probability } 3/8; \\ 1, & \text{with probability } 4/8; \\ 2, & \text{with probability } 1/8. \end{cases}$$

Remark 3.2. *We discussed phases, where the growth of i is systematically increasing toward n. However, there is no limit to how bizarre the sequence $i = i(n)$ can be. For example, $i(n)$ might be a sequence alternating between two (or more values), such as the sequence $i(n) = 5 + (- 1)^n$, in which case the degree of node i does not converge to a limit. Or, $i(n)$ might alternate between low and high values, such as*

$$i(n) = \begin{cases} 1 & n, \text{ even}; \\ \lfloor \ln n + 0.76884 \rfloor, & n \text{ odd}. \end{cases}$$

Even worse, $i(n)$ may not have any structure at all. We reckon that such sequences are not interesting and do not appear in practice.

4. Concluding remarks

We discussed an exponentially preferential model of recursive trees, wherein node i recruits with probability proportional to a^i. The proportionality constant is time dependent, which captures the reality of dynamic change in networks. The radix a governs the behavior of the tree. The case $a = 1$ is critical and corresponds to the standard well-studied uniform model of random recursive trees. This criticality creates three regimes, the subcritical ($0 < a < 1$), the critical ($a = 1$), and the supercritical ($a > 1$). Each regime has its own early, intermediate, and late phases. They are not the same. For instance, the early phase in the subcritical regime extends to $O(\ln n)$, whereas in the critical regime, the early phase stops at $n/i \to \infty$ and in the supercritical regime the early phase is restricted to fixed i.

Gaussian behavior appears only in the early phases of the subcritical and critical regimes. Other behavior is detected, too. For instance, Poisson approximations are the appropriate asymptotic behavior in the intermediate phases of both the subcritical and critical regimes, with oscillations in the case of the subcritical regime.

Consistently, in the late phase of all three regimes (noting they start differently), $\Delta_{n,i}$ converges to a constant, with constant being 0 in the subcritical and critical regimes, while it is a positive constant in the supercritical regime. The intuition behind these constants is that in the subcritical and critical regimes, the nodes in the early and intermediate phases gobble up the largest share of recruits, but in the supercritical regime, the highest labeled nodes in the late stage are the most attractive.

The dependence on the radix a hosts a broad range of applicability. For example, with $a = 1$ alone, the entire class of uniform recursive trees comes in with its plethora of known applications. The range $0 < a < 1$ corresponds to applications where early nodes are the most attractive, such as the growth of companies and subsidiary branches, where the headquarters and its closest branches acquire influence and wealth over time. The case $a > 1$ corresponds to applications where late nodes are the most attractive, such as some social networks in which newcomers are the most anxious to expand their circles.

Future research may extend to deal with other tree properties, such as the maximal degree, the depth of nodes, and the total path length (among many others). We may also consider other weights and generalizations.

Supplementary material. The supplementary material for this article can be found at https://doi.org/10.1017/nws.2025.3.

Notes

1 The superscript r is often dropped, when it is 1.
2 The number $\gamma \approx 0.5772$ is Euler–Mascheroni constant.
3 The reader should be alerted to that the words "early," "intermediate," and "late" mean different things in the different regimes.
4 One should take note that i is a node index and is always an integer.
5 The ultimate formula is unwieldy, but can be used to discover the probability for small k, such as, for example that an intermediate node is a leaf ($k = 0$).
6 The ultimate formula is unwieldy, but can be used to discover the probability for small k, such as, for example that an intermediate node is a leaf ($k = 0$).
7 For $k = 1$, the sum is for a product that does not exist, to be interpreted as 1.

References

Barabási, A., & Albert, R. (1999). Emergence of scaling in random networks. *Science*, *286*(5439), 509–512.

Barbour, A., & Hall, P. (1984). On the rate of Poisson convergence. *Mathematical Proceedings of Cambridge Philosophical Society*, *95*(03), 473–480.

Barbour, A., Holst, L., & Janson, S. (1992). *Poisson Approximation*. Oxford, UK: Clarendon Press.

Bergeron, F., Flajolet, P., & Salvy, B. (1992). Varieties of increasing trees. In *Lecture Notes in Computer Science*, Vol. *581*, (pp. 24–48). Berlin, Germany: Springer.

Billingsley, P. (2012). *Probability and measure*. Anniversary (ed.), Hoboken, New Jersey: Wiley.

Dereich, S., & Ortgiese, M. (2014). Robust analysis of preferential attachment models with fitness. *Combinatorics, Probability and Computing*, *23*(3), 386–411.

Drmota, M. (2009). *Random Trees: An Interplay Between Combinatorics and Probability*. Wien New York, New York: Springer-Verlag.

Frieze, A., & Karoński, M. (2016). *Introduction to Random Graphs* (2nd ed.). Cambridge, UK: Cambridge University Press.

Gastwirth, J., & Bhattacharya, P. (1984). Two probability models of pyramid or chain letter schemes demonstrating that their promotional claims are unreliable. *Operations Research*, *32*(3), 527–536.

Hofri, M., & Mahmoud, H. (2018). *Algorithmics of Nonuniformity: Tools and Paradigms*. Boca Raton, Florida: CRC Press.

Javanian, M., & Asl, M. (2003). Note on the outdegree of a node in random recursive trees. *Journal of Applied Mathematics and Computing*, *13*, 99–103.

Karr, A. (1993). *Probability*. New York: Springer-Verlag.

Kuipers, L., & Niederreiter, H. (1974). *Uniform Distribution of Sequences*. New York: Wiley.

Lyon, M., & Mahmoud, H. (2020). Trees grown under young-age preferential attachment. *Journal of Applied Probability*, *57*(3), 911–927.

Lyon, M., & Mahmoud, H. (2022). Insertion depth in power-weight trees. *Information Processing Letters*, *176*, 106227.

Mahmoud, H. (2019). Local and global degree profiles of randomly grown self-similar hooking networks under uniform and preferential attachment. *Advances in Applied Mathematics*, *111*, 101930.

Mahmoud, H., Smythe, R., & Szymański, J. (1993). On the structure of plane-oriented recursive trees and their branches. *Random Structures & Algorithms*, *4*(2), 151–176.

Najock, D., & Heyde, C. (1982). On the number of terminal vertices in certain random trees with an application to stemma construction in philology. *Journal of Applied Probability*, *19*(3), 675–680.

Smythe, R., & Mahmoud, H. (1995). A survey of recursive trees. *Theory of Probability and Mathematical Statistics*, *51*, 1–27.

Szymański, J. (1987). On a nonuniform random recursive tree. *North-Holland Mathematics Studies*, *144*, 297–306.

Cite this article: Aguech R., Mahmoud H., Mohamed H. and Yang Z. (2025). Exponentially preferential trees. *Network Science* **13**, 1–20. https://doi.org/10.1017/nws.2025.3