COMPETITION OF CONNECTIVITY IN EVOLVING RANDOM TREES

BAS LODEWIJKS

Contents

1. Introduction and motivation

Tree-like mathematical structures play a central role in several fields of mathematics, with applications ranging from phylogenetic trees and social networks to algorithms in computer science. The versatility of mathematical trees lies in their ability to encode hierarchical relationships and the tractability of their analysis.

In this mini-course we will investigate a family of models known as *preferential attachment trees*. These models have been introduced under different names, most famously by Barabási and Albert [3] as a way to model the evolution of the internet. Though these models have become very popular as a general model for real-world networks. See this video by Veritasium for a nice popular science video on modelling real-world networks that also discusses preferential attachment.

Heuristically, preferential attachment trees are created by adding vertices one by one to a tree, and attaching new vertices to existing vertices randomly. Here, one is more likely to attach to vertices that have a large degree, i.e. there is a *preference* to *attach* to high-degree vertices. These simple dynamics lead to (some of) these models to exhibit certain properties that are often also found (to a certain extent) in real-world networks.

The main property that we discuss in these lectures is the formation of hubs. We can think of hubs as vertices in a network (or a preferential attachment tree) that have a very large degree, much larger than the average degree. Think of influencers in social-media networks, who have many more followers, subscribers, or friends compared to the typical user of such platforms. Or a network of protein interaction in a cell, where the proteins form the vertices and connections are formed when two proteins together take part in a chemical reaction. Here, hubs are proteins that can react with many other proteins and thus play an essential role in the functioning of the cell. Or take the human population, where hubs are well-connected people that can dramatically speed up the spread of viruses such as COVID-19.

The formation of hubs can be understood in terms of the *empirical degree distribution*, which denotes the proportion of vertices in a network that has a certain degree. Hubs cause the degree distribution to have a heavy tail. That is, even for large degrees k, much larger than the average degree, say, the proportion of vertices with degree k and larger tends to zero slowly. Note that this is different from the bell curve of a normal distribution, where deviations from the mean decay much faster (at a super-exponential rate).

Take the example in Figure 1. Here, we see the tail of the empirical degree distribution of a collaboration network of condensed matter physicists. The vertices represent researchers in condensed matter physics and two vertices are connected by an edge if the two corresponding physicists have written a scientific article together. We see that, on a log-log scale, the proportion of vertices with degree at least k decays at roughly a linear rate (for k large). This indicates that the proportion p_k of individuals with degree at least k satisfies approximately $p_k \approx k^{-\tau}$ for some $\tau > 0$. This is known as a *power-law* distribution, also called a *scale-free* distribution. There is yet another interesting video of Veritasium here on the emergence of power-laws in many real-world systems.

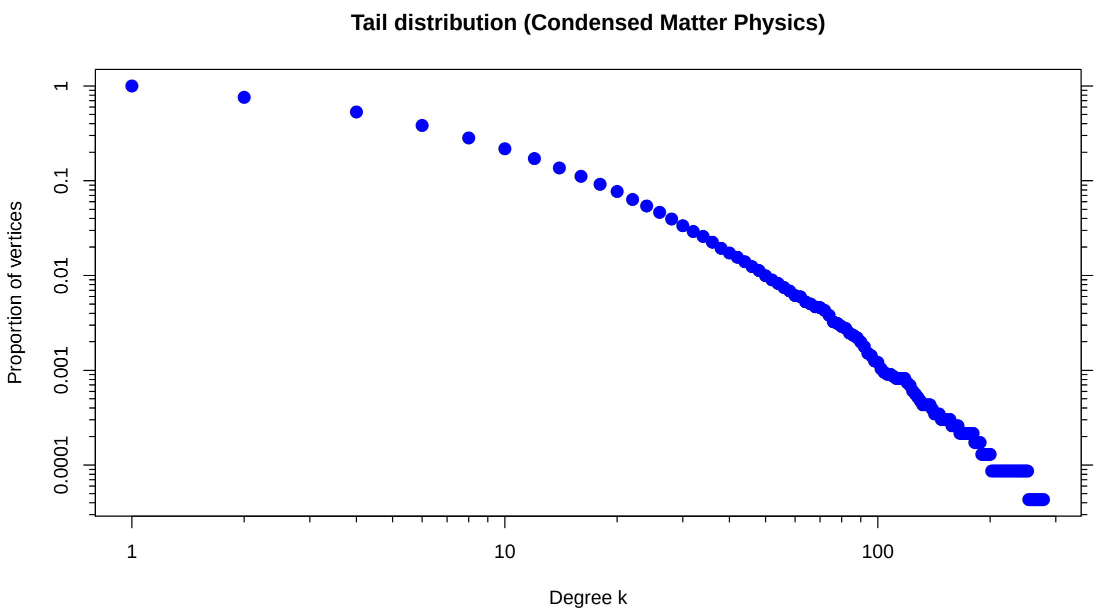

FIGURE 1. The tail of the empirical degree distribution for a collaboration network of condensed matter physicists.

Of course, an exact power-law cannot occur in real-world systems, since that would require degrees of any size to be present, and the blue dots in Figure 1 to line up as a line exactly. Still, the models that we use to understand these system can provide a great insight into their behaviour.

In the next two section, we are going to study preferential attachment models, and investigate the empirical degree distribution and study the *persistence* of hubs.

2. Model definition

Notation. We set $\mathbb{N}_0 \coloneqq \{0, 1, \dots, \}$ and for $n \in \mathbb{N}$ define $[n] \coloneqq \{1, \dots, n\}$. For a tree T, let $V(T)$ and $E(T)$ denote its vertex set and edge set, respectively. For a directed tree T and a vertex $v \in V(T)$, let $\deg(v, T)$ denote the in-degree of v in T. We let $\Gamma\colon (0, \infty) \to \mathbb{R}$ with $\Gamma(t) \coloneqq \int_0^\infty z^{t-1} \mathrm{e}^{-z}\,\mathrm{d}z$ denote the gamma function.

Let $f\colon \mathbb{N}_0 \to [0, \infty)$ such that $f(0) > 0$. We shall refer to f as the *attachment function.* We define the preferential attachment model as follows.

Definition 2.1 (Preferential attachment model)**.** Let $f\colon \mathbb{N}_0 \to [0, \infty)$ such that $f(0) > 0$ be the attachment function. We construct a sequence of trees $(T_n)_{n\in\mathbb{N}}$ as follows.

- Let T_1 denote a single isolated vertex.
- For each $n \geq 2$, we construct T_n conditionally on T_{n-1}. Select a vertex v in T_{n-1} with probability
$$\frac{f(\deg(v, T_{n-1}))}{\sum_{u\in V(T_{n-1})} f(\deg(u, T_{n-1}))}.$$
Construct T_n from T_{n-1} by adding a directed edge from n to v.

3. Uniform attachment and affine preferential attachment: The degree distribution

In this section we focus on two special cases of the preferential attachment:

- The *uniform attachment* model, where $f(i) = 1$ for all $i \in \mathbb{N}_0$,
- The *affine preferential attachment* model, where $f(i) = i + \delta$ for some $\delta > 0$.

We can treat these two models (almost) at once by setting $f(i) = \theta i + (1 - \theta)$ with $\theta \in [0, 1)$. We obtain the uniform attachment model by taking $\theta = 0$ and, as the model definition is invariant under scalar multiplication of the attachment function f, we obtain the affine preferential attachment model by setting $\delta = (1 - \theta)/\theta$ (or $\theta = (1 + \delta)^{-1}$).

In this section we investigate the degrees of (fixed) vertices and the empirical degree distribution $P_n(k)$, which is defined as

$$P_n(k) \coloneqq \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}_{\{\deg(v, T_n)=k\}} \qquad \text{for } k \in \mathbb{N}_0. \tag{3.1}$$

3.1. Degree growth of fixed vertices.

Take $v \in [n]$. We start by calculating the expected in-degree of vertex v in T_n.

Proposition 3.1. *Fix $v \in \mathbb{N}$. For $\theta > 0$, there exists a random variable ξ_v such that*

$$\deg(v, T_n) n^{-\theta} \xrightarrow{\text{a.s.}} \xi_v.$$

When $\theta = 0$, let N be a standard normal random variable. Then,

$$\frac{\deg(v, T_n) - \log n}{\sqrt{\log n}} \xrightarrow{\text{d}} N.$$

To prove Proposition 3.1, we use the following result.

Lemma 3.2. *Fix $v \in \mathbb{N}$. For $\theta \in (0, 1)$, the random variable*

$$M_n \coloneqq \Big(\deg(v, T_n) + \frac{1-\theta}{\theta}\Big) \prod_{j=v}^{n-1} \Big(1 + \frac{\theta}{j - \theta}\Big)^{-1}$$

is a martingale with respect to $(\deg(v, T_n))_{n\geq v}$.

Proof. Let $(I_k)_{v+1\le k\le n}$ be a sequence of indicator random variables, where I_k equals one if vertex k connects to v. We can then write

$$\deg(v,T_n)=\sum_{k=v+1}^{n} I_k.$$

Conditionally on T_{n-1}, we thus have

$$\mathbb{E}\left[\deg(v,T_n)\,|\,T_{n-1}\right]=\deg(v,T_{n-1})+\mathbb{E}\left[I_n\,|\,T_{n-1}\right]. \tag{3.2}$$

Since $f(i)=\theta i+(1-\theta)$, it follows that

$$\sum_{i=1}^{n} f(\deg(i,T_{n-1}))=\theta\sum_{i=1}^{n}\deg(i,T_{n-1})+(1-\theta)(n-1)=(n-1)-\theta.$$

As a result,

$$\mathbb{E}\left[I_n\,|\,T_{n-1}\right]=\frac{\theta\deg(v,T_{n-1})+(1-\theta)}{(n-1)-\theta}.$$

Using this in (3.2) yields

$$\begin{aligned}\mathbb{E}\left[\deg(v,T_n)+\frac{1-\theta}{\theta}\,\middle|\,T_{n-1}\right]&=\Big(1+\frac{\theta}{(n-1)-\theta}\Big)\deg(v,T_{n-1})+\frac{1-\theta}{(n-1)-\theta}+\frac{1-\theta}{\theta}\\&=\Big(1+\frac{\theta}{(n-1)-\theta}\Big)\Big(\deg(v,T_{n-1})+\frac{1-\theta}{\theta}\Big).\end{aligned}$$

Note that it also suffices to condition on $\deg(v,T_{n-1})$ only, not on the entire tree T_{n-1}. By the definition of M_n, it thus follows that

$$\mathbb{E}\left[M_n\mid\deg(v,T_{n-1})\right]=\Big(\deg(v,T_{n-1})+\frac{1-\theta}{\theta}\Big)\prod_{j=v}^{n-2}\Big(1+\frac{\theta}{j-\theta}\Big)^{-1}=M_{n-1},$$

as desired. Since it is clear that $\mathbb{E}\left[|M_n|\right]<\infty$ for all $n\ge v$, since $\deg(v,T_n)<n$, we have that M_n is a martingale with respect to $(\deg(v,T_n))_{n\ge v}$. ☐

To apply the result in Proposition 3.1, we use Doob's well-known martingale convergence theorem.

Theorem 3.3 (Martingale convergence theorem, Theorems 4.2.11 and 4.2.12 in [6])**.** *Let $(X_n)_{n\in\mathbb{N}}$ be a supermartingale, such that*

$$\sup_{n\in\mathbb{N}}\mathbb{E}\left[\max\{0,-X_n\}\right]<\infty.$$

Then, X_n converges almost surely to a limit X with $\mathbb{E}\left[|X|\right]<\infty$. In particular, any positive (super)martingale X_n converges almost surely to a limit X with $\mathbb{E}[X]\le\mathbb{E}[X_1]$.

Proof of Proposition 3.1. Since M_n, as defined in Lemma 3.2, is positive a martingale, it follows from the martingale convergence theorem that there exists a random variable M_∞ such that $M_n\xrightarrow{\text{a.s.}}M_\infty$. Moreover, for $\theta>0$, we have

$$\prod_{j=v}^{n-1}\Big(1+\frac{\theta}{j-\theta}\Big)^{-1}=\prod_{j=v}^{n-1}\frac{j-\theta}{j}=\frac{\Gamma(v)\Gamma(n-\theta)}{\Gamma(v-\theta)\Gamma(n)},$$

where Γ denotes the gamma function. Using Stirling's approximation, we can write

$$\frac{\Gamma(n-\theta)}{\Gamma(n)}=n^{-\theta}(1+\mathcal{O}(1/n)),$$

see also Exercise 3.2. Hence,

$$\lim_{n\to\infty}n^{\theta}\prod_{j=v}^{n-1}\Big(1+\frac{\theta}{j-\theta}\Big)^{-1}=\frac{\Gamma(v)}{\Gamma(v-\theta)}. \tag{3.3}$$

Combined with the almost sure convergence of M_n, we obtain that

$$\deg(v,T_n)n^{-\theta} \xrightarrow{\text{a.s.}} \frac{\Gamma(v-\theta)}{\Gamma(v)} M_\infty =: \xi_v.$$

For the case $\theta = 0$, let again $(I_k)_{v+1\le k\le n}$ be a sequence of indicator random variables, where I_k equals one if vertex k connects to v. We can then write

$$\deg(v,T_n) = \sum_{k=v+1}^{n} I_k.$$

We observe in this case that $\mathbb{P}(I_k = 1) = 1/(k-1)$ and that the I_k are mutually independent. We thus have

$$\mathbb{E}\left[\deg(v,T_n)\right] = \sum_{k=v+1}^{n} \frac{1}{k-1} = \log n + \gamma - \sum_{k=1}^{v-1}\frac{1}{k} + o(1) = \log n + \mathcal{O}(1), \tag{3.4}$$

and

$$\mathrm{Var}(\deg(v,T_n)) = \sum_{k=v+1}^{n} \frac{1}{k-1}\Big(1-\frac{1}{k-1}\Big) = \log n + \gamma - \sum_{k=1}^{v-1}\frac{1}{k} - \sum_{k=v}^{n}\frac{1}{k^2} + o(1) = \log n + \mathcal{O}(1),$$

where γ denotes the Euler-Mascheroni constant. Since the degree of v equals a sum of independent indicator random variables, it is readily verified that the conditions for Lindeberg's central limit theorem are satisfied (see Exercise 3.4). Combined with Slutsky's theorem, it thus follows that

$$\frac{\deg(v,T_n) - \log n}{\sqrt{\log n}} = \frac{\deg(v,T_n) - \mathbb{E}\left[\deg(v,T_n)\right]}{\sqrt{\mathrm{Var}(\deg(v,T_n))}}(1+o(1)) + o(1) \xrightarrow{\text{d}} N,$$

which concludes the proof. □

Since $\deg(v,T_v) = 0$, it follows directly that $\mathbb{E}\left[M_n\right] = \mathbb{E}\left[M_v\right] = (1-\theta)/\theta$ when $\theta > 0$, and hence that

$$\mathbb{E}\left[\deg(v,T_n)\right] = \frac{1-\theta}{\theta}\Bigg(\prod_{j=v}^{n-1}\Big(1+\frac{\theta}{j-\theta}\Big) - 1\Bigg).$$

We thus observe that $\mathbb{E}\left[\deg(v,T_n)\right]$ is of the order $(n/v)^\theta$ when $\theta > 0$ by using (3.3). Similarly, by using (3.4), we see that $\mathbb{E}\left[\deg(v,T_n)\right]$ is of the order $\log(n/v)$ when $\theta = 0$. Vertices with smaller labels thus have the largest degrees in expectation. However, most vertices have a label of order n, so that the majority of vertices has a rather small degree in T_n (in expectation). To this end we study the empirical degree distribution $P_n(k)$, as defined in (3.1), which gives us the proportion of vertices that have degree k for each fixed $k \in \mathbb{N}_0$.

3.2. **The degree distribution.** Fix $k \in \mathbb{N}$ and recall the empirical degree distribution $P_n(k)$ from (3.1). We have the following result.

Theorem 3.4 (Convergence of the empirical degree distribution)**.** *Let $\theta \ge 0$ and consider a preferential attachment tree T_n, as in Definition 2.1, with attachment function $f(i) = \theta i + (1-\theta)$. Recall the empirical degree distribution $P_n(k)$ for $k \in \mathbb{N}_0$ from (3.1). Then,*

$$\sup_{k\in\mathbb{N}_0} |P_n(k) - p_k| \xrightarrow{\text{a.s.}} 0,$$

where

$$p_k := \begin{cases} \frac{1}{2-\theta}\frac{\Gamma(k-1+1/\theta)\Gamma(2/\theta)}{\Gamma(1/\theta-1)\Gamma(k+2/\theta)} & \text{if } \theta > 0,\\ 2^{-(k+1)} & \text{if } \theta = 0.\end{cases}$$

Remark 3.5. The proof of this theorem is an adaptation of [8, Section 8.4], where the case $\theta > 0$ is treated (though also for non-tree preferential attachment models, see also Exercise 3.4). ◀

First, we investigate the mean empirical degree distribution. We observe that the number of vertices with in-degree $k \neq 0$ increases by one when a new vertex connects to an existing vertex with in-degree $k-1$ and decreases by one when a new vertex connects to an existing vertex with in-degree k. For the special case $k = 0$ (i.e. the number of leaves), it increases by one when a new vertex does not connect to a leaf and stays the same otherwise. Hence, with $N_n(k) \coloneqq nP_n(k)$ for $k \in \mathbb{N}_0$ and $N_n(-1) \coloneqq 0$, the number of vertices with in-degree k in T_n, we have

$$\mathbb{E}\left[N_{n+1}(k) - N_n(k) \,|\, T_n\right] = -N_n(k)\frac{\theta k + (1-\theta)}{n-\theta} + N_n(k-1)\frac{\theta(k-1) + (1-\theta)}{n-\theta}\mathbb{1}_{\{k\geq 1\}} + \mathbb{1}_{\{k=0\}}.$$

By taking the expected value with respect to T_n on both sides, and writing $\overline{N}_n(k) \coloneqq \mathbb{E}\left[N_n(k)\right]$, we thus obtain

$$\overline{N}_{n+1}(k) - \overline{N}_n(k) = -\overline{N}_n(k)\frac{\theta k + (1-\theta)}{n-\theta} + \overline{N}_n(k-1)\frac{\theta(k-1) + (1-\theta)}{n-\theta}\mathbb{1}_{\{k\geq 1\}} + \mathbb{1}_{\{k=0\}}. \tag{3.5}$$

If we believe that $\overline{N}_n(k) \approx np_k$, then $\overline{N}_{n+1}(k) - \overline{N}_n(k) \approx p_k$ and $\overline{N}_n(k)/(n-\theta) \approx p_k$, so that we obtain that p_k should satisfy the recurrence relation

$$p_k = -(\theta k + (1-\theta))p_k + (\theta(k-1) + (1-\theta))p_{k-1}\mathbb{1}_{\{k\geq 1\}} + \mathbb{1}_{\{k=0\}}. \tag{3.6}$$

By reordering, we obtain

$$p_k = \frac{\theta(k-1) + (1-\theta)}{\theta(k+1) + 2(1-\theta)}p_{k-1}\mathbb{1}_{\{k\geq 1\}} + \frac{1}{\theta(k+1) + 2(1-\theta)}\mathbb{1}_{\{k=0\}}.$$

By recursively applying the equality (with $p_{-1} = 0$), we thus obtain

$$p_k = \frac{1}{2-\theta}\prod_{\ell=1}^{k}\frac{\theta(\ell-1) + (1-\theta)}{\theta(\ell+1) + 2(1-\theta)} = \frac{1}{2-\theta}\prod_{\ell=1}^{k}\frac{\ell - 2 + 1/\theta}{\ell - 1 + 2/\theta} = \frac{1}{2-\theta}\frac{\Gamma(k-1+1/\theta)\Gamma(2/\theta)}{\Gamma(1/\theta - 1)\Gamma(k+2/\theta)} \tag{3.7}$$

for $\theta > 0$, and

$$p_k = 2^{-(k+1)} \tag{3.8}$$

for $\theta = 0$, as desired.

It remains to show that (1) $\overline{N}_n(k)$ is close to np_k and (2) $\overline{N}_n(k)$ is close to $N_n(k)$. Let us set

$$\varepsilon_n(k) \coloneqq \overline{N}_n(k) - np_k.$$

Our goal is the following result.

Lemma 3.6. *For $\theta \geq 0$, there exists $C > 0$ such that*

$$\sup_{k\in\mathbb{N}_0}\sup_{n\in\mathbb{N}}|\varepsilon_n(k)| \leq C.$$

Proof. We can use the recurrence relation in (3.6) to write

$$\begin{aligned}(n+1)p_k &= np_k + p_k\\ &= np_k - (\theta k + (1-\theta))p_k + (\theta(k-1) + (1-\theta))p_{k-1}\mathbb{1}_{\{k\geq 1\}} + \mathbb{1}_{\{k=0\}}\\ &= np_k - \frac{\theta k + (1-\theta)}{n-\theta}np_k + \frac{\theta(k-1) + (1-\theta)}{n-\theta}np_{k-1}\mathbb{1}_{\{k\geq 1\}} + \mathbb{1}_{\{k=0\}}\\ &\quad - \Big(1 - \frac{n}{n-\theta}\Big)(\theta k + (1-\theta))p_k + \Big(1 - \frac{n}{n-\theta}\Big)(\theta(k-1) + (1-\theta))p_{k-1}\mathbb{1}_{\{k\geq 1\}}\\ &= np_k - \frac{\theta k + (1-\theta)}{n-\theta}np_k + \frac{\theta(k-1) + (1-\theta)}{n-\theta}np_{k-1}\mathbb{1}_{\{k\geq 1\}} + \mathbb{1}_{\{k=0\}}\\ &\quad - \frac{\theta}{n-\theta}\big[-(\theta k + (1-\theta))p_k + (\theta(k-1) + (1-\theta))p_{k-1}\mathbb{1}_{\{k\geq 1\}}\big].\end{aligned} \tag{3.9}$$

By defining

$$\kappa_n(k) \coloneqq -\frac{\theta}{n-\theta}\big[-(\theta k + (1-\theta))p_k + (\theta(k-1) + (1-\theta))p_{k-1}\mathbb{1}_{\{k\geq 1\}}\big], \tag{3.10}$$

we can subtract (3.9) from (3.5) to obtain

$$\varepsilon_{n+1}(k) = \Big(1 - \frac{\theta k + (1-\theta)}{n-\theta}\Big)\varepsilon_n(k) + \frac{\theta(k-1) + (1-\theta)}{n-\theta}\varepsilon_n(k-1)\mathbb{1}_{\{k\geq 1\}} + \kappa_n(k). \tag{3.11}$$

Let us first bound $|\kappa_n(k)|$. From its definition in (3.10) and since $p_k \leq 1$ for all $k \in \mathbb{N}_0$ and

$$\begin{aligned}(\theta(k-1) + (1-\theta))p_{k-1} &= \frac{\theta(k-1)+(1-\theta)}{2-\theta}\prod_{\ell=1}^{k-1}\frac{\theta(\ell-1)+(1-\theta)}{\theta(\ell+1)+2(1-\theta)} \\ &= \frac{\theta(k+1)+2(1-\theta)}{2-\theta}\prod_{\ell=1}^{k}\frac{\theta(\ell-1)+(1-\theta)}{\theta(\ell+1)+2(1-\theta)} \\ &= (\theta(k+1)+2(1-\theta))p_k,\end{aligned}$$

for all $k \in \mathbb{N}$, it is clear that there exists a constant $C_\kappa > 0$ such that

$$\sup_{k\in\mathbb{N}_0}|\kappa_n(k)| \leq \frac{C_\kappa}{n}. \tag{3.12}$$

We prove Lemma 3.6 by showing that there exists a universal constant $C > 0$ such that for each k,

$$\sup_{n\in\mathbb{N}}|\varepsilon_n(k)| \leq C.$$

We prove this by induction in both k and n, using the recursive equality in (3.11). For $k = 0$, we first have $\varepsilon_1(0) = 1 - 1/(2-\theta) = (1-\theta)/(2-\theta)$. Take $C \geq (1-\theta)/(2-\theta)$ and suppose that $|\varepsilon_n(0)| \leq C$ for some n. We then use (3.11) to write

$$\varepsilon_{n+1}(0) = \Big(1 - \frac{1-\theta}{n-\theta}\Big)\varepsilon_n(0) + \kappa_n(0).$$

Since $1 - \frac{1-\theta}{n-\theta} \geq 0$ for all $n \in \mathbb{N}$ and using (3.12), we thus have

$$|\varepsilon_{n+1}(0)| \leq \Big(1 - \frac{1-\theta}{n-\theta}\Big)|\varepsilon_n(0)| + \frac{C_\kappa}{n} \leq C\Big(1 - \frac{1-\theta}{n-\theta}\Big) + \frac{C_\kappa}{n}.$$

It follows that $|\varepsilon_{n+1}(0)| \leq C$ whenever

$$\frac{C_\kappa}{n} - C\frac{1-\theta}{n-\theta} \leq 0$$

which is true for all $n \in \mathbb{N}$ by choosing $C \geq \max\{(1-\theta)/(2-\theta), C_\kappa/(1-\theta)\}$. We thus arrive at

$$\sup_{n\in\mathbb{N}}|\varepsilon_n(0)| \leq C.$$

We now assume that

$$\sup_{n\in\mathbb{N}}|\varepsilon_n(k-1)| \leq C \qquad \text{for some } k \in \mathbb{N},$$

and prove that the same bound then holds for $\varepsilon_n(k)$. To start, we note that $\overline{N}_n(k) = 0$ for all $n \leq k$. Hence, we assume first that $n \geq k+1$, in which case

$$1 - \frac{\theta k + (1-\theta)}{n-\theta} \geq 0.$$

We can hence again use (3.11) to obtain

$$|\varepsilon_{n+1}(k)| \leq \Big(1 - \frac{\theta k + (1-\theta)}{n-\theta}\Big)|\varepsilon_n(k)| + \frac{\theta(k-1)+(1-\theta)}{n-\theta}|\varepsilon_n(k-1)| + |\kappa_n(k)|.$$

Using the induction hypothesis on $\varepsilon_n(k-1)$ and (3.12) yields

$$|\varepsilon_{n+1}(k)| \leq \Big(1 - \frac{\theta k + (1-\theta)}{n-\theta}\Big)|\varepsilon_n(k)| + \frac{C\theta(k-1) + (1-\theta)C + C_\kappa}{n-\theta}. \tag{3.13}$$

We first show that $|\varepsilon_{k+1}(k)| \leq C$. Note that $\overline{N}_{k+1}(k) \leq 1$, as there can be at most one vertex with in-degree k in T_{k+1}. Furthermore,

$$p_k = \frac{1}{2-\theta}\frac{\Gamma(k-1+1/\theta)\Gamma(2/\theta)}{\Gamma(1/\theta-1)\Gamma(k+2/\theta)} = \frac{1}{2-\theta}\frac{\Gamma(2/\theta)}{\Gamma(1/\theta-1)}k^{-(1+1/\theta)}\Big(1+\mathcal{O}\Big(\frac{1}{k}\Big)\Big),$$

when $\theta > 0$, and

$$p_k = 2^{-(k+1)}$$

when $\theta = 0$. Hence,

$$|\varepsilon_{k+1}(k)| \leq 1 + (k+1)p_k \leq C \qquad \text{for any } k \in \mathbb{N}_0, \tag{3.14}$$

if we take C sufficiently large. Now, assume that $|\varepsilon_n(k)| \leq C$ for some $n \geq k+1$. Then, by (3.13)

$$|\varepsilon_{n+1}(k)| \leq \frac{1}{n-\theta}\big[(n-\theta)C - C\theta + C_\kappa\big],$$

which is at most C when $C \geq C_\kappa/\theta$. This proves that

$$\sup_{n\geq k+1} |\varepsilon_n(k)| \leq C. \tag{3.15}$$

We now take care of values $n < k+1$. Recall that $\overline{N}_n(k)$ equals zero for such values of n, so that $|\varepsilon_n(k)| = np_k$. We again using the asymptotic expressions for p_k in (3.7) and (3.8). First, we have that p_k is decreasing in k for both $\theta > 0$ and $\theta = 0$. The latter case is directly clear, as $p_k = 2^{-(k+1)}$ for $\theta = 0$. In the former case, we can write

$$p_{k+1} = \frac{k-1+1/\theta}{k+2/\theta}p_k \leq p_k.$$

Hence, we have for $n \leq k$, that

$$|\varepsilon_n(k)| = np_k \leq np_n \leq C,$$

where the final step follows from (3.14) and the choice of C. Combined with (3.15), we finally arrive at

$$\sup_{n\in\mathbb{N}} |\varepsilon_n(k)| \leq C \qquad \text{and thus that} \qquad \sup_{k\in\mathbb{N}_0}\sup_{n\in\mathbb{N}} |\varepsilon_n(k)| \leq C,$$

as desired. □

What remains is to show that $\overline{N}_n(k)$ is close to $N_n(k)$ for k fixed and n large, as in the following lemma.

Lemma 3.7. *Fix $\theta \geq 0$ and $k \in \mathbb{N}_0$. Then, for any $C > 0$,*

$$\mathbb{P}\Big(\sup_{k\in\mathbb{N}_0} |N_n(k) - \overline{N}_n(k)| \geq C\sqrt{n\log n}\Big) = \mathcal{O}(n^{-(C^2/8-1)}).$$

To prove this result, we use the Azuma-Hoeffding inequality for (super)martingales.

Theorem 3.8 (Azuma-Hoeffding inequality)**.** *Let $(X_i)_{i\in\mathbb{N}}$ be a (super)martingale and let $(c_i)_{i\in\mathbb{N}}$ be a sequence of positive real numbers such that*

$$|X_{i+1} - X_i| \leq c_i \qquad \text{almost surely, for each } i \in \mathbb{N}.$$

Then, for any $t > 0$,

$$\mathbb{P}(|X_n - X_1| \geq t) \leq \exp\Big(-\frac{t^2}{2\sum_{i=1}^{n-1} c_i^2}\Big).$$

Proof of Lemma 3.7. We use a concentration results of martingale to prove the desired bound. Define the random variable

$$M_\ell := \mathbb{E}\left[N_n(k) \,|\, T_\ell\right] \qquad \text{for } \ell \in [n].$$

It directly follows from the tower property of the expected value that M_ℓ is a martingale with respect to $(T_\ell)_{\ell\in[n]}$, since $M_\ell \leq n < \infty$ for all $\ell \in [n]$ and

$$\mathbb{E}\left[M_{\ell+1} \,|\, T_\ell\right] = \mathbb{E}\left[\mathbb{E}\left[N_n(k) \,|\, T_{\ell+1}\right] \,|\, T_\ell\right] = \mathbb{E}\left[N_n(k) \,|\, T_\ell\right] = M_\ell.$$

The martingale M_ℓ is called a *Doob martingale*, which satisfies $M_1 = \mathbb{E}[N_n(k)]$ and $M_n = N_n(k)$. Our aim is to apply the Azuma-Hoeffding inequality, as in Theorem 3.8, to $|M_n - M_1|$. To this end, we need to find a sequence $(c_\ell)_{\ell\in[n-1]}$, such that

$$|M_{\ell+1} - M_\ell| \le c_\ell \qquad \text{for all } \ell \in [n-1], \text{ almost surely.}$$

The difference between $M_{\ell+1}$ and M_ℓ lies in the fact that we condition on more information in $M_{\ell+1}$ compared to M_ℓ. By the definition of $N_n(k)$, we can write

$$M_\ell = \sum_{i=1}^{n} \mathbb{P}(\deg(i, T_n) = k \,|\, T_\ell)\,.$$

To bound their difference, we define two preferential attachment tree processes $(T_i)_{i\in[n]}$ and $(T'_i)_{i\in[n]}$, where $T'_i = T_i$ for all $i \le \ell$, and T'_i evolves *independently* from T_i for $i \in \{\ell+1, \dots, n\}$. Their marginal distributions are the same, they agree up to step ℓ, and evolve independently afterwards. Since T'_i evolves independently of T_i for $i > \ell$, we have

$$M_{\ell+1} = \sum_{i=1}^{n} \mathbb{P}(\deg(i, T_n) = k \,|\, T_{\ell+1})$$

and

$$M_\ell = \sum_{i=1}^{n} \mathbb{P}(\deg(i, T'_n) = k \,|\, T'_\ell) = \sum_{i=1}^{n} \mathbb{P}(\deg(i, T'_n) = k \,|\, T_{\ell+1})\,.$$

We then have

$$|M_{\ell+1} - M_\ell| \le \sum_{i=1}^{n} \left|\mathbb{P}(\deg(i, T_n) = k \,|\, T_{\ell+1}) - \mathbb{P}(\deg(i, T'_n) = k \,|\, T_{\ell+1})\right|$$

As the evolution of the degree $\deg(i, T'_t)$ for $t \ge \ell + 1$ depends only on $\deg(i, T'_{\ell+1})$, we can write

$$\mathbb{P}\big(\deg(i, T'_n) = k \,\big|\, T_{\ell+1}\big) = \mathbb{E}\left[\mathbb{P}\big(\deg(i, T'_n) = k \,\big|\, \deg(i, T'_{\ell+1})\big) \,\big|\, T_{\ell+1}\right],$$

and

$$\begin{aligned}\mathbb{P}(\deg(i, T_n) = k \,|\, T_{\ell+1}) &= \mathbb{P}\big(\deg(i, T_n) = k \,\big|\, \deg(i, T_{\ell+1})\big)\\ &= \mathbb{E}\left[\mathbb{P}\big(\deg(i, T_n) = k \,\big|\, \deg(i, T_{\ell+1})\big) \,|\, T_{\ell+1}\right].\end{aligned}$$

As a result,

$$|M_{\ell+1} - M_\ell| \le \sum_{i=1}^{n} \mathbb{E}\left[\left|\mathbb{P}(\deg(i, T_n) = k \mid \deg(i, T_{\ell+1})) - \mathbb{P}\big(\deg(i, T_n) = k \mid \deg(i, T'_{\ell+1})\big)\right| \,\Big|\, T_{\ell+1}\right].$$

We observe that $\mathbb{P}(\deg(i, T_n) = k \mid \deg(i, T_{\ell+1})) = \mathbb{P}\big(\deg(i, T_n) = k \mid \deg(i, T'_{\ell+1})\big)$ when $\deg(i, T_{\ell+1}) = \deg(i, T'_{\ell+1})$, since both tree processes evolve according to the same rules. Hence,

$$\left|\mathbb{P}(\deg(i, T_n) = k \mid \deg(i, T_{\ell+1})) - \mathbb{P}\big(\deg(i, T_n) = k \mid \deg(i, T'_{\ell+1})\big)\right| \le \mathbb{1}_{\{\deg(i, T_{\ell+1}) \ne \deg(i, T'_{\ell+1})\}}.$$

We thus arrive at

$$|M_{\ell+1} - M_\ell| \le \sum_{i=1}^{n} \mathbb{E}\left[\mathbb{1}_{\{\deg(i, T_{\ell+1}) \ne \deg(i, T'_{\ell+1})\}}\right] \le 2,$$

where the final inequality follows from the fact that $\deg(i, T_\ell) = \deg(i, T'_\ell)$ for all $i \in [n]$, since $T_\ell = T'_\ell$, and we change the in-degree of one vertex in each tree in step $\ell + 1$.

We then use the Azuma-Hoeffding inequality with $c_\ell = 2$ to arrive at

$$\begin{aligned}\mathbb{P}\Big(\sup_{k\in\mathbb{N}_0}|N_n(k)-\overline{N}_n(k)|\geq C\sqrt{n\log n}\Big) &= \mathbb{P}\Big(\sup_{k\leq n-1}|N_n(k)-\overline{N}_n(k)|\geq C\sqrt{n\log n}\Big)\\ &\leq n\sup_{k\leq n-1}\mathbb{P}\Big(|N_n(k)-\overline{N}_n(k)|\geq C\sqrt{n\log n}\Big)\\ &\leq 2n\exp\Big(-\frac{C^2 n\log(n)}{8n}\Big)\\ &= \mathcal{O}(n^{-(C^2/8-1)}),\end{aligned}$$

which concludes the proof. □

We are now ready to prove Theorem 3.4.

Proof of Theorem 3.4. By the triangle inequality,

$$\begin{aligned}\sup_{k\in\mathbb{N}_0}|P_n(k)-p_k| &\leq \frac{1}{n}\sup_{k\in\mathbb{N}_0}|N_n(k)-np_k| + \frac{1}{n}\sup_{k\in\mathbb{N}_0}|N_n(k)-\overline{N}_n(k)|\\ &\leq \frac{1}{n}\sup_{k\in\mathbb{N}_0}\sup_{n\in\mathbb{N}}|\varepsilon_n(k)| + \frac{1}{n}\sup_{k\in\mathbb{N}_0}|N_n(k)-\overline{N}_n(k)|.\end{aligned}$$

By combining Lemmas 3.6 and 3.7, we thus obtain for any $C' > C'' > 4$ and all n large, that

$$\begin{aligned}\mathbb{P}\Big(\sup_{k\in\mathbb{N}_0}|P_n(k)-p_k|\geq C'\sqrt{\frac{\log n}{n}}\Big) &\leq \mathbb{P}\Big(C+\sup_{k\in\mathbb{N}_0}|N_n(k)-\overline{N}_n(k)|\geq C'\sqrt{n\log n}\Big)\\ &\leq \mathbb{P}\Big(\sup_{k\in\mathbb{N}_0}|N_n(k)-\overline{N}_n(k)|\geq C''\sqrt{n\log n}\Big) = \mathcal{O}(n^{-(C''/8-1)}).\end{aligned}$$

By the choice of C'', we thus obtain by the Borel-Cantelli lemma that

$$\sup_{k\in\mathbb{N}_0}|P_n(k)-p_k| \overset{\text{a.s.}}{\longrightarrow} 0,$$

which concludes the proof. □

3.3. **Exercises.**

Exercise 3.1. (Gamma function) Fix $z > 0$ and recall that

$$\Gamma(z) \coloneqq \int_0^\infty t^{z-1}\mathrm{e}^{-t}\,\mathrm{d}t. \tag{3.16}$$

Show that $\Gamma(z+1) = z\Gamma(z)$. Use this to conclude that, for $\alpha > 0$

$$\prod_{i=0}^{n}(i+\alpha) = \frac{\Gamma(n+1+\alpha)}{\Gamma(\alpha)}.$$

Exercise 3.2. (Gamma function approximation) With Γ as in (3.16), use Stirling's approximation to show that for any $\alpha\in\mathbb{R}$ fixed,

$$\frac{\Gamma(z+\alpha)}{\Gamma(z)} = z^\alpha\Big(1+\mathcal{O}\Big(\frac{1}{z}\Big)\Big) \qquad \text{as } z\to\infty.$$

Exercise 3.3. (Upper bound on maximum degree of affine preferential attachment) Let $\theta > 0$. Show that for any $k\in\mathbb{N}$ and any $v\in[n]$, the random variable

$$Z_n(v) \coloneqq \binom{\deg(v,T_n)+(1-\theta)/\theta+k-1}{k}\prod_{j=v}^{n-1}\Big(1+\frac{\theta k}{j-\theta}\Big)^{-1}$$

is a martingale with respect to $(\deg(v,T_n))_{n\geq v}$. Here, we define, for $a,b>-1$ such that $a-b>-1$,

$$\binom{a}{b} := \frac{\Gamma(a+1)}{\Gamma(b+1)\Gamma(a-b+1)}.$$

As $\binom{\deg(v,T_n)+(1-\theta)/\theta+k-1}{k} \approx \deg(v,T_n)^k$ for n large, use this to show that

$$\lim_{n\to\infty} \frac{\log(\max_{v\in[n]}\deg(v,T_n))}{\log n} = \theta \qquad \text{almost surely.}$$

Remark 3.9. In fact, it can be proved that $\max_{v\in[n]}\deg(v,T_n)n^{-\theta}$ converges almost surely to $\max_{v\in\mathbb{N}}\xi_v$, with ξ_v as in Lemma 3.1.

Exercise 3.4. (Lindeberg's CLT) Let $(X_k)_{k\in\mathbb{N}}$ be a sequence of independent random variables. Assume that

$$\mu_k := \mathbb{E}[X_k] \quad \text{and} \quad \sigma_k := \mathrm{Var}(X_k)$$

exist and are finite for all $k\in\mathbb{N}$. Set

$$s_n^2 := \sum_{k=1}^n \sigma_k^2.$$

If the sequence of random variables $(X_k)_{k\in\mathbb{N}}$ satisfies Lindeberg's condition: For any $\varepsilon>0$,

$$\lim_{n\to\infty}\frac{1}{s_n^2}\sum_{k=1}^n \mathbb{E}\left[(X_k-\mu_k)^2\mathbb{1}_{\{|X_k-\mu_k|\geq \varepsilon s_n\}}\right]=0,$$

then

$$\frac{\sum_{k=1}^n (X_k-\mu_k)}{s_n} \xrightarrow{\mathrm{d}} N,$$

with N a standard normal random variable.

Check that Lindeberg's condition is satisfied for $\deg(v,T_n)$ and any $v\in\mathbb{N}$ fixed when $\theta=0$.

Exercise 3.5. Define

$$P_n^{\geq}(k) := \sum_{i=k}^{n-1} P_n(i) \quad \text{and} \quad p_{\geq k} := \sum_{i=k}^{\infty} p_i.$$

Find an explicit expression for $p_{\geq k}$ for both $\theta>0$ and $\theta=0$. For $\theta>0$, it could help to prove the following: For $n\in\mathbb{N}$ and $a,b\in\mathbb{R}$ not negative integers and $b\neq 1+a$, show that

$$\sum_{i=0}^n \frac{\Gamma(i+a)}{\Gamma(i+b)} = \frac{1}{1+a-b}\Big(\frac{\Gamma(n+1+a)}{\Gamma(n+b)} - \frac{\Gamma(a)}{\Gamma(b-1)}\Big).$$

By adapting the arguments in Lemmas 3.6 and 3.7, then show that, for any $k\in\mathbb{N}_0$,

$$P_n^{\geq}(k) \xrightarrow{\text{a.s.}} p_{\geq k}.$$

For Lemma 3.6, find a similar recurrence relation as in (3.6) and with $\varepsilon_n^{\geq}(k) := nP_n^{\geq}(k) - np_{\geq k}$, use induction to show that for any $k\in\mathbb{N}_0$ there exists $C_k>0$ such that $|\varepsilon_n^{\geq}(k)|\leq C_k$ for all $n\in\mathbb{N}$. For Lemma 3.7, adapt the argument for the martingale

$$M_\ell^{\geq} := \mathbb{E}\left[N_n^{\geq}(k)\,|\,T_\ell\right].$$

Exercise 3.6. (Uniform and affine preferential attachment graphs) Fix an integer $m\geq 2$ and adjust the preferential attachment model by letting each vertex attach to m vertices. That is, construct a sequence of graphs $(G_n)_{n\in\mathbb{N}}$ where G_1 an isolated vertex, and for each $n\geq 2$ we add a vertex with label n to G_{n-1} and attach, for each $i\in[m]$, its i^{th} edge to a vertex v, conditionally on G_{n-1} and the first $i-1$ edges of n, chosen with probability

$$\frac{\deg(v,i-1,G_{n-1})}{\sum_{u\in V(G_{n-1})}\deg(u,i-1,G_{n-1})},$$

where $\deg(v,i-1,G_{n-1})$ denotes the in-degree of vertex v when vertex n has connected $i-1$ edges to the vertices in G_{n-1}.

All results proved in this section can be adjusted to also hold for the graphs G_n. Can you adjust the definitions and proofs?

3.4. **Further reading.** The linear preferential attachment and uniform attachment model are by far the most studied models among the family of preferential attachment models. This is because of the interesting properties they have, but also because these models tend to be more tractable and somewhat easier to analyse. The two books of Van der Hofstad [8, 9] on random graphs and complex networks provide an excellent overview of the linear preferential attachment model (including plenty of other interesting random graphs!) as well as many references to models related to preferential attachment.

For further reading on the uniform attachment model (also known as the random recursive tree model), here are some suggestions, some of which use different tools and ideas to analyse the model:

- Lodewijks [12] studies various properties of high-degree vertices in the uniform attachment model, such as their label, depth (distance to the root), and the graph distance between high-degree vertices. The analysis relies on a time-reversed construction of the uniform attachment tree called the Kingman coalescent construction.
- Addario-Berry and Ford [1] obtain very precise estimates for the height of the uniform attachment tree (the maximal distance of all leafs to the root). The embedding into continuous-time branching process that we use in Section 4 is used here.
- Briend, Calvillo, and Lugosi [5] aim to find the root vertex in random directed acyclic graphs and linear preferential attachment graphs (versions of the uniform and linear preferential attachment models discussed here, where a new vertex connects to more than one existing vertex) when one does not know the labels of the vertices.

4. Persistent hubs in preferential attachment trees

We have seen in Section 3 that the (expected) degree of a vertex is decreasing in its label for the uniform and linear attachment cases. This is intuitively clear for any function f, since vertices that arrive late have fewer possibilities to create new connection and have to compete with more vertices for these connections. The ‘rich-get-richer’ phenomenon that we have discussed earlier is therefore also often called the ‘old-get-richer’ phenomenon. It is the old vertices that are able to obtain large degrees (become rich) and then grow their degrees even further (get richer).

However, we see from Theorem 3.4 that $p_k \approx k^{-(1+1/\theta)}$ when $\theta > 0$ and $p_k = 2^{-(k+1)}$ when $\theta = 0$. Intuitively, the largest degree k_n in the tree T_n should so large that

$$\sum_{i=k_n}^{\infty} p_i = \frac{1}{n}.$$

That is, $k_n \approx n^\theta$ for $\theta > 0$ and $k_n \approx \log(n)/\log(2)$ for $\theta = 0$. Observe that the size of k_n agrees with the degrees of fixed vertices, as in Proposition 3.1 when $\theta > 0$, but that this is not the case when $\theta = 0$. For $\theta > 0$ our results thus far seem to agree with this ‘old-get-richer’ phenomenon, whereas this seems not to be the case for $\theta = 0$.

A natural question to ask is the following:

Are the vertices with the largest degrees always among the oldest vertices?

To make this question more precise, let us define

$$I_n := \min\{i \in [n] : \deg(i, T_n) \geq \deg(j, T_n) \text{ for all } j \in [n]\}.$$

We say that a *persistent hub* exists in $(T_n)_{n\in\mathbb{N}}$ when there exists an integer-valued random variable I such that

$$I_n \xrightarrow{\text{a.s.}} I.$$

The following result reveals an interesting phase transition based on the attachment function f with regards to the existence of persistent hubs.

Theorem 4.1 (Existence of persistent hubs)**.** *Let $(T_n)_{n\in\mathbb{N}}$ be a sequence of preferential attachment trees with attachment function f, as in Definition 2.1. Assume that*

$$\limsup_{i\to\infty} \frac{f(i)}{i} < \infty. \tag{LS-f}$$

Then, there exists a persistent hub if and only if

$$\sum_{i=0}^{\infty} \frac{1}{f(i)^2} < \infty. \tag{S}$$

Remark 4.2. The proof of the theorem is an adaptation of [11, Theorem 2.10]. ◀

Remark 4.3. The assumption in (LS-f) on the growth-rate is not (strictly) necessary, in the sense that the existence of persistent hubs has been proved for certain families of attachment functions f that grow faster than linear as well. A complete understanding that does not require any other condition other than (S) is currently not present, however. We use condition (LS-f) here because it covers most natural and interesting models, such as the linear preferential attachment model and uniform attachment model discussed in Section 3. ◀

We can directly verify that the assumption in (LS-f) is satisfied for both the uniform and affine preferential attachment cases studied in Section 3. Moreover, the affine case satisfies Condition (S), whereas the uniform case does not, which agrees with the observations we made at the start of this section regarding the degree distribution and the degrees of fixed vertices.

In the remainder of the section, we prove Theorem 4.1.

4.1. **Embedding preferential attachment trees in continuous-time branching processes.**

To make the analysis of general preferential attachment trees easier, we will use a trick first introduced by Athreya and Karlin, and first applied to preferential attachment trees by Rudas Toth and Valko, where one 'embeds' the discrete sequence of trees $(T_n)_{n\in\mathbb{N}}$ into a branching process that evolves in continuous time. We will then analyse the branching process instead, which turns out to have several advantages.

Let us first introduce some notation. We define

$$\mathrm{UH} := \{\varnothing\} \cup \bigcup_{k\in\mathbb{N}} \mathbb{N}^k \tag{4.1}$$

to be the *Ulam-Harris* tree. For $u = u_1\cdots u_k \in \mathrm{UH}$ for some $k\in\mathbb{N}_0$ and $u_1,\ldots,u_k\in\mathbb{N}$, we view u as the u_k^{th} child of $u_1\cdots u_{k-1}$, which is the u_{k-1}^{th} child of $u_1\cdots u_{k-2}$,..., which is the u_1^{th} child of $\varnothing$. In this section, we view a (finite) tree T as a (finite) subset of UH that satisfies:

- $\varnothing \in T$;
- T is parent closed: If $uk\in T$ for some $u\in\mathrm{UH}$ and $k\in\mathbb{N}$, then $u\in T$;
- T is sibling closed: If $uk\in T$ for some $u\in\mathrm{UH}$ and $k\in\mathbb{N}$, then $ui\in T$ for all $i\in[k-1]$.

We define the branching process $(\mathcal{T}_t)_{t\geq 0}$ as follows. For an attachment function f, let $\mathbf{E}=(E_i)_{i\in\mathbb{N}}$ be a sequence of independent exponential random variables, where $E_i\sim\mathrm{Exp}(f(i-1))$ for each $i\in\mathbb{N}$. Here, we have $E_i=\infty$ almost surely when $f(i-1)=0$. We assign to each individual $u\in\mathrm{UH}$ an independent copy $\mathbf{E}^{(u)}$ of $\mathbf{E}$, where the $\mathbf{E}^{(u)}$ are mutually independent among the $u\in\mathrm{UH}$. The random variables $E_i^{(u)}$ are the *inter-birth times* of individual u. That is, $E_i^{(u)}$ denotes the time between the birth of child number $i-1$ and i of individual u. Then, define for each $u\in\mathrm{UH}$ the sequence of partial sums $\mathbf{S}^{(u)}=(S_k^{(u)})_{k\in\mathbb{N}}$, where

$$S_k^{(u)} := \sum_{i=1}^{k} E_i^{(u)} \qquad \text{for } k\in\mathbb{N}. \tag{4.2}$$

We define the *birth-time* of individuals recursively as $B(\varnothing) \coloneqq 0$, and

$$B(uk) \coloneqq B(u) + S_k^{(u)} \qquad \text{for } k \in \mathbb{N}, u \in \text{UH}. \tag{4.3}$$

Finally, we define

$$\mathcal{T}_t \coloneqq \{u \in \text{UH} : B_u \le t\} \qquad \text{for } t \ge 0, \tag{4.4}$$

as the set of all individuals that are born up to time t. Here, we view individuals as being connected by a directed edge to their parents. It is clear that $\mathcal{T}_t$ satisfies all conditions for a tree almost for all $t \ge 0$, almost surely.

The relation between the branching process $(\mathcal{T}_t)_{t\ge 0}$ and the preferential attachment trees $(T_n)_{n\in\mathbb{N}}$ is a follows. First, define the stopping times

$$\tau_n \coloneqq \inf\{t \ge 0 : |\mathcal{T}_t| = n\} \qquad \text{for } n \in \mathbb{N}.$$

Clearly, the τ_n are almost surely increasing in n, so that we can define their limit

$$\tau_\infty \coloneqq \lim_{n\to\infty} \tau_n.$$

Then, define the bijection L which assign a label in $\mathbb{N}$ to the individuals contained in $\mathcal{T}_t$, by listing them in increasing order of their birth-times (i.e. the n^{th} individual born in $\mathcal{T}_t$ receives the label n). Let $L(\mathcal{T}_t)$ denote the branching process $\mathcal{T}_t$ with the individuals relabelled by L. We then have the following result.

Proposition 4.4 (Embedding of preferential attachment trees in branching processes)**.** *We have*

$$\{L(\mathcal{T}_{\tau_n}) : n \in \mathbb{N}\} \overset{\mathrm{d}}{=} \{T_n : n \in \mathbb{N}\}.$$

Proof. We prove the result by induction. It is clear that $L(\mathcal{T}_{\tau_1}) = L(\mathcal{T}_0) = T_1$. Given that $L(\mathcal{T}_{\tau_n}) \overset{\mathrm{d}}{=} T_n$ for some $n \in \mathbb{N}$, we prove that $L(\mathcal{T}_{\tau_{n+1}}) \overset{\mathrm{d}}{=} T_{n+1}$. We view the branching process $(\mathcal{T}_t)_{t\ge 0}$ as follows. Each individual has an exponential clock where, if an individual has $i \in \mathbb{N}_0$ many children, its clock rings with rate $f(i)$, independently of all other individuals. If the clock rings, the individual produces a child and updates the rate of its clock. By the strong Markov property and the memoryless property of the exponential distribution, conditionally on $\mathcal{T}_{\tau_n}$ we construct $\mathcal{T}_{\tau_{n+1}}$ by adding a child to the first individual in $\mathcal{T}_{\tau_{n+1}}$ whose exponential clock rings.

Now, let T be a tree (in the sense of a subset of UH) on n vertices, such that $\mathbb{P}(\mathcal{T}_{\tau_n} = T) > 0$. Suppose that $\mathcal{T}_{\tau_n} = T$. If we let $c(u, T)$ denote the number of children of an individual u in T, it follows that, conditionally on $\mathcal{T}_{\tau_n} = T$, the clock of $u \in T$ rings first among all individuals in T with probability

$$\frac{f(c(u, \mathcal{T}_{\tau_n}))}{\sum_{v\in\mathcal{T}_{\tau_n}} f(c(v, \mathcal{T}_{\tau_n}))} = \frac{f(c(u, T))}{\sum_{v\in T} f(c(v, T))} \qquad \text{for all } u \in T.$$

Observe that

$$\frac{f(c(u, T))}{\sum_{v\in T} f(c(v, T))} = \frac{\deg(u, T)}{\sum_{v\in T} \deg(i, T)}.$$

As a result, the transition probabilities of $L(\mathcal{T}_{\tau_n})$ to $L(\mathcal{T}_{\tau_{n+1}})$ and of T_n to T_{n+1} are equal, so that $L(\mathcal{T}_{\tau_{n+1}}) \overset{\mathrm{d}}{=} T_{n+1}$, which concludes the proof. □

4.2. **Preliminary results on sums of inter-birth times.** Before we are ready to prove Theorem 4.1, we provide several useful technical tools.

Lemma 4.5. *Recall the branching process* $(\mathcal{T}_t)_{t\ge 0}$ *defined in* (4.4)*. Assume that* f *satisfies* (LS-f)*. Then,* $|\mathcal{T}_t| < \infty$ *for all* $t \in [0, \infty)$ *almost surely.*

To prove Lemma 4.5, we first state the following technical result, whose proof we leave to Exercise 4.1.

Lemma 4.6. *Recall the inter-birth time random variables* $(E_i)_{i\in\mathbb{N}}$ *and define the quantity*

$$\mu(\lambda) := \sum_{j=1}^{\infty} \mathbb{E}\left[\exp\Big(-\lambda\sum_{i=1}^{j} E_i\Big)\right] = \sum_{j=1}^{\infty}\prod_{i=1}^{j}\frac{f(i-1)}{f(i-1)+\lambda} \qquad \text{for } \lambda \geq 0. \tag{4.5}$$

When f *satisfies* (LS-f), *there exists* $\lambda > 0$ *such that* $\mu(\lambda) < 1$.

Proof of Lemma 4.5. For any $v = v_1 \cdots v_m \in \mathrm{UH}$, with $m \in \mathbb{N}$ and $v_1, \dots, v_m \in \mathbb{N}$, and any $t > 0$,

$$\mathbb{P}(\mathcal{B}(v) \leq t) = \mathbb{P}\left(\sum_{\ell=1}^{m}\sum_{i=1}^{v_\ell} E_i^{(v_1\cdots v_{\ell-1})} \leq t\right).$$

We use the independence of the inter-birth times and apply Chernoff's inequality with parameter $\lambda > 0$ such that $\mu(\lambda) < 1$ by Lemma 4.6 to bound the probability from above by

$$\mathrm{e}^{\lambda t}\prod_{\ell=1}^{m}\mathbb{E}\left[\exp\Big(-\lambda\sum_{i=1}^{v_\ell} E_i\Big)\right].$$

By summing over all individuals $v \in \mathrm{UH}$ (where $\mathbb{P}(\mathcal{B}(\varnothing) \leq t) = 1$ as $\mathcal{B}(\varnothing) = 0$), we thus obtain

$$\mathbb{E}\left[|\mathcal{T}_t|\right] = \sum_{v\in\mathrm{UH}}\mathbb{P}(\mathcal{B}(v) \leq t) \leq 1 + \mathrm{e}^{\lambda t}\sum_{m=1}^{\infty}\sum_{v_1=1}^{\infty}\cdots\sum_{v_m=1}^{\infty}\prod_{\ell=1}^{m}\mathbb{E}\left[\exp\Big(-\lambda\sum_{i=1}^{v_\ell} E_i\Big)\right].$$

As the terms in the product depend on distinct summation indices, we can interchange summation and product. to yield

$$1 + \mathrm{e}^{\lambda t}\sum_{m=0}^{\infty}\left(\sum_{j=1}^{\infty}\mathbb{E}\left[\exp\Big(-\lambda\sum_{i=1}^{j} E_i\Big)\right]\right)^{m} = 1 + \mathrm{e}^{\lambda t}\sum_{m=1}^{\infty}\mu(\lambda)^m,$$

where we recall $\mu(\lambda)$ from (4.5). Since $\mu(\lambda) < 1$, the expected value is finite. Hence, $|\mathcal{T}_t| < \infty$ almost surely. □

Let $\mathbf{E} = (E_i)_{i\in\mathbb{N}}$ and $\mathbf{E}' = (E_i')_{i\in\mathbb{N}}$ be two i.i.d. sequences of exponential random variables, with $E_i \sim \mathrm{Exp}(f(i))$. Define

$$Y_i := E_i - E_i' \qquad \text{for } i \in \mathbb{N}. \tag{4.6}$$

We have the following results for the symmetric random variables Y_i.

Lemma 4.7. *Consider the random variables* $(Y_i)_{i\in\mathbb{N}}$*, as in* (4.6). *We have that*

$$\sum_{i=1}^{\infty} Y_i \text{ converges almost surely if and only if } \sum_{i=0}^{\infty}\frac{1}{f(i)^2} < \infty.$$

Lemma 4.8 (Proposition 1.16 in [10]). *Consider the random variables* $(Y_i)_{i\in\mathbb{N}}$*, as in* (4.6). *If* $\sum_{i=1}^{\infty} Y_i$ *diverges almost surely, then*

$$\limsup_{k\to\infty}\sum_{i=1}^{k} Y_i = \infty \qquad \text{almost surely,}$$
$$\liminf_{k\to\infty}\sum_{i=1}^{k} Y_i = -\infty \qquad \text{almost surely.}$$

The proof of both results relies upon Kolmogorov's three series theorem, which provides necessary and sufficient conditions for the convergence of random sums.

Theorem 4.9 (Kolmogorov's three series theorem, Theorem 2.5.8 in [6]). *Let* $(X_i)_{i\in\mathbb{N}}$ *be a sequence of independent random variables. Fix* $C > 0$ *and set* $Z_i := X_i \mathbb{1}_{\{|X_i|\leq C\}}$. *Then,* $\sum_{i=1}^{\infty} X_i$ *converges almost surely if for some* $C > 0$,

$$(i)\quad \sum_{i=1}^{\infty}\mathbb{P}(|X_i| \geq C) < \infty \qquad (ii)\quad \sum_{i=1}^{\infty}\mathbb{E}\left[Z_i\right] < \infty \qquad (iii)\quad \sum_{i=1}^{\infty}\mathrm{Var}(Z_i) < \infty.$$

If $\sum_{i=1}^\infty X_i$ converges almost surely, then (i), (ii), and (iii) hold for any $C>0$.

We now prove the two lemmas.

Proof of Lemma 4.7. Without loss of generality we can assume that $f(i)>0$ for all $i\in\mathbb{N}_0$, so that each Y_i is finite almost surely. For the 'if' direction, we compute the moment generating function of $\sum_{i=K}^\infty Y_i$ for some large K. Take $\lambda>0$ and K large such that $f(i-1)>\lambda$ for all $i\geq K$. Note that this is possible since f tends to infinity when we assume that $\sum_{i=0}^\infty 1/f(i)^2<\infty$. Then,

$$\begin{aligned}\mathbb{E}\left[\exp\left(\lambda\sum_{i=K}^\infty Y_i\right)\right]&=\prod_{i=K}^\infty \mathbb{E}\left[\mathrm{e}^{\lambda(E_i-\lambda E_i')}\right]\\&=\prod_{i=K}^\infty\frac{f(i-1)^2}{(f(i-1)-\lambda)(f(i-1)+\lambda)}=\prod_{i=K}^\infty\frac{1}{1-\lambda/f(i-1)^2}.\end{aligned}\tag{4.7}$$

It follows from standard results in analysis, this infinite product is finite if and only if $\sum_{i=K}^\infty 1/f(i)^2<\infty$. As we assume that $f(i)>0$ for all i, we have that $\sum_{i=1}^{K-1}Y_i<\infty$ almost surely as well, so that we obtain the desired result.

For the 'only if' direction, we use the three-series theorem of Kolmogorov, which states that $\sum_{i=1}^\infty Y_i$ of independent random variables Y_i converges almost surely if and only if for any $C>0$,

$$\sum_{i=1}^\infty \mathbb{P}(|Y_i|\geq C)<\infty,\quad \sum_{i=1}^\infty \mathbb{E}\left[Y_i\mathbb{1}_{\{|Y_i|\leq C\}}\right]<\infty,\quad\text{and}\quad \sum_{i=1}^\infty \mathrm{Var}(Y_i\mathbb{1}_{\{|Y_i|\leq C\}})<\infty.$$

As the Y_i are symmetric, the second condition is always satisfied. We argue by contradiction and assume that

$$\sum_{i=1}^\infty Y_i\text{ diverges almost surely but }\sum_{i=0}^\infty\frac{1}{f(i)^2}<\infty.$$

The first part, combined with the three-series theorem implies that, for any $C>0$,

$$\text{either}\quad\sum_{i=1}^\infty\mathbb{P}(|Y_i|\geq C)=\infty\quad\text{or}\quad\sum_{i=1}^\infty\mathrm{Var}(Y_i\mathbb{1}_{\{|Y_i|\leq C\}})=\infty.$$

We show that either of these two cases leads to a contradiction. As $Y_i=E_i-E_i'$, the first case is is equivalent to

$$\sum_{i=1}^\infty 2\mathbb{P}(E_i\geq E_i'+C)=\infty.$$

The probability in each term equals

$$\int_0^\infty f(i-1)\mathrm{e}^{-f(i-1)s-f(i-1)(s+C)}\,\mathrm{d}s=\frac{1}{2}\mathrm{e}^{-f(i-1)C}.$$

Hence, it follows that

$$\sum_{i=0}^\infty \mathrm{e}^{-f(i)C}=\infty.$$

Now, since we assumed that $\sum_{i=0}^\infty 1/f(i)^2$ is finite, it follows that $f(i)$ tends to infinity with i, so that $f(i)^{-2}\geq \mathrm{e}^{-f(i)C}$ for all $i\geq I$ and some finite $I=I(C,f)\in\mathbb{N}$. Combined with the above that leads to a contradiction.

In the second case, we compute

$$\mathrm{Var}(Y_i^2\mathbb{1}_{\{|Y_i|\leq C\}})=\mathbb{E}\left[Y_i^2\mathbb{1}_{\{|Y_i|\leq C\}}\right]\leq\mathbb{E}\left[Y_i^2\right]=2\mathbb{E}\left[E_i^2\right]-2\mathbb{E}\left[E_i\right]^2=2\mathrm{Var}(E_i)=\frac{2}{f(i-1)^2}.$$

By a similar reasoning as in the first case, the combination of

$$\sum_{i=1}^\infty\mathrm{Var}(Y_i^2\mathbb{1}_{\{|Y_i|\leq C\}})=\infty\quad\text{and}\quad\sum_{i=1}^\infty\frac{1}{f(i-1)^2}<\infty$$

thus also leads to a contradiction, which concludes the proof. □

For the proof of Lemma 4.8 we use the following result regarding martingales with bounded increments, which states that martingales with bounded increments either converge or oscillate between ∞ and $-\infty$. We leave the proof to Exercise 4.2.

Lemma 4.10 (Theorem 4.3.1 in [6]). *Let $(M_k)_{k\in\mathbb{N}}$ be a martingale and $C>0$ a constant such that $|M_{k+1}-M_k|\le C$ for all $k\in\mathbb{N}$ almost surely. Then,*

$$\mathbb{P}\big(\big\{\lim_{k\to\infty} M_k \text{ exists and is finite}\big\}\cup\big(\big\{\limsup_{k\to\infty} M_k=\infty\big\}\cap\big\{\liminf_{k\to\infty} M_k=-\infty\big\}\big)\big)=1.$$

Proof of Lemma 4.8. Let us write

$$\mathcal{E}_+ := \bigg\{\limsup_{k\to\infty}\sum_{i=1}^k Y_i=\infty\bigg\},\qquad \mathcal{E}_- := \bigg\{\liminf_{k\to\infty}\sum_{i=1}^k Y_i=-\infty\bigg\}.$$

As both $\mathcal{E}_+$ and $\mathcal{E}_-$ are tail events with respect to the sequence of random variables $(Y_i)_{i\in\mathbb{N}}$, it follows from Kolmogorov's 0–1 law that $\mathcal{E}_+$ and $\mathcal{E}_-$ occur with probability 0 or 1. Since the random variables Y_i are symmetric, we have $Y_i\overset{\mathrm{d}}{=}-Y_i$. Hence, as

$$\limsup_{k\to\infty}\sum_{i=1}^k Y_i=\infty\quad\Leftrightarrow\quad\liminf_{k\to\infty}\sum_{i=1}^k -Y_i=-\infty,$$

it follows that $\mathcal{E}_+\cap\mathcal{E}_-$ and $\mathcal{E}_+\cup\mathcal{E}_-$ coincide up to sets of measure zero. As $\sum_{i=1}^\infty Y_i$ diverges, it follows from Kolmogorov's three series theorem that for any $C>0$ one of the following three conditions holds:

$$\sum_{i=1}^\infty \mathbb{P}(|Y_i|\ge C)=\infty,\quad \sum_{i=1}^\infty \mathbb{E}\left[Y_i\mathbb{1}_{\{|Y_i|\le C\}}\right]=\infty,\quad \sum_{i=1}^\infty \mathrm{Var}(Y_i\mathbb{1}_{\{|Y_i|\le C\}})=\infty. \tag{4.8}$$

By the symmetry of the Y_i, the second condition does not hold, so we can restrict ourselves to the first and third.

Suppose the first condition holds. Then, $|Y_i|\ge C$ for infinitely many i. By taking a countable union over all $C\in\mathbb{N}$, we thus obtain

$$\mathbb{P}\Bigg(\bigcap_{C\in\mathbb{N}}\{|Y_i|\ge C \text{ for infinitely many } i\}\Bigg)=1. \tag{4.9}$$

Now, fix $a,b\in\mathbb{Z}$. We then have

$$\mathbb{P}\Bigg(\limsup_{k\to\infty}\sum_{i=1}^k Y_i\le a,\liminf_{k\to\infty}\sum_{i=1}^k Y_i\ge b\Bigg)=0. \tag{4.10}$$

We argue by contradiction and assume that the event in the probability has positive probability. Then, with positive probability there exists $K\in\mathbb{N}$ such that

$$\sum_{i=1}^k Y_i\le a+1\qquad\text{and}\qquad\sum_{i=1}^k Y_i\ge b-1\qquad\text{for all } k\ge K.$$

However, using (4.9) with $C=|a|+|b|+2$, it follows that $|Y_i|\ge|a|+|b|+2$ for infinitely many i, which yields a contradiction. Hence, by a union bound and using (4.10),

$$\begin{aligned}\mathbb{P}\big(\mathcal{E}_+^c\cap\mathcal{E}_-^c\big)&=\mathbb{P}\Bigg(\bigcup_{a,b\in\mathbb{Z}}\bigg\{\limsup_{k\to\infty}\sum_{i=1}^k Y_i<a,\liminf_{k\to\infty}\sum_{i=1}^k Y_i>b\bigg\}\Bigg)\\&\le\sum_{a,b\in\mathbb{Z}}\mathbb{P}\Bigg(\limsup_{k\to\infty}\sum_{i=1}^k Y_i<a,\liminf_{k\to\infty}\sum_{i=1}^k Y_i>b\Bigg)=0.\end{aligned}$$

So, $\mathbb{P}(\mathcal{E}_+\cup\mathcal{E}_-)=1$ and thus $\mathbb{P}(\mathcal{E}_+\cap\mathcal{E}_-)=1$.

Now, suppose that the third condition in (4.8) holds for any $C > 0$ but the first condition is not met for some $C > 0$. So, there exists $C > 0$ such that

$$\sum_{i=1}^{\infty} \mathbb{P}(|Y_i| \geq C) < \infty \qquad \text{and} \qquad \sum_{i=1}^{\infty} \mathrm{Var}(Y_i \mathbb{1}_{\{|Y_i| \leq C\}} = \infty.$$

As the Y_i are independent, the reverse Borel-Cantelli lemma yields that $|Y_i| \geq C$ for only finite many indices i almost surely. Since $|Y_i| < \infty$ for all $i \in \mathbb{N}$ almost surely, it follows that $\mathbb{P}(\mathcal{E}_+ \cap \mathcal{E}_-) = 1$ is equivalent to

$$\mathbb{P}\bigg(\limsup_{k\to\infty} \sum_{i=1}^{k} Y_i \mathbb{1}_{\{|Y_i| \leq C\}} = \infty, \liminf_{k\to\infty} \sum_{i=1}^{k} Y_i \mathbb{1}_{\{|Y_i| \leq C\}} = -\infty\bigg) = 1.$$

Note that

$$M_k \coloneqq \sum_{i=1}^{k} Y_i \mathbb{1}_{\{|Y_i| \leq C\}}$$

is a martingale that satisfies $|M_{k+1} - M_k| \leq C$ for all $k \in \mathbb{N}$. Hence, by Lemma 4.10,

$$\mathbb{P}\Big(\Big\{\lim_{n\to\infty} M_n \text{ exists and is finite}\Big\} \cup (\mathcal{E}_+ \cap \mathcal{E}_-)\Big) = 1.$$

As both the event $\{\lim_{n\to\infty} M_n \text{ exists and is finite}\}$ and $\mathcal{E}_+ \cap \mathcal{E}_-$ occur with either probability 0 or 1, it suffices to show that the former has probability 0. We thus show that

$$\sum_{k=1}^{\infty} Y_i \mathbb{1}_{\{|Y_i| \leq C\}}$$

does not converge almost surely. This follows directly from applying Kolmogorov's three series theorem to the random variables $Z_i \coloneqq Y_i \mathbb{1}_{\{|Y_i| \leq C\}}$ for any $C > 0$, since

$$\sum_{i=1}^{\infty} \mathrm{Var}(Z_i \mathbb{1}_{\{|Z_i| \leq C\}}) = \sum_{i=1}^{\infty} \mathrm{Var}(Y_i \mathbb{1}_{\{|Y_i| \leq C\}}) = \infty,$$

so that we conclude the proof. □

4.3. **The (non)-existence of hubs.** With the preliminary results in Section 4.2 at hand, we now turn to proving Theorem 4.1.

Recall the Ulam-Harris tree UH from (4.1). We say that an individual $u = u_1 \cdots u_k \in \mathrm{UH}$ is K-moderate when $u_i \leq K$ for all $i \in [k]$. Further, recall that $\deg(v, T)$ denotes the number of children v has in the tree $T \subseteq \mathrm{UH}$. If $v \notin V(T)$, then we set $\deg(v, T) = -\infty$. We now define the event

$$\mathrm{Win}(u, v) \coloneqq \{\exists t_0 \geq 0 : \deg(u, \mathcal{T}_t) \geq \deg(v, \mathcal{T}_t) \text{ for all } t \geq t_0\}.$$

The following result provides sufficient conditions under which $\mathrm{Win}(u, v)$ occurs almost surely for any $u, v \in \mathrm{UH}$.

Proposition 4.11 (Lemma 3.6 in [11])**.** *Recall the random variables $(E_i)_{i\in\mathbb{N}}$, $(E'_i)_{i\in\mathbb{N}}$, and $(Y_i)_{i\in\mathbb{N}}$ from (4.6). The following two claims hold.*

(a) If $\sum_{i=1}^{\infty} Y_i$ converges almost surely and f satisfies (LS-f), then for any $u, v \in \mathrm{UH}$, we have

$$\mathbb{P}(\mathrm{Win}(u, v) \cup \mathrm{Win}(v, u)) = 1.$$

(b) If $\sum_{i=1}^{\infty} Y_i$ diverges almost surely, then there exists a function $\phi \colon \mathbb{N} \to \mathbb{N}$ such that $\phi(j) > j$ and

$$\mathbb{P}\bigg(\exists k \in \{j+1, \ldots, \phi(j)\} : \sum_{i=1}^{k} E_i \leq \sum_{i=j+1}^{k} E'_i\bigg) \geq \frac{1}{2} \qquad \text{for all } j \in \mathbb{N}.$$

Before we prove the proposition, let us discuss the intuition behind the results. (a) states that for any two individuals u and v, eventually one will have a larger degree than the other. As such, for any finite set S of individuals, we know that there exists a unique winner $u^* \in S$ that eventually attains the largest degree amongst all $u \in S$. Our branching process, however, contains a growing number of individuals, so that this result is not yet sufficient. The goal will be to construct a set S in a smart way such that we can prove that it is finite and that it contains the individual with the largest degree among all individuals, from which it follows that a persistent hub exists.

The result in (b) should be understood as follows: Let $(E_i')_{i\in\mathbb{N}}$ correspond to the inter-birth times of an individual $u \in \mathrm{UH}$, and let the $(E_i)_{i\in\mathbb{N}}$ correspond to the inter-birth times of the individual uj, which is the j^{th} child of u. Then, the result states that the probability that the individual uj produces $k \geq j+1$ children before u produces k children is at least $1/2$. That is, the probability that a child catches up to its parent, despite being born much later, is bounded away from zero. This suggests that 'many' children of u should catch up to u, so that u cannot attain the largest degree for all but finite time. As the choice of u is arbitrary, this suggests that no individual should attain the largest degree for for all but finite time, so that no persistent hub exists.

We make both of these claims precise in the remainder of this section, after proving Proposition 4.11.

Proof. (a) Recall that $\mathcal{B}(u)$ and $\mathcal{B}(v)$ denote the birth-times of u and v, respectively. Also, we recall $S_k^{(u)}$ and $S_k^{(v)}$ from (4.2) to denote the time u and v need to produce k children, respectively, for $k \in \mathbb{N}$. We then write, using (4.3), for $k \in \mathbb{N}$,

$$\mathcal{B}(uk) - \mathcal{B}(uv) = (\mathcal{B}(u) - \mathcal{B}(v)) + \sum_{i=1}^{k} (E_i^{(u)} - E_i^{(v)}).$$

It is clear that both $\mathcal{B}(u)$ and $\mathcal{B}(v)$ are finite almost surely. Furthermore, we note that $E_i^{(u)} - E_i^{(v)} \overset{\mathrm{d}}{=} Y_i$. Since $\sum_{i=1}^{\infty} Y_i$ converges almost surely, it thus follows that there exists $K_0 \in \mathbb{N}$ such that almost surely

$$\text{either} \quad \mathcal{B}(uk) - \mathcal{B}(uk) < 0 \quad \text{for all } k \geq K_0, \qquad \text{or} \qquad \mathcal{B}(uk) - \mathcal{B}(uk) > 0 \quad \text{for all } k \geq K_0.$$

By the condition in (LS-f) we can apply Lemma 4.5, so that $|\mathcal{T}_t| < \infty$ for all $t < \infty$ almost surely. This ensures that u and v both produce an infinite offspring as $t \to \infty$. In the first case, it now follows that the k^{th} child of u is born before the k^{th} child of v for all $k \geq K_0$, so that by the time the K_0^{th} child of u is born, its degree is always at least the degree of v. So, $\mathrm{Win}(u, v)$ occurs almost surely with $t_0 = \mathcal{B}(uK_0)$. In the second case, the event $\mathrm{Win}(v, u)$ occurs almost surely with $t_0 = \mathcal{B}(vK_0)$.

(b) As $\sum_{i=1}^{\infty} Y_i$ diverges, it follows from Lemma 4.8 that $\limsup_{k\to\infty} \sum_{i=1}^{k} Y_i = \infty$ almost surely. As

$$\left| \sum_{i=1}^{j} Y_i \right| < \infty$$

for any $j \in \mathbb{N}$ almost surely, we thus also have

$$\limsup_{k\to\infty} \sum_{i=j+1}^{k} Y_i = \infty.$$

Since $E_i' - E_i \overset{\mathrm{d}}{=} Y_i$ and $\sum_{i=1}^{j} E_i < \infty$ for any $j \in \mathbb{N}$ almost surely, we arrive at

$$\mathbb{P}\bigg(\exists k > j : \sum_{i=1}^{j} E_i \leq \sum_{i=j+1}^{k} (E_i' - E_i) \bigg) = 1.$$

By rearranging terms in the inequality and using the monotone convergence theorem, we obtain

$$\lim_{n\to\infty} \mathbb{P}\bigg(\exists k \in \{j+1, \ldots, n\} : \sum_{i=1}^{k} E_i \leq \sum_{i=j+1}^{k} E_i' \bigg) = 1.$$

We can now choose $\phi(j)$ as

$$\phi(j) := \inf\bigg\{n > j : \mathbb{P}\bigg(\exists k \in \{j+1,\dots,n\} : \sum_{i=1}^{k} E_i \le \sum_{i=j+1}^{k} E_i'\bigg) \ge \frac{1}{2}\bigg\},$$

for $j \in \mathbb{N}$, which concludes the proof. $\square$

With Proposition 4.11 at hand, we can now prove the 'only if' direction of Theorem 4.1.

Proof of Theorem 4.1 assuming (S) *does not hold.* As we assume (S) is not satisfied, we conclude from Lemma 4.7 that $\sum_{i=1}^{\infty} Y_i$ diverges almost surely, and thus we can apply Lemma 4.8 and part (b) of Proposition 4.11.

We prove the result by contradiction. Fix $u \in \mathrm{UH}$ and suppose that u is a persistent hub with positive probability. That is,

$$\mathbb{P}(\exists t_0 > 0\ \forall t \ge t_0\ \forall v \in \mathrm{UH}\colon\ \deg(u, \mathcal{T}_t) \ge \deg(v, \mathcal{T}_t)) > 0.$$

In particular, this implies that u eventually has a larger degree than all of its children with positive probability. As $\mathcal{T}_t$ is almost surely finite for all $t > 0$ by Lemma 4.5, we thus have

$$\{\exists t_0 > 0\ \forall t \ge t_0\ \forall v \in \mathrm{UH}\colon\ \deg(u, \mathcal{T}_t) \ge \deg(v, \mathcal{T}_t)\} \subseteq \bigcup_{K\in\mathbb{N}} \{\forall j \in \mathbb{N}\ \forall k \ge K : \mathcal{B}(uk) \le \mathcal{B}(ujk)\}.$$

The event on the right-hand side thus also occurs with positive probability. Note that, for $j < k$,

$$\mathcal{B}(uk) \le \mathcal{B}(ujk) \quad \Leftrightarrow \quad \mathcal{B}(uj) + \sum_{i=j+1}^{k} E_i^{(u)} \le \mathcal{B}(uj) + \sum_{i=1}^{k} E_i^{(uj)} \quad \Leftrightarrow \quad \sum_{i=j+1}^{k} E_i^{(u)} \le \sum_{i=1}^{k} E_i^{(uj)}.$$

Now, we use Proposition 4.11(b) and recall the function ϕ. We define the subsequence $j_1 = K + 1$ and $j_{\ell+1} = \phi(j_\ell)$ for $\ell \in \mathbb{N}$ and the events

$$\mathcal{E}_\ell(u) := \{\exists k \in \{j_\ell + 1, \dots, \phi(j_\ell)\} \colon \mathcal{B}(uj_\ell k) \le \mathcal{B}(uk)\} \qquad \text{for } \ell \in \mathbb{N}.$$

We can rewrite these events as

$$\mathcal{E}_\ell(u) = \bigg\{\exists k \in \{j_\ell + 1, \dots, \phi(j_\ell)\} \colon \sum_{i=1}^{k} E_i^{(uj_\ell)} \le \sum_{i=j_\ell+1}^{k} E_i^{(u)}\bigg\}.$$

Note that, since $\phi(j) > j$ and thus $j_{\ell+1} = \phi(j_\ell) > j_\ell$, the events $(\mathcal{E}_\ell(u))_{\ell\in\mathbb{N}}$ are mutually exclusive. Moreover, we have $\mathbb{P}(\mathcal{E}_\ell(u)) \ge 1/2$ for any $\ell \in \mathbb{N}$. The reverse Borel Cantelli lemma thus yields that infinitely many events $\mathcal{E}_\ell(u)$ occur almost surely. Since $j_{\ell+1} = \phi(j_\ell) > j_\ell$ is increasing in ℓ, this contradicts the fact that the event

$$\bigcup_{K\in\mathbb{N}} \{\forall j \in \mathbb{N}\ \forall k \ge K : \mathcal{B}(uk) \le \mathcal{B}(ujk)\}$$

occurs with positive probability, so that

$$\mathbb{P}(u \text{ is a persistent hub}) = 0.$$

A union bound over all $u \in \mathrm{UH}$ thus yields

$$\mathbb{P}(\text{There exists a persistent hub}) \le \sum_{u\in\mathrm{UH}} \mathbb{P}(u \text{ is a persistent hub}) = 0,$$

which concludes the proof. $\square$

We now are going to prove the 'if' direction of Theorem 4.1. First, we state the following general result.

Lemma 4.12 (Lemma 3.4 in [11])**.** *Let Z, $(X_i)_{i\in\mathbb{N}}$, and $(X_i')_{i\in\mathbb{N}}$ be mutually independent random variables, where $(X_i)_{i\in\mathbb{N}} \overset{\mathrm{d}}{=} (X_i')_{i\in\mathbb{N}}$. Suppose there exists $\lambda > 0$ and $k \in \mathbb{N}$ such that*

$$\prod_{i=k}^{\infty} \mathbb{E}\left[\mathrm{e}^{\lambda(X_i'-X_i)}\right] < \infty. \tag{4.11}$$

Then,

$$\mathbb{P}\Big(\exists j \in \mathbb{N}_0 : Z + \sum_{i=1}^{k+j} X_i \le \sum_{i=k}^{k+j} X_i'\Big) \le \Big(\prod_{i=k}^{\infty} \mathbb{E}\Big[\mathrm{e}^{\lambda(X_i'-X_i)}\Big]\Big)\mathbb{E}\Big[\exp\Big(-\lambda\Big(Z + \sum_{i=1}^{k-1} X_i\Big)\Big)\Big].$$

Remark 4.13. Let us set $X_i = E_i$ and $X_i' = E_i'$ to be independent exponential random variables with rate $f(i-1)$, and let us assume that f satisfies (S). Then, for any $\lambda > 0$ we can choose K large enough such that $f(i) > \lambda$ for all $i \ge K-1$ (notice that this is possible since f diverges to infinity by (S)). It then follows from (4.7) that (4.11) holds. ◀

Proof. First, note that

$$\mathbb{P}\Big(\exists j \in \mathbb{N}_0 : Z + \sum_{i=1}^{k+j} X_i \le \sum_{i=k}^{k+j} X_i'\Big) = \lim_{n\to\infty} \mathbb{P}\Big(\exists j \in \{0,\dots,n\} : Z + \sum_{i=1}^{k+j} X_i \le \sum_{i=k}^{k+j} X_i'\Big).$$

We then define

$$M_j := \exp\Big(\lambda \sum_{i=k}^{k+j} X_i' - \lambda \sum_{i=1}^{k+j} X_i - \lambda Z\Big) \qquad \text{for } j \in \mathbb{N}_0.$$

By Jensen's inequality, it follows that M_j is a submartingale. We then rewrite

$$\mathbb{P}\Big(\exists j \in \{0,\dots,n\} : Z + \sum_{i=1}^{k+j} X_i \le \sum_{i=k}^{k+j} X_i'\Big) = \mathbb{P}(\exists j \in \{0,\dots,n\} : M_j \ge 1) = \mathbb{P}\big(\max_{j\in\{0,\dots,n\}} M_j \ge 1\big).$$

By applying Doob's maximal inequality for submartingale to the right-hand side, we can bound it from above by $\mathbb{E}[M_n]$ and thus obtain

$$\mathbb{P}\Big(\exists j \in \{0,\dots,n\} : Z + \sum_{i=1}^{k+j} X_i \le \sum_{i=k}^{k+j} X_i'\Big) \le \Big(\prod_{i=k}^{\infty} \mathbb{E}\Big[\mathrm{e}^{\lambda(X_i'-X_i)}\Big]\Big)\mathbb{E}\Big[\exp\Big(-\lambda\Big(Z + \sum_{i=1}^{k-1} X_i\Big)\Big)\Big],$$

as desired. □

We define for $u, v \in \mathrm{UH}$ such that $\mathcal{B}(v) \ge \mathcal{B}(u)$ the event

$$\{v \text{ catches up to } u\} = \{\exists k \in \mathbb{N} \colon \mathcal{B}(vk) \le \mathcal{B}(uk)\}. \tag{4.12}$$

That is, an individual v that is born after u catches up to u if at some point v has at least as many children as u does (not including both having no children). We also define

$$\mathrm{Pre}(v) := \{u \in \mathrm{UH} \colon \mathcal{B}(u) \le \mathcal{B}(v)\} \qquad \text{for } v \in \mathrm{UH},$$

as the set of individuals born before (or at the same time as) v, also called its predecessors. Now, we finally define

$$\mathcal{P} := \{u \in \mathrm{UH} \colon u \text{ catches up to all } v \in \mathrm{Pre}(u)\}.$$

This set $\mathcal{P}$ is the set that we hinted at after Proposition 4.11. Indeed, it is clear that if some individual $u^* \in \mathrm{UH}$ were to be a persistent hub, then it has to catch up to all of its predecessors in $\mathrm{Pre}(u^*)$. So, if u^* is a persistent hub, then $u^* \in \mathcal{P}$. As discussed after Proposition 4.11, showing that the set $\mathcal{P}$ is finite almost surely will then allow us to conclude that a persistent hub exists. This is done is the following result, after which we prove the 'if' direction of Theorem 4.1.

Proposition 4.14 (Lemma 3.8 in [11])**.** *If Condition (S) is satisfied and f satisfies (LS-f), then $|\mathcal{P}|$ is finite almost surely.*

Proof of Theorem 4.1 assuming (S) *holds.* Since (S) is satisfied, we conclude from Lemma 4.7 that $\sum_{i=1}^\infty Y_i$ converges almost surely. Since f satisfies (LS-f) as well, we can thus apply part (a) from Proposition 4.11, so that for any $u, v \in \mathrm{UH}$, we have

$$\mathbb{P}(\mathrm{Win}(u,v) \cup \mathrm{Win}(v,u)) = 1.$$

As a result,

$$\mathbb{P}\bigg(\bigcap_{\substack{S\subset \mathrm{UH}\\ |S|<\infty}} \bigcap_{u,v\in S} \mathrm{Win}(u,v) \cup \mathrm{Win}(v,u)\bigg) = 1,$$

as the intersections are over countable sets. For any finite set $S \subset \mathrm{UH}$,

$$\bigcap_{u,v\in S} \mathrm{Win}(u,v)\cup\mathrm{Win}(v,u) = \{\exists w\in S\ \exists t_0>0\colon\ \deg(w,\mathcal{T}_t) = \max_{u\in S}\deg(u,\mathcal{T}_t) \text{ for all } t\geq t_0\}.$$

As a result, choosing $S = \mathcal{P}$ and using that $\mathcal{P}$ is almost surely finite by Proposition 4.14 yields

$$\mathbb{P}\bigg(\exists w\in\mathcal{P}\ \exists t_0>0\colon\ \deg(w,\mathcal{T}_t) = \max_{u\in\mathcal{P}}\deg(u,\mathcal{T}_t) \text{ for all } t\geq t_0\bigg) = 1.$$

As a persistent hub eventually attains the largest degree in $\mathcal{T}_t$ forever onwards, it must be in $\mathcal{P}$, so that we conclude that a persistent hub exists almost surely. □

It remains to prove Proposition 4.14, for which we use the following lemma.

Lemma 4.15 (Proposition 3.9 in [11])**.** *Suppose that f satisfies* (S) *and* (LS-f)*. Then, there exists $K \in \mathbb{N}$ such that for any $u \in \mathrm{UH}$,*

$$\mathbb{E}\left[|\{v\in \mathrm{UH}\setminus\{\varnothing\}\colon v_1>K, uv \textit{ catches up to } u\}|\right] < \infty.$$

Proof. Fix $\lambda > 0$ as in Lemma 4.5 and $K = K(\lambda) \in \mathbb{N}$ large enough, as in Remark 4.13, so that

$$\prod_{i=K}^\infty \mathbb{E}\left[\mathrm{e}^{\lambda(E_i - E_i')}\right] < \infty.$$

We first write

$$\begin{aligned}
&\mathbb{E}[|\{v\in\mathrm{UH}\setminus\{\varnothing\}\colon v_1>K, uv \text{ catches up to } u\}|]\\
&\quad= \sum_{\substack{v\in\mathrm{UH}\setminus\{\varnothing\}\\ v_1>K}} \mathbb{P}(uv \text{ catches up to } v)\\
&\quad= \sum_{m=1}^\infty \sum_{v_1=K+1}^\infty \sum_{v_2=1}^\infty \cdots \sum_{v_m=1}^\infty \mathbb{P}(uv_1\cdots v_m \text{ catches up to } u)\,.
\end{aligned} \tag{4.13}$$

Our aim is hence to show that these probabilities decays sufficiently fast as we increase m and $v_1, \ldots, v_m$. By the definition of the event that $uv_1\cdots v_m$ catches up to u in (4.12), we have

$$\begin{aligned}
\{uv_1\cdots v_m \text{ catches up to } u\} &= \{\exists k\in\mathbb{N}\colon \mathcal{B}(uv_1\cdots v_m k)\leq \mathcal{B}(uk)\}\\
&= \bigg\{\exists k>v_1\colon\ \sum_{\ell=1}^{m-1}\sum_{i=1}^{v_{\ell+1}} E_i^{(uv_1\cdots v_\ell)} + \sum_{i=1}^k E_i^{(uv)} \leq \sum_{i=v_1+1}^k E_i^{(u)}\bigg\}.
\end{aligned}$$

To bound the probability of this event, we apply Lemma 4.12 with

$$Z = \sum_{\ell=1}^{m-1}\sum_{i=1}^{v_{\ell+1}} E_i^{(uv_1\cdots v_\ell)}, \quad X_i = E_i^{(uv)}, \quad X_i' = E_i^{(u)}, \quad \text{and} \quad k = v_1,$$

and choose λ as in Lemma 4.6. Since $v_1 > K$ This yields the upper bound

$$\begin{aligned}
&\mathbb{P}(uv_1\cdots v_m \text{ catches up to } u)\\
&\leq \bigg(\prod_{i=v_1+1}^\infty \mathbb{E}\left[\mathrm{e}^{\lambda(E_i^{(u)} - E_i^{(uv)})}\right]\bigg)\mathbb{E}\bigg[\exp\bigg(-\lambda\bigg(\sum_{\ell=1}^{m-1}\sum_{i=1}^{v_{\ell+1}} E_i^{(uv_1\cdots v_\ell)} + \sum_{i=1}^{v_1-1} E_i^{(uv)}\bigg)\bigg)\bigg].
\end{aligned}$$

Since $v_1 \geq K$, we can further bound this from above by replacing v_1 by K in the product, since each term in the product is at least one by Jensen's inequality. We now substitute this upper bound in (4.13) to arrive at

$$\begin{aligned}
&\mathbb{E}[|\{v\in \mathrm{UH}\setminus\{\varnothing\}\colon v_1>K, uv \text{ catches up to } u\}|]\\
&\quad\leq \prod_{i=K}^{\infty}\mathbb{E}\Big[\mathrm{e}^{\lambda(E_i-E_i')}\Big]\sum_{m=1}^{\infty}\sum_{v_1=K+1}^{\infty}\sum_{v_2=1}^{\infty}\cdots\sum_{v_m=1}^{\infty}\mathbb{E}\bigg[\exp\bigg(-\lambda\Big(\sum_{\ell=1}^{m-1}\sum_{i=1}^{v_{\ell+1}}E_i^{(uv_1\cdots v_\ell)}+\sum_{i=1}^{v_1-1}E_i\Big)\bigg)\bigg],
\end{aligned}$$

where we changed some superscripts, since the expected values do not depend on the superscripts $(u),(uv)$, etc., as the random variables $(E_i^{(u)})_{i\in\mathbb{N},u\in\mathrm{UH}}$ are mutually independent. Using this independence also yields

$$\begin{aligned}
&\sum_{v_1=K+1}^{\infty}\sum_{v_2=1}^{\infty}\cdots\sum_{v_m=1}^{\infty}\mathbb{E}\bigg[\exp\bigg(-\lambda\Big(\sum_{\ell=1}^{m-1}\sum_{i=1}^{v_{\ell+1}}E_i^{(uv_1\cdots v_\ell)}+\sum_{i=1}^{v_1-1}E_i\Big)\bigg)\bigg]\\
&\qquad=\sum_{v_1=K+1}^{\infty}\sum_{v_2=1}^{\infty}\cdots\sum_{v_m=1}^{\infty}\mathbb{E}\bigg[\exp\Big(-\lambda\sum_{i=1}^{v_1-1}E_i\Big)\bigg]\prod_{\ell=1}^{m-1}\bigg(\mathbb{E}\bigg[\exp\Big(-\lambda\sum_{i=1}^{v_{\ell+1}}E_i^{(uv_1\cdots v_\ell)}\Big)\bigg]\bigg)\\
&\qquad=\sum_{v_1=K+1}^{\infty}\mathbb{E}\bigg[\exp\Big(-\lambda\sum_{i=1}^{v_1-1}E_i\Big)\bigg]\prod_{\ell=2}^{m}\bigg(\sum_{v_\ell=1}^{\infty}\mathbb{E}\bigg[\exp\Big(-\lambda\sum_{i=1}^{v_\ell}E_i\Big)\bigg]\bigg)\\
&\qquad=\sum_{v_1=K+1}^{\infty}\mathbb{E}\bigg[\exp\Big(-\lambda\sum_{i=1}^{v_1-1}E_i\Big)\bigg]\bigg(\sum_{v=1}^{\infty}\mathbb{E}\bigg[\exp\Big(-\lambda\sum_{i=1}^{v}E_i\Big)\bigg]\bigg)^{m-1}.
\end{aligned}$$

By our choice of λ (as in Lemma 4.6), the sum in brackets is equals a constant $c\in(0,1)$, and the first sum on the right-hand side is at most c. We thus obtain an upper bound c^m. Finally, we arrive at

$$\mathbb{E}[|\{v\in \mathrm{UH}\setminus\{\varnothing\}\colon v_1>K, uv \text{ catches up to } u\}|]\leq \prod_{i=K}^{\infty}\mathbb{E}\Big[\mathrm{e}^{\lambda(E_i-E_i')}\Big]\sum_{m=1}^{\infty}c^m,$$

which is finite since $c<1$ and by the choice of K, as desired. □

We have thus seen that there are finitely many descendants uv of an individual u such that $v_1>K$ that manage to catch up to u almost surely when (S) is satisfied. In the proof of Proposition 4.14 we use and further extend this result to show that the set of all individuals that catch up to all ancestors is finite almost surely as well.

Proof of Proposition 4.14. Fix $K\in\mathbb{N}$, which is to be determined later. We then define

$$\mathrm{UH}_{\leq K}\coloneqq\{u\in\mathrm{UH}\colon u_i\leq K \text{ for all } i\in[|u|]\}.$$

In other words, $\mathrm{UH}_{\leq K}$ is the infinite K-ary tree. We call individuals $u\in\mathrm{UH}_{\leq K}$ K-moderate. Also, let $T\coloneqq\mathcal{B}(K)$ be the time at which the root individual $\varnothing$ has produced K children.

As $\mathcal{B}(K)<\infty$ almost surely, it follows from Lemma 4.5 that $|\mathcal{T}_T|<\infty$ almost surely. Now, let M be the maximum length of any K-moderate individual born in $\mathcal{T}_T$. That is,

$$M\coloneqq\max\{|u|\colon u\in\mathcal{T}_T\cap\mathrm{UH}_{\leq K}\}.$$

We define the sets

$$A_+\coloneqq\{u\in\mathrm{UH}_{\leq K}\colon |u|=M+1\}\qquad\text{and}\qquad A_-\coloneqq\{u\in\mathrm{UH}_{\leq K}\colon |u|\leq M\}.$$

Observe that both sets are $\mathcal{F}_T$ measurable, with $\mathcal{F}_T$ the σ-algebra generated by $(\mathcal{T}_t)_{t\leq T}$. We further define, for $u\in\mathrm{UH}$,

$$\mathcal{C}_u\coloneqq\{uv\colon v\in\mathrm{UH}, v_1>K, \text{ and } uv \text{ catches up to } u\},$$

and

$$\mathcal{D}_u\coloneqq\{uv\colon v\in\mathrm{UH}\setminus\{\varnothing\} \text{ and } uv \text{ catches up to } \varnothing\}.$$

Now, we claim that

$$\mathcal{P} \subseteq \mathcal{C}_\varnothing \cup A_- \cup A_+ \cup \Big(\bigcup_{u\in A_-} \mathcal{C}_u \Big) \cup \Big(\bigcup_{u\in A_+} \mathcal{D}_u \Big). \tag{4.14}$$

This is true, since

- $\mathcal{C}_\varnothing$ contains all individuals u with u_1 that catch up to $\varnothing$.
- $A_- \cup (\cup_{u\in A_-}\mathcal{C}_u)$ contains individuals that catch up to one of their predecessors in A_-.
- $A_+ \cup (\cup_{u\in A_+}\mathcal{D}_u)$ contains individuals that catch up to one of their predecessors in A_+.

As $\mathcal{P}$ contains the individuals that catch up to all of their predecessors, each such individual is contained in one of the three sets outlined above.

To show that $|\mathcal{P}|$ is finite almost surely, it thus suffices to show that the size of each of the sets in the union in (4.14) is finite.

It is directly clear that

$$|A_-| = \frac{K^{M+1}-1}{K-1} \qquad \text{and} \qquad |A_+| = K^{M+1}.$$

Again using Lemma 4.5, it follows that M is finite almost surely, so that $|A_-|$ and $|A_+|$ are finite almost surely. Also, by Lemma 4.15 we have that $|\mathcal{C}_u|$ is finite almost surely for any $u \in \mathrm{UH}$. It thus remains to prove that the $\mathcal{D}_u$ are finite sets for any $u \in A_+$ almost surely. Here, we use the fact that $u \in A_+$ and any of its descendants are not yet born in $\mathcal{T}_T$. We write using (4.12),

$$\begin{aligned}\{uv \text{ catches up to } \varnothing\} &= \{\exists k\in\mathbb{N}\colon \mathcal{B}(uvk) \leq \mathcal{B}(k)\}\\ &= \Big\{\exists k > K\colon \mathcal{B}(u) + \sum_{\ell=0}^{m-1}\sum_{i=1}^{v_{\ell+1}} E_i^{(uv_1\cdots v_\ell)} + \sum_{i=1}^k E_i^{(uv)} \leq \sum_{i=1}^k E_i^{(\varnothing)}\Big\}.\end{aligned}$$

We now notice that $\mathcal{B}(u) > \mathcal{B}(K)$ since $u \in A_+$. Hence, conditionally on $\mathcal{F}_T$ and for $u\in A_+$ we obtain the inclusion

$$\{uv \text{ catches up to } \varnothing\} \subseteq \Big\{\exists k > K\colon \sum_{\ell=0}^{m-1}\sum_{i=1}^{v_{\ell+1}} E_i^{(uv_1\cdots v_\ell)} + \sum_{i=1}^k E_i^{(uv)} \leq \sum_{i=K+1}^k E_i^{(\varnothing)}\Big\}.$$

Hence,

$$\mathbb{P}(uv \text{ catches up to } \varnothing \,|\, \mathcal{F}_T) \leq \mathbb{P}\bigg(\exists k > K\colon \sum_{\ell=0}^{m-1}\sum_{i=1}^{v_{\ell+1}} E_i^{(uv_1\cdots v_\ell)} + \sum_{i=1}^k E_i^{(uv)} \leq \sum_{i=K+1}^k E_i^{(\varnothing)} \,\bigg|\, \mathcal{F}_T\bigg).$$

By the definition of T, we observe that all random variables in the probability on the right-hand side are independent of $\mathcal{F}_T$. We can thus apply Lemma 4.12 with $\lambda > 0$ as in Lemma 4.5 and $K = K(\lambda) \in \mathbb{N}$ large enough as in Remark 4.13, to arrive at the upper bound

$$\mathbb{P}(uv \text{ catches up to } \varnothing \,|\, \mathcal{F}_T) \leq \prod_{i=K+1}^\infty \mathbb{E}\Big[\mathrm{e}^{\lambda(E_i - E_i')}\Big]\,\mathbb{E}\Big[\exp\Big(-\lambda\Big(\sum_{\ell=0}^{m-1}\sum_{i=1}^{v_{\ell+1}} E_i^{(uv_1\cdots v_\ell)} + \sum_{i=1}^K E_i^{(\varnothing)}\Big)\Big)\Big].$$

Using independence of the exponential random variables, we can write the right-hand side as

$$\prod_{i=K+1}^\infty \mathbb{E}\Big[\mathrm{e}^{\lambda(E_i - E_i')}\Big]\prod_{\ell=1}^m \mathbb{E}\Big[\exp\Big(-\lambda\sum_{i=1}^{v_\ell} E_i\Big)\Big]\,\mathbb{E}\Big[\exp\Big(-\lambda\sum_{i=1}^K E_i\Big)\Big].$$

By now summing over all $v \in \mathrm{UH}$, we arrive at

$$\begin{aligned}&\mathbb{E}[|\mathcal{D}_u| \,|\, \mathcal{F}_T]\\ &\leq \prod_{i=K+1}^\infty \mathbb{E}\Big[\mathrm{e}^{\lambda(E_i - E_i')}\Big]\,\mathbb{E}\Big[\exp\Big(-\lambda\sum_{i=1}^K E_i\Big)\Big]\sum_{m=1}^\infty\sum_{v_1=1}^\infty\cdots\sum_{v_m=1}^\infty\prod_{\ell=1}^m \mathbb{E}\Big[\exp\Big(-\lambda\sum_{i=1}^{v_\ell} E_i\Big)\Big]\\ &= \prod_{i=K+1}^\infty \mathbb{E}\Big[\mathrm{e}^{\lambda(E_i - E_i')}\Big]\,\mathbb{E}\Big[\exp\Big(-\lambda\sum_{i=1}^K E_i\Big)\Big]\sum_{m=1}^\infty\Big(\sum_{v=1}^\infty \mathbb{E}\Big[\exp\Big(-\lambda\sum_{i=1}^v E_i\Big)\Big]\Big)^m.\end{aligned}$$

We now choose λ as in (4.5), so that the sum of expected values is strictly smaller than 1, and so the sum over m is finite. Since A_+ is finite almost surely, we thus obtain that

$$\mathbb{P}\left(\left|\bigcup_{u\in A_+}\mathcal{D}_u\right|<\infty\,\middle|\,\mathcal{F}_T\right)=1\qquad\text{almost surely,}$$

and thus, by taking the expected value with respect to $\mathcal{F}_T$,

$$\mathbb{P}\left(\left|\bigcup_{u\in A_+}\mathcal{D}_u\right|<\infty\right)=1,$$

which yields the desired result and concludes the proof. □

4.4. **Exercises.**

Exercise 4.1. Prove Lemma 4.6. *Hint:* Prove first that there exists $\lambda'>0$ such that $\mu(\lambda')<\infty$, and use this to argue that there must then exist $\lambda\geq\lambda'$ such that $\mu(\lambda)<1$.

Exercise 4.2. Prove Lemma 4.10. *Hint:* We can assume without loss of generality that $X_0=0$. Now, for $K\in(0,\infty)$ define the stopping time

$$N=N(K):=\inf\{n\in\mathbb{N}\colon X_n\leq -K\}.$$

Show that we can apply Theorem 3.3 to $X_{\min\{n,N\}}+K+C$, so that $\lim_{n\to\infty}X_n$ exists on the event $\{N=\infty\}$. Argue that this idea, by letting $K\to\infty$, yields that $\lim_{n\to\infty}$ exists on the event $\{\liminf_{n\to\infty}X_n>-\infty\}$. Apply the same idea to $-X_n$ to show that $\lim_{n\to\infty}X_n$ exists on the event $\{\limsup_{n\to\infty}X_n<\infty\}$.

Exercise 4.3. The proofs of the main results in Section 4 did not rely on the exact distribution of the random variables E_i and Y_i was not important. All that mattered was whether or not $\sum_{i=1}^\infty Y_i$ converged of diverged almost surely. Can you think of other examples of inter-birth time distributions for which you can prove this convergence/divergence? Can you also prove the results in Lemmas 4.5, 4.6, and 4.7 for these examples?

Exercise 4.4. Suppose that f is such that

$$\sum_{i=0}^\infty\frac{1}{f(i)}<\infty.$$

Why can the premise of Lemma 4.5 not be true for such functions f?

4.5. **Further reading.**

There are various other works that study the (non-)existence of persistent hubs in preferential attachment models. We do not aim to provide a full overview here, but give some suggestions for interested students.

- Bäumler and Iyer [4] study, in a very recent paper, competition in a balls-in-bins process. This is related to preferential attachment models, and some of their results inform behaviour of preferential attachment models: When Condition (S) is not satisfied, for any $K\in\mathbb{N}$ the degrees $\deg(1,T_n),\deg(2,T_n),\dots,\deg(K,Tn)$ ordered by size appear in any permutation for infinitely many $n\in\mathbb{N}$ almost surely.
- Banerjee and Bhamidi [2] study the (non-)existence of persistent hubs in non-tree preferential attachment models. The analyses presented in these lecture notes is limited to the tree case only, as this allows for the embedding of the discrete tree process into the continuous-time branching process, as in Proposition 4.4. Banerjee and Bhamidi can, under some additional assumptions, also prove results for the non-tree case. Furthermore, in the case of non-existence of persistent hubs in the tree case, they are able to provide precise asymptotic results for the growth rate of I_n.

- Lodewijks [14] and Heydenreich and Lodewijks [7] study the non-existence of persistent hubs for a more general models known as *preferential attachment trees with vertex death*. In these models, at every step either a new vertex is added to the tree and connected to an alive vertex selected preferentially, or an existing alive vertex is selected preferentially and killed. The presence of killing vertices can change the behaviour of the model significantly and allow for much richer behaviour to be observed. This generalises some results and ideas of Banerjee and Bhamidi [2] as well as of Iyer [11], and presents new results that cannot be observed in preferential attachment trees without vertex death.
- When f grows faster than linear, results about persistent hubs can still be proved, but require different ideas and techniques. Both Iyer [11] and Lodewijks [13] prove the existence of persistent hubs in preferential attachment trees for cases where f grows super-linear.

References

[1] L. Addario-Berry and K. Ford. Poisson–Dirichlet branching random walks. 2013.

[2] S. Banerjee and S. Bhamidi. Persistence of hubs in growing random networks. *Probability Theory and Related Fields*, 180(3):891–953, 2021.

[3] A.-L. Barabási and R. Albert. Emergence of scaling in random networks. *science*, 286(5439):509–512, 1999.

[4] J. Bäumler and T. Iyer. Permutations in competing growth processes and balls-in-bins. *arXiv preprint arXiv:2510.07104*, 2025.

[5] S. Briend, F. Calvillo, and G. Lugosi. Archaeology of random recursive dags and Cooper-Frieze random networks. *Combinatorics, Probability and Computing*, 32(6):859–873, 2023.

[6] R. Durrett. *Probability: theory and examples*, volume 49. Cambridge university press, 1989.

[7] M. Heydenreich and B. Lodewijks. Preferential Attachment Trees with Vertex Death: Lack of Persistence of the Maximum Degree. *arXiv preprint arXiv:2503.08675*, 2025.

[8] R. van der Hofstad. *Random Graphs and Complex Networks. Volume 1*. Cambridge University Press, 2017.

[9] R. van der Hofstad. *Random Graphs and Complex Networks. Volume 2*. Cambridge University Press, 2024.

[10] T. Iyer. Fixation of leadership in non-Markovian growth processes. *arXiv preprint arXiv:2408.11516*, 2024.

[11] T. Iyer. Persistent hubs in CMJ branching processes with independent increments and preferential attachment trees. *arXiv preprint arXiv:2410.24170*, 2024.

[12] B. Lodewijks. On joint properties of vertices with a given degree or label in the random recursive tree. *Electronic Journal of Probability*, 27:1–45, 2022.

[13] B. Lodewijks. A star is born: Explosive Crump–Mode–Jagers branching processes in random environment. *arXiv preprint arXiv:2411.18749*, 2024.

[14] B. Lodewijks. Preferential Attachment Trees with Vertex Death: Persistence of the Maximum Degree. *arXiv preprint arXiv:2505.06187*, 2025.