

MASTER

Distances in power-law random graphs

Dommers, S.

Award date:
2009

[Link to publication](#)

Disclaimer

This document contains a student thesis (bachelor's or master's), as authored by a student at Eindhoven University of Technology. Student theses are made available in the TU/e repository upon obtaining the required degree. The grade received is not published on the document as presented in the repository. The required complexity or quality of research of student theses may vary by program, and the required minimum study period may vary in duration.

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain

EINDHOVEN UNIVERSITY OF TECHNOLOGY
DEPARTMENT OF MATHEMATICS AND COMPUTER SCIENCE

Distances in power-law random graphs

Sander Dommers

Supervisor: prof. dr. R.W. van der Hofstad

January 12, 2009

Abstract

In many real-world networks, such as the Internet and social networks, power-law degree sequences have been observed. This means that, when the graph is large, the proportion of vertices with degree k is asymptotically proportional to $k^{-\tau}$, for some $\tau \geq 1$. These networks are often *small worlds*, which means that distances in these networks are small. We will study two random graph models, the configuration model and the preferential attachment model, which will have power-law degree sequences when the number of vertices tends to infinity. An overview is given of known results about distances in these graph models. Also some new results will be presented, among which a log log lower bound on the diameter of preferential attachment graphs with $\tau > 2$.

Contents

1	Introduction	4
2	Random graph models	4
2.1	Configuration model	5
2.2	Preferential attachment models	5
3	Power-laws	6
3.1	Configuration model	6
3.2	Preferential attachment models	7
4	Distances and diameters	8
4.1	Configuration model	8
4.2	Preferential attachment models	10
5	A log log lower bound on the diameter in PA models	12
5.1	The first moment	13
5.2	The second moment	17
6	Late vertices have small degrees in PA models	19
7	Average distances for $\delta = 0$ in PA models	22
8	Distances for $\tau = 3$ in the configuration model	23
9	Conclusion	25

1 Introduction

There has been a lot of interest lately, in the topology of real-world networks, such as the Internet, social networks and biological networks. Empirical studies show that these networks have some remarkable similarities. The degree sequence of many of these networks obey a *power-law*. This means that, when the graph is large, the proportion of vertices with degree k is asymptotically proportional to $k^{-\tau}$, for some $\tau \geq 1$. Networks with this property are also called *scale-free* in the literature.

Power-laws have for example been observed in the topology of the Internet by Faloutsos, Faloutsos and Faloutsos in [12]. They show that the out-degree of domains in the Internet obeys a power-law with exponent $\tau \approx 2.2$ between the end of 1997 and 1998, although the network grew 45% in that time period.

An other property of these networks is that they seem to be *small worlds*, which means that distances in these networks are small compared to the size of the network. A famous result in this direction was from Milgram ([21]). He sent letters to random people in Nebraska and asked them to get the letter to a friend in Boston, but they were only allowed to sent it to people they know on first basis. Milgram discovered that the average number of steps necessary to get the letter back to his friend was about 6. Hence the famous phrase “six degrees of separation”. Many question marks can be placed about the interpretation of this experiment, but it created a huge interest in this phenomenon. For an extensive review of properties of complex networks, see [23] and the references therein.

Since these networks are often rather complex and large, various random graph models have been proposed to study them. The classical Erdős-Rényi random graph ([10]), does not suffice, because there power-law degree sequences can not be obtained. The two models that we study are the configuration model and the preferential attachment model.

The configuration model was originally proposed by Newman et al. in [24]. In this model, a fixed number of vertices all have an independent and identically distributed (i.i.d.) number of half-edges attached to it. Here the distribution will be a power-law distribution. These half-edges are then connected to each other uniformly at random to form a graph.

The preferential attachment model was introduced by Barabási and Albert in [3]. This model is based on two principles: the network grows continuously and new vertices are more likely to connect to a vertex that already has a large number of connections. This last phenomenon is also called the *rich-get-richer* effect. This model does not only give rise to a graph with a power-law degree sequence, but also explains *why* power-laws arise. See [2] for a popular account of preferential attachment and its consequences.

This rest of this thesis is structured as follows. We will first formally introduce the configuration model and the preferential attachment models we will study in Section 2. In Section 3, we will show that these models indeed give rise to power-law degree sequence, while in Section 4 we will give an overview of the known results on distances in these models. Sections 5 and further contain new results, starting with a log log lower bound on the diameter of preferential attachment graphs with $\tau \in (2, 3)$. A log log upper bound also exists. In Section 6 we will show that late vertices have small degrees, a result we need to prove this upper bound. In Section 7 we will generalize the log / log log upper and lower bound on diameters in preferential attachment models with $\tau = 3$, to bounds on average distances. Finally, we will show in Section 8, that average distances in the configuration model with $\tau = 3$ also have a log / log log lower bound. Section 9 will contain some concluding remarks and some ideas for future research.

2 Random graph models

In the next two sections, we will first introduce the random graph models we will study.

2.1 Configuration model

In the configuration model an undirected random graph with n vertices is constructed as follows. Let $\{D_i\}_{i=1}^n$ be a sequence of i.i.d. random variables with distribution D . Let vertex $i, i = 1, \dots, n$, be a vertex with D_i half-edges, also called *stubs*, attached to it, i.e. vertex i has degree D_i . Let $L_n = \sum_{i=1}^n D_i$ be the total degree, which we will assume to be even in order to be able to construct a graph. When L_n is odd we will increase the degree of D_n by 1. For n large, this will hardly change the results and we will therefore ignore this effect.

Now connect one of the half-edges uniformly at random to one of the remaining $L_n - 1$ half-edges. Repeat this procedure until all half-edges have been connected. We will denote the resulting graph by $\text{CM}_n(\{D_i\}_{i=1}^n)$.

Since we are interested in power-law random graphs, we will usually let D have a power-law distribution. This construction, however, also works for other distributions of D .

Note that the above construction will not necessarily result in a simple graph. Both self-loops and multiple edges may occur. Two ways to overcome this are to delete all loops and multiple edges, this model is also called the *erased* configuration model, and to perform the configuration model until it produces a simple graph, which is also called the *repeated* configuration model.

2.2 Preferential attachment models

In 1999 Barabási and Albert first introduced the preferential attachment model in [3] as follows:

... starting with a small number (m_0) of vertices, at every time step we add a new vertex with m ($\leq m_0$) edges that link the new vertex to m different vertices already present in the system. To incorporate preferential attachment, we assume that the probability Π that a new vertex will be connected to vertex i depends on the connectivity k_i of that vertex, so that $\Pi(k_i) = k_i / \sum_j k_j$. After t time steps, the model leads to a random network with $t + m_0$ vertices and mt edges.

This definition, however, is rather imprecise. For instance, it is not clear how the process starts, since at time $t = 0$, the degrees of the first m_0 vertices are 0, so the connecting probabilities are not properly defined. It also is not clear whether the m vertices should be added independently of each other, or if the connectivity coefficients should be updated after each edge that was added. In [5] Bollobás and Riordan gave a rigorous definition, starting with one vertex with m self-loops and specifying that the degrees should be updated in the process of attaching the m edges. We will study a generalization of this model, which we will call model (a), and two variants hereof; models (b) and (c). These three models are defined below, as proposed by Van der Hofstad and Hooghiemstra in [16]. In all these models a graph process $\{\text{PA}_{m,\delta}(t)\}$ is defined for $t \geq 1$ or $t \geq 2$. We will label the vertices of $\text{PA}_{m,\delta}(t)$ as $1, \dots, t$. Let $[t] = \{1, \dots, t\}$. The integer parameter $m \geq 1$ is the number of connections a new vertex that is added makes. At time t the random graphs have exactly t vertices and mt edges. The parameter $\delta \geq -m$ allows us to control the exponent of the power-law, as will be shown in Section 3.2.

- (a) Start with $\text{PA}_{m,\delta}^{(a)}(1)$ consisting of a single vertex with m self-loops. For $1 \leq j \leq m$, the end points of each of the m edges of vertex $t + 1$ are chosen with the following probabilities:

$$\mathbb{P} \left[j^{\text{th}} \text{ edge of } t + 1 \rightarrow i \mid \text{PA}_{m,\delta}^{(a)}(t, j - 1) \right] = \begin{cases} \frac{D_i(t, j-1) + \delta}{t(2m+\delta) + 2(j-1) + 1 + j\delta/m}, & \text{for } i \in [t], \\ \frac{D_i(t, j-1) + 1 + j\delta/m}{t(2m+\delta) + 2(j-1) + 1 + j\delta/m}, & \text{for } i = t + 1. \end{cases} \quad (1)$$

Here, $\text{PA}_{m,\delta}^{(a)}(t, j)$ is the graph and $D_i(t, j)$ is the degree of vertex i after the j -th edge of vertex $t + 1$ has been added. This model with $\delta = 0$ is the same as the model defined in [5].

- (b) For $m \geq 2$, start with $\text{PA}_{m,\delta}^{(b)}(1)$ consisting of a single vertex with m self-loops. For $m = 1$ let

$\text{PA}_{m,\delta}^{(b)}(1)$ undefined and let $\text{PA}_{m,\delta}^{(b)}(2)$ consist of vertices 1 and 2 joined by 2 edges.

$$\mathbb{P}\left[j^{\text{th}} \text{ edge of } t+1 \rightarrow i \mid \text{PA}_{m,\delta}^{(b)}(t, j-1)\right] = \begin{cases} \frac{D_i(t, j-1) + \delta}{t(2m+\delta) + 2(j-1) + (j-1)\delta/m}, & \text{for } i \in [t], \\ \frac{D_i(t, j-1) + (j-1)\delta/m}{t(2m+\delta) + 2(j-1) + (j-1)\delta/m}, & \text{for } i = t+1. \end{cases} \quad (2)$$

Again, $\text{PA}_{m,\delta}^{(b)}(t, j)$ is the graph and $D_i(t, j)$ is the degree of vertex i after the j -th edge of vertex $t+1$ has been added. The advantage of this model is that the graph $\text{PA}_{m,\delta}^{(b)}(t)$ is a connected random graph, while this is not necessarily the case in model (a). This is because the first edge that is added of a vertex t creates a self-loop with probability 0 in model (b).

(c) Let $\text{PA}_{m,\delta}^{(c)}(1)$ undefined and start with $\text{PA}_{m,\delta}^{(c)}(2)$, with the vertices 1 and 2 joined together by $2m$ edges. For $1 \leq j \leq m$, the end points of each of the m edges of vertex $t+1$ are chosen, conditionally on $\text{PA}_{m,\delta}^{(c)}(t)$, independently, with

$$\mathbb{P}\left[j^{\text{th}} \text{ edge of } t+1 \rightarrow i \mid \text{PA}_{m,\delta}^{(c)}(t)\right] = \frac{D_i(t) + \delta}{t(2m + \delta)}, \quad \text{for } i \in [t]. \quad (3)$$

Here, $D_i(t)$ is the degree of vertex i after vertex t and all its m edges have been added. This model will also result in a connected random graph $\text{PA}_{m,\delta}^{(c)}(t)$ and the resulting graph will not have any self-loops.

Often, models (a) and (b) are defined for $m = 1$ as above and the model for $m > 1$ is then derived by identifying, for $i \in [t]$, vertices $(i-1)m+1, \dots, im$ of $\text{PA}_{1,\delta'}(mt)$ to be vertex i in $\text{PA}_{m,\delta}(t)$, where $\delta' = \delta/m$. This will result in the same models as above. For model (c) this method is also possible, but then (3) has to be replaced with

$$\mathbb{P}\left[t+1 \rightarrow i \mid \text{PA}_{1,\delta'}^{(c)}(t)\right] = \frac{D_i(m\lfloor t/m \rfloor) + \delta'}{m\lfloor t/m \rfloor(2 + \delta')}, \quad \text{for } i \in [m\lfloor t/m \rfloor], \quad (4)$$

because the degrees are only updated after each m -th vertex that was added. We often use the explicit formulas (1-3), because it is easier to give bounds on the connecting probabilities.

In the literature many other preferential attachment models have been studied. For example, other functions of the degree can be taken to replace the numerators of (1-3). For an overview, see for example [4] and the references therein.

3 Power-laws

Before we investigate distances in the random graph models described in Sections 2.1 and 2.2, we will first show that these models result in power-law random graphs.

3.1 Configuration model

Define p_k^n , for $k = 1, 2, \dots$, by

$$p_k^n = \frac{1}{n} \sum_{i=1}^n \mathbb{I}\{D_i = k\}. \quad (5)$$

By the strong law of large numbers, we have that

$$p_k^n \xrightarrow{a.s.} \mathbb{P}[D = k], \quad (6)$$

so that the degree distribution of $\text{CM}_n(\{D_i\}_{i=1}^n)$ converges almost surely to the distribution of D . Thus, when D satisfies a power-law, so will the degree distribution.

In [7], Britton et al. study the degree distribution of the erased and the repeated configuration model. They show that, in the erased configuration model, the degree distribution converges to the distribution of D when D has finite mean. When D has a finite second moment, they show that this also holds for the repeated configuration model.

3.2 Preferential attachment models

Most papers about preferential attachment models study the degree sequence, and show that power-laws arise. Denote the number of vertices with degree k at time t as

$$N_k(t) = \sum_{i=1}^t \mathbb{I}\{D_i(t) = k\}, \quad (7)$$

and define $p_k(t) = N_k(t)/t$.

Most of these papers prove that $N_k(t)$ is concentrated around its mean by a martingale argument from Bollobás et al. ([6]). Then it is shown that the expected degree sequence obeys a power-law. We will give a heuristic here that can be found in [9] for a more general version of model (c).

Note that the expected number of vertices with degree k at time t satisfies

$$\mathbb{E}[N_k(t)|\text{PA}_{m,\delta}(t-1)] = N_k(t-1) + \mathbb{E}[N_k(t) - N_k(t-1)|\text{PA}_{m,\delta}(t-1)]. \quad (8)$$

When t gets large, it becomes unlikely that two or more edges from vertex t will attach to the same vertex, so let us assume that this does not happen. Then $N_k(t) - N_k(t-1)$ can only be unequal to zero if one of the following events happens:

- When one of the m edges of vertex t attaches to a vertex with degree $k-1$, the number of vertices with degree k increases by one. The probability that a fixed edge attaches to a vertex with degree $k-1$ can be derived from (3). Thus, the expected number of times this happens is, conditionally on $\text{PA}_{m,\delta}(t-1)$,

$$m \frac{(k-1+\delta)N_{k-1}(t-1)}{t(2m+\delta)}; \quad (9)$$

- When one of the m edges of vertex t attaches to a vertex with degree k , the number of vertices with degree k decreases by one. The expected number of times this happens is, conditionally on $\text{PA}_{m,\delta}(t-1)$,

$$m \frac{(k+\delta)N_k(t-1)}{t(2m+\delta)}; \quad (10)$$

- When $k=m$ the number of vertices with degree k increases by one.

We thus have that

$$\begin{aligned} & \mathbb{E}[N_k(t)|\text{PA}_{m,\delta}(t-1)] \\ & \approx N_k(t-1) + m \frac{(k-1+\delta)N_{k-1}(t-1)}{t(2m+\delta)} - m \frac{(k+\delta)N_k(t-1)}{t(2m+\delta)} + \mathbb{I}\{k=m\}, \end{aligned} \quad (11)$$

where there is an approximation sign because we ignore the possibility that two or more edges from vertex t attach to the same vertex. Taking expectations on both sides gives

$$\mathbb{E}[N_k(t)] \approx \mathbb{E}[N_k(t-1)] + \frac{(k-1+\delta)}{t\theta} \mathbb{E}[N_{k-1}(t-1)] - \frac{(k+\delta)}{t\theta} \mathbb{E}[N_k(t-1)] + \mathbb{I}\{k=m\}, \quad (12)$$

with $\theta = 2 + \frac{\delta}{m}$.

When we assume that $p_k(t)$ converges to some limit p_k as $t \rightarrow \infty$, we have that $\mathbb{E}[N_k(t)] - \mathbb{E}[N_k(t-1)] \rightarrow p_k$, so also that $\frac{1}{t} \mathbb{E}[N_k(t)] \rightarrow p_k$. We thus have the following recursion

$$p_k = \frac{(k-1+\delta)}{\theta} p_{k-1} - \frac{(k+\delta)}{\theta} p_k + \mathbb{I}\{k=m\}, \quad k \geq m. \quad (13)$$

By iteration we get that, for $k \geq m$,

$$p_k = \frac{\theta}{k + \delta + \theta} \prod_{j=1}^{k-m} \frac{k-j+\delta}{k-j+\delta+\theta} = \theta \frac{\Gamma(k+\delta)}{\Gamma(m+\delta)} \frac{\Gamma(m+\delta+\theta)}{\Gamma(k+1+\delta+\theta)}. \quad (14)$$

By Stirling's formula, $\Gamma(k+a)/\Gamma(k) \sim k^a$, so

$$p_k \sim ck^{-(\theta+1)} = ck^{-(3+\frac{\delta}{m})}, \quad (15)$$

for some $c > 0$. Thus, p_k obeys a power law with exponent $\tau = 3 + \frac{\delta}{m}$. By choosing $\delta > -m$, we can get any power law with exponent $\tau > 2$.

4 Distances and diameters

In the sequel, we will denote the event that a vertex i is an element of the vertex set of a graph G by $i \in G$. The distance between vertices i and j in the graph G will be denoted by $\text{dist}_G(i, j)$. We will define the diameter of a graph G as

$$\text{diam}(G) = \max_{i, j \in G} \{ \text{dist}_G(i, j) \mid \text{dist}_G(i, j) < \infty \}, \quad (16)$$

i.e., the largest distance between two connected vertices. A sequence of events $\{\mathcal{A}_t\}$ is said to hold *with high probability (whp)*, when $\mathbb{P}[\mathcal{A}_t] \rightarrow 1$ as $t \rightarrow \infty$.

4.1 Configuration model

Distances in the configuration model have been studied by Van der Hofstad et al. in a series of papers [17], [18] and [11]. In these papers they study the distance or hopcount H_n between vertices 1 and 2 in $\text{CM}_n(\{D_i\}_{i=1}^n)$, where D satisfies

$$\mathbb{P}[D > x] = x^{1-\tau} L(x), \quad x = 1, 2, \dots, \quad (17)$$

for $\tau > 3$, $\tau \in (2, 3)$ and $\tau \in [1, 2]$ respectively and where $L(x)$ is a slowly varying function, i.e., $\lim_{x \rightarrow \infty} \frac{L(cx)}{L(x)} = 1$ for all $c > 0$. Note that the distance between two vertices chosen uniformly at random, has the same distribution as H_n , because all vertices are exchangeable. If two vertices are not connected, then the distance between them is defined as ∞ .

In [17] the configuration model is studied with a degree distribution satisfying

$$\mathbb{P}[D \geq x] \leq cx^{1-\tau}, \quad x = 1, 2, \dots, \quad (18)$$

where c is a positive constant and $\tau > 3$. This covers all cases where $\mathbb{P}[D > x] = x^{1-\tau} L(x)$, for $\gamma > 3$ and $L(x)$ a slowly varying function, because of Potter's theorem ([13], Lemma 2, p. 277) which states that, for $x \rightarrow \infty$, $x^{-\varepsilon} < L(x) < x^\varepsilon$ for all $\varepsilon > 0$. The main result of [17] is the following.

Theorem 1. *Let D be a random variable satisfying (18). Let $\nu = \frac{\mathbb{E}[D(D-1)]}{\mathbb{E}[D]}$ and suppose $\nu > 1$. Then, conditionally on $H_n < \infty$, **whp**,*

$$(1 - \varepsilon) \frac{\log n}{\log \nu} \leq H_n \leq (1 + \varepsilon) \frac{\log n}{\log \nu}, \quad (19)$$

for all $\varepsilon > 0$.

Also the fluctuations of H_n around $\frac{\log n}{\log \nu}$ are determined. To prove this theorem, the neighborhood of vertices 1 and 2 is investigated. We start from vertex 1 which has D_1 stubs. Next, look at the vertices that connect to these stubs, which will be the vertices at distance 1. *Free* stubs are those stubs that do not connect to previously studied vertices. Continue in the same fashion. Let $Z_k^{(1)}$ be the number of free stubs of vertices at distance k from vertex 1. Do the same with vertex 2. As long as $Z_k^{(i)}$ is small compared to n , the probability that a new vertex with degree $j + 1$ connects to such a free stub is approximately equal to the fraction of stubs that belong to a vertex with degree $j + 1$. This equals:

$$\frac{1}{L_n} \sum_{l=1}^n (j+1) \mathbb{I}\{D = j+1\}. \quad (20)$$

Let $\mu = \mathbb{E}[D]$. Then, by the strong law of large numbers,

$$\frac{L_n}{n} \xrightarrow{a.s.} \mu \quad \text{and} \quad \frac{\sum_{l=1}^n \mathbb{I}\{D = j+1\}}{n} \xrightarrow{a.s.} \mathbb{P}[D = j+1], \quad (21)$$

so that (20) converges almost surely to

$$(j+1) \frac{\mathbb{P}[D_1 = j+1]}{\mu} \equiv g_j, \quad (22)$$

say. $Z_k^{(i)}$, $i = 1, 2$, thus behave like branching processes with offspring distribution D in the first generation and $\{g_j\}_{j=0}^\infty$, in all further generations. The condition $\nu > 1$ makes sure that this branching process is supercritical. The process $\{Z_k^{(i)}/\mu\nu^{k-1}\}$ is a martingale with uniformly bounded expectation, and thus the martingale convergence theorem tells us that there exist random variables $W^{(1)}$ and $W^{(2)}$, such that

$$\frac{Z_k^{(i)}}{\mu\nu^{k-1}} \xrightarrow{a.s.} W^{(i)}, \quad i = 1, 2. \quad (23)$$

It can be shown that the free stubs of vertices at distance $k - 1$ from vertex 1 are likely to connect to the free stubs of vertices at distance $l - 1$ from vertex 2 if $Z_k^{(1)}Z_l^{(2)}$ is of order n . Since $Z_k^{(1)}$ and $Z_l^{(2)}$ both grow at the same rate, both should be of order \sqrt{n} , so take $k = l = \frac{1}{2} \log_\nu n$. Thus, the distance between vertices 1 and 2 is approximately $k + l = \log_\nu n$. The random variables $W^{(1)}$ and $W^{(2)}$ can be used to describe the fluctuations of H_n around $\log_\nu n$.

In [18] it is shown that the above methodology can also be applied for $\tau \in (2, 3)$. The only difficulty is, that the corresponding branching process has an infinite mean offspring distribution. Under some extra conditions, Davies shows in [8] that there exist random variables $Y^{(1)}$ and $Y^{(2)}$ such that

$$(\tau - 2)^k \log \left(Z_k^{(i)} + 1 \right) \xrightarrow{a.s.} Y^{(i)}, \quad i = 1, 2. \quad (24)$$

These conditions hold, if there exist $\gamma \in [0, 1)$ and $C > 0$ such that

$$x^{-\tau+1-C(\log x)^{\gamma-1}} \leq \mathbb{P}[D > x] \leq x^{-\tau+1+C(\log x)^{\gamma-1}}, \quad \text{for large } x. \quad (25)$$

Because of this double exponential behavior, the following theorem can be proven.

Theorem 2. *Let D be a random variable satisfying (25) for $\tau \in (2, 3)$. Then, conditionally on $H_n < \infty$, **whp**,*

$$(1 - \varepsilon) \frac{2 \log \log n}{|\log(\tau - 2)|} \leq H_n \leq (1 + \varepsilon) \frac{2 \log \log n}{|\log(\tau - 2)|}, \quad (26)$$

for all $\varepsilon > 0$.

Interestingly, for $\tau > 2$ the diameter of $\text{CM}_n(\{D_i\}_{i=1}^n)$ is of order $\log n$, when $\mathbb{P}[D = 1] + \mathbb{P}[D_1 = 2] > 0$ and $\mathbb{P}[D = 1] < 1$, as is shown in [19]. This is because, **whp**, there are long chains of vertices with degree 2 in this case. When $\mathbb{P}[D = 1] + \mathbb{P}[D_1 = 2] = 0$, these long chains do not exist and the diameter will be of order $\log \log n$ for $\tau \in (2, 3)$.

For $\tau \in (1, 2)$ the corresponding branching process does not exist, because the degrees have an infinite mean. In this case, there is a finite number of vertices which have a *giant* degree, which together form a complete graph. All other vertices connect **whp** to at least one of these vertices with giant degree. So, the average hopcount H_n is 2 if the two vertices are connected to the same vertex with a giant degree and 3 otherwise. Based on this heuristic, the following theorem is proven in [11].

Theorem 3. *Let D be a random variable satisfying $\mathbb{P}[D > x] = x^{1-\tau}L(x)$ for $\tau \in (1, 2)$. Then,*

$$\lim_{n \rightarrow \infty} \mathbb{P}[H_n = 2] = 1 - \lim_{n \rightarrow \infty} \mathbb{P}[H_n = 2] = p_D \in (0, 1), \quad (27)$$

where p_D only depends on the distribution of D .

In [11], also the boundary cases $\tau = 1$ and $\tau = 2$ have been studied. For $\tau = 1$ the hopcount H_n turns out to be 2 **whp**. For $\tau = 2$ the distribution of H_n depends on the slowly varying function $L(x)$. An example has been given where H_n behaves the same as in Theorem 1 and an example is given where the average distance behaves as in Theorem 2.

For the boundary case $\tau = 3$ results are, to the best of our knowledge, not known. In Section 8, we will show that H_n is **whp** bounded from below by $(1 - \varepsilon) \log n / \log \log n$, for all $\varepsilon > 0$. This behavior is as expected, because of the known result on the preferential attachment model with $\tau = 3$ as is formulated in Theorem 5 below.

4.2 Preferential attachment models

For $m = 1$, where the graphs are trees, distances have been studied in various papers. See for example [25], where Pittel studies models (b) and (c), which are equivalent for $m = 1$, for $\delta = 0$ and $\delta = \infty$. He shows that in these models the height of the tree, h_t say, satisfies:

$$\frac{h_t}{\log t} \xrightarrow{\mathbb{P}} c_M, \quad t \rightarrow \infty, \quad (28)$$

where c_M is a constant depending on the model and its parameters. This holds for a wide variety of preferential attachment models. For an overview, see [4] and the references therein.

This result can also be shown to hold for all $\delta > -1$. Moreover, this immediately gives a $\log t$ upper bound on the diameter for $m \geq 2$, since this model can be constructed from the case $m = 1$ by grouping certain vertices, which can only decrease the distances between them.

For $m \geq 2$ and $\delta > 0$, so for $\tau > 3$, there also exists a $\log t$ lower bound on the diameter as was shown by Van der Hofstad and Hooghiemstra in [16]. We thus have the following theorem.

Theorem 4. *Fix $m \geq 1$ and $\delta > 0$. Then there exist constants $c_1, c_2 > 0$ such that, **whp**,*

$$c_1 \log t \leq \text{diam}(PA_{m,\delta}(t)) \leq c_2 \log t. \quad (29)$$

Their proof is an extension of an earlier proof by Bollobás and Riordan given in [5], where they prove the following for the case where $\tau = 3$.

Theorem 5. *Fix $m \geq 2$ and $\delta = 0$. Then, **whp**,*

$$(1 - \varepsilon) \frac{\log t}{\log \log t} \leq \text{diam}(PA_{m,0}(t)) \leq (1 + \varepsilon) \frac{\log t}{\log \log t}, \quad (30)$$

for some constant $C > 0$ and any constant $\varepsilon > 0$.

These proofs investigate the distance between vertices $t - 1$ and t and prove that this is, **whp**, at least equal to the lower bounds given. See also Section 7, where we will extend the result of Theorem 5 to average distances. Van der Hofstad and Hooghiemstra also did this with Theorem 4 for $\delta > 0$ in [16].

In [16] also the diameter is studied when $m \geq 2$ and $-m < \delta < 0$, so for $\tau \in (2, 3)$. The diameter in this case is shown to be at most of order $\log \log t$ as is stated in the following theorem.

Theorem 6. *Fix $m \geq 2$ and $\delta \in (-m, 0)$. Then, for every $\sigma > \frac{1}{3-\tau}$,*

$$\text{diam}(PA_{m,\delta}(t)) \leq C + \frac{4 \log \log t}{|\log(\tau - 2)|} + \frac{4\sigma \log \log t}{\log m}, \quad (31)$$

for some constant C .

Note that $\sigma > 1$. This theorem is proven for time $2t$ rather than time t , which will not make any difference for the result. The proof consists of several parts. First, it is shown that distances in $PA_{m,\delta}(2t)$ between vertices with a large degree at time t is small, and then it is shown that all other vertices are within a small distance of one of these vertices. The set of vertices with large degree is called the *core* and is defined as follows. Let $\sigma > \frac{1}{3-\tau}$, then

$$\text{Core}_t = \{i \in [t] \mid D_i(t) \geq (\log t)^\sigma\}. \quad (32)$$

The vertices in Core_t are split up into sets $\mathcal{N}^{(1)}, \dots, \mathcal{N}^{(k)}$, for some k , such that the degrees of the vertices in set $\mathcal{N}^{(i)}$ are larger than the degrees of the vertices in set $\mathcal{N}^{(j)}$ whenever $i < j$. When

$$k = \left\lfloor \frac{\log \log t}{|\log(\tau - 2)|} \right\rfloor, \quad (33)$$

this can be done in such a way that the maximum distance between any two vertices in $\mathcal{N}^{(1)}$ is at most C . Further, for every vertex $v \in \mathcal{N}^{(j)}$ there exists a vertex in $[2t]/[t]$ that is connected to both the vertex v and some vertex in $\mathcal{N}^{(j-1)}$. This shows that the distance between any two vertices in Core_t is at most

$$C + 4k \leq C + \frac{4 \log \log t}{|\log(\tau - 2)|}. \quad (34)$$

Next, the neighborhood of vertices $i \in [t]$ is studied. That is, starting from vertex i , connect its $m \geq 2$ edges. Then, successively, connect the m edges from each of the at most m vertices that i has connected to and have not yet been explored. Continue in the same fashion. The arising process, where all vertices up to distance k from vertex i have been explored, will be called the *k-exploration tree*, $\mathcal{T}_i^{(k)}$, of vertex i . We will call the event where an edge connects to a vertex that already was in the tree a *collision*.

Let $C' = \sigma / \log m$ and $k = C' \log \log t - 2$. Then, the probability that there are many collisions before the k -exploration tree hits $\text{Core}_t \cup \llbracket t^b \rrbracket$, for some $b > 0$, is small. When the k -exploration tree hits the core we are done. When it does not, it can be shown that with probability $1 - o(t^{-1})$, there exists a vertex $v \in [2t]/[t]$ which connects to both some vertex in $\mathcal{T}_i^{(k)}$ and a vertex in $\text{Core}_t \cup \llbracket t^b \rrbracket$. The constant b can be chosen such that $\llbracket t^b \rrbracket \subseteq \text{Core}_t$. So the distance between any vertex $i \in [t]$ and Core_t is, **whp**, at most

$$k + 2 = \frac{\sigma \log \log t}{\log m}. \quad (35)$$

A similar approach can be used for vertices $i \in [2t]/[t]$. It can be shown that in this case, with probability $1 - o(t^{-1})$, the $(k+1)$ -exploration tree hits $\text{Core}_{2t} \cup [t]$, or at least one of the vertices at distance $k+1$ attaches to some vertex in $\text{Core}_{2t} \cup [t]$. In Section 6 we will show that $\text{Core}_{2t} \subseteq [t]$, so, **whp**, the distance between any vertex $i \in [2t]/[t]$ and $[t]$ is at most

$$k + 2 = \frac{\sigma \log \log t}{\log m}. \quad (36)$$

These three bounds together give Theorem 6. In Section 5 we will prove that there is also a log log lower bound on the diameter for $\tau \in (2, 3)$.

It is not possible for the preferential attachment models (a-c) to have a power law with $\tau \in [1, 2]$. Deijfen et al. ([9]) proposed a model where a new vertex starts not with m edges, but with a random number of edges W . In this case, they conjecture that the degree sequence of the resulting graph obeys a power-law with exponent $\tau \in [1, 2]$, by letting W have a power law with this exponent τ . Results on distances, however, are not known for this model.

5 A log log lower bound on the diameter in PA models

In this section we will give a log log lower bound on the diameter of preferential attachment graphs with $m \geq 2$ and $\delta > -m$. All results apply to models (a), (b) and (c) simultaneously. The main result is stated in the following theorem:

Theorem 7. Fix $m \geq 2$ and $\delta > -m$. Let $k = \frac{\varepsilon}{\log m} \log \log t$, with $0 < \varepsilon < 1$. Then, **whp**,

$$\text{diam}(\text{PA}_{m,\delta}(t)) \geq k. \quad (37)$$

We will again prove this theorem for time $2t$ rather than time t . To show that the diameter of the graph is, **whp**, at least k , we will study, at time $2t$, the k -exploration trees of vertices $i \in [2t] \setminus [t]$, $\mathcal{T}_i^{(k)}$, as defined above. We again will call the event where an edge connects to a vertex that already was in the tree a *collision*. We call such a tree *proper* if the following conditions hold:

- The k -exploration tree has no collisions;
- All vertices of $\mathcal{T}_i^{(k)}$ are in $[2t] \setminus [t]$;
- No other vertex connects to a vertex in $\mathcal{T}_i^{(k)}$.

When such a tree exists in $\text{PA}_{m,\delta}(2t)$ for a certain vertex i then we know that the diameter is at least k , since the distance between the root of the tree i and the vertices at depth k is exactly k ; there cannot be a shorter route.

To prove that a proper k -exploration tree exists in $\text{PA}_{m,\delta}(2t)$, we will use the second moment method. Let $\mathfrak{T}_m^k(2t)$ be the set of all possible k -exploration trees that can exist in $\text{PA}_{m,\delta}(2t)$ and satisfy the first two conditions. Note that the order in which the edges are added matters: if two edges are added in a different order, then the arising exploration tree will be considered a different tree. Let $Z_{m,\delta}^k(2t)$ be the number of proper k -exploration trees in $\text{PA}_{m,\delta}(2t)$, i.e.,

$$Z_{m,\delta}^k(2t) = \sum_{\mathcal{T} \in \mathfrak{T}_m^k(2t)} \mathbb{I}\{\mathcal{T} \subseteq \text{PA}_{m,\delta}(2t) \text{ and } \mathcal{T} \text{ is proper}\}, \quad (38)$$

where $\mathbb{I}\{A\}$ is the indicator function of the event A . $\mathcal{T} \subseteq \text{PA}_{m,\delta}(2t)$ denotes the event that all edges of \mathcal{T} have been formed in $\text{PA}_{m,\delta}(2t)$.

In Section 5.1 we will investigate the first moment of $Z_{m,\delta}^k(2t)$ and prove the following:

Proposition 8. Fix $m \geq 2$ and $\delta > -m$. Let $k = \frac{\varepsilon}{\log m} \log \log t$, with $0 < \varepsilon < 1$. Then

$$\lim_{t \rightarrow \infty} \mathbb{E} [Z_{m,\delta}^k(2t)] = \infty. \quad (39)$$

The variance of $Z_{m,\delta}^k(2t)$ will be the subject of Section 5.2, where we will prove the following:

Proposition 9. Fix $m \geq 2$, $\delta > -m$ and $0 \leq k \leq \frac{\log \log t}{\log m}$. Then there exists a constant $c_{m,\delta} > 0$, such that, for t sufficiently large,

$$\text{Var} [Z_{m,\delta}^k(2t)] \leq c_{m,\delta} \frac{(\log t)^2}{t} \mathbb{E} [Z_{m,\delta}^k(2t)]^2 + \mathbb{E} [Z_{m,\delta}^k(2t)]. \quad (40)$$

We use these two propositions to prove the main result of this section:

Proof of Theorem 7. We first use the Chebychev inequality to obtain that

$$\mathbb{P}[\text{diam}(\text{PA}_{m,\delta}(2t)) < k] \leq \mathbb{P}[Z_{m,\delta}^k(2t) = 0] \leq \frac{\text{Var}[Z_{m,\delta}^k(2t)]}{\mathbb{E}[Z_{m,\delta}^k(2t)]^2}. \quad (41)$$

By Proposition 9, this is, for some constant $c_{m,\delta} > 0$, at most

$$\frac{c_{m,\delta} \frac{(\log t)^2}{t} \mathbb{E}[Z_{m,\delta}^k(2t)]^2 + \mathbb{E}[Z_{m,\delta}^k(2t)]}{\mathbb{E}[Z_{m,\delta}^k(2t)]^2} = c_{m,\delta} \frac{(\log t)^2}{t} + \frac{1}{\mathbb{E}[Z_{m,\delta}^k(2t)]} = o(1), \quad (42)$$

by Proposition 8. □

5.1 The first moment

Let $\mathcal{B}_{\mathcal{T}}$ denote the event that no vertex outside a tree \mathcal{T} connects to a vertex in this tree. We can then write that the expected number of proper k -exploration trees in $\text{PA}_{m,\delta}(2t)$ equals

$$\begin{aligned} \mathbb{E}[Z_{m,\delta}^k(2t)] &= \mathbb{E}\left[\sum_{\mathcal{T} \in \mathfrak{T}_m^k(2t)} \mathbb{I}\{\mathcal{T} \subseteq \text{PA}_{m,\delta}(2t) \text{ and } \mathcal{T} \text{ is proper}\}\right] \\ &= \sum_{\mathcal{T} \in \mathfrak{T}_m^k(2t)} \mathbb{E}[\mathbb{I}\{\mathcal{T} \subseteq \text{PA}_{m,\delta}(2t) \text{ and } \mathcal{T} \text{ is proper}\}] \\ &= \sum_{\mathcal{T} \in \mathfrak{T}_m^k(2t)} \mathbb{P}[\mathcal{T} \subseteq \text{PA}_{m,\delta}(2t) \text{ and } \mathcal{T} \text{ is proper}] \\ &= \sum_{\mathcal{T} \in \mathfrak{T}_m^k(2t)} \mathbb{P}[\mathcal{T} \text{ is proper} | \mathcal{T} \subseteq \text{PA}_{m,\delta}(2t)] \mathbb{P}[\mathcal{T} \subseteq \text{PA}_{m,\delta}(2t)] \\ &= \sum_{\mathcal{T} \in \mathfrak{T}_m^k(2t)} \mathbb{P}[\mathcal{B}_{\mathcal{T}} | \mathcal{T} \subseteq \text{PA}_{m,\delta}(2t)] \cdot \mathbb{P}[\mathcal{T} \subseteq \text{PA}_{m,\delta}(2t)]. \end{aligned} \quad (43)$$

We will first give a lower bound on the probability that a given k -exploration tree exists in the graph at time $2t$. For convenience we will write $a_{m,\delta} = \frac{m+\delta}{3(2m+\delta)}$.

Lemma 10. *Fix $m \geq 2$, $\delta > -m$ and $k \geq 0$. Given a possible proper k -exploration tree $\mathcal{T} \in \mathfrak{T}_m^k(2t)$, then, for t sufficiently large,*

$$\mathbb{P}[\mathcal{T} \subseteq \text{PA}_{m,\delta}(2t)] \geq \left(\frac{a_{m,\delta}}{t}\right)^{\frac{m^{k+1}-1}{m-1}-1}. \quad (44)$$

Proof. Since every vertex is added before time $2t$, the denominator in (1), (2) and (3) is at most $3t(2m+\delta)$. The degree of all vertices already in the graph is at least m , so the probability that a certain given edge is formed is at least

$$\frac{m+\delta}{3t(2m+\delta)} = \frac{a_{m,\delta}}{t}. \quad (45)$$

Since exactly $\frac{m^{k+1}-1}{m-1} - 1$ edges have to be formed to form the given tree \mathcal{T} , we have that

$$\mathbb{P}[\mathcal{T} \subseteq \text{PA}_{m,\delta}(2t)] \geq \left(\frac{a_{m,\delta}}{t}\right)^{\frac{m^{k+1}-1}{m-1}-1}. \quad (46)$$

□

We will now give a lower bound on the probability that no other vertex connects to a given tree. We will write $m_\delta = m + 1 + \delta$.

Lemma 11. *Fix $m \geq 2$, $\delta > -m$ and $0 \leq k \leq \frac{\log \log t}{\log m}$. Given a possible proper k -exploration tree $\mathcal{T} \in \mathfrak{T}_m^k(2t)$, then, for t sufficiently large,*

$$\mathbb{P}[\mathcal{B}_\mathcal{T} | \mathcal{T} \subseteq \text{PA}_{m,\delta}(2t)] \geq \left(1 - \frac{m_\delta m^{k+1}}{t}\right)^{mt}. \quad (47)$$

Proof. First note that for $k \leq \frac{\log \log t}{\log m}$ and t sufficiently large, $m_\delta m^{k+1} \leq m_\delta m \log t \leq t$. So $0 \leq 1 - \frac{m_\delta m^{k+1}}{t} \leq 1$. Further note that vertices $[t]$ cannot connect to a vertex in \mathcal{T} , since $\mathcal{T} \subseteq [2t] \setminus [t]$. In the remainder of the proof we will refer to *outside* edges as those edges that do not belong to \mathcal{T} , of which there are exactly $mt - \left(\frac{m^{k+1}-1}{m-1} - 1\right)$ added after time t . Let $\mathcal{E}_n(A)$ denote the event that the n -th outside edge added after time t connects to a vertex in A and let $\bar{\mathcal{E}}_n(A)$ be the negation of $\mathcal{E}_n(A)$. We use induction on the number of outside edges that did not connect to the tree \mathcal{T} , i.e., we show that:

$$\mathbb{P}\left[\bigcap_{i=1}^n \bar{\mathcal{E}}_i(\mathcal{T}) \mid \mathcal{T} \subseteq \text{PA}_{m,\delta}(2t)\right] \geq \left(1 - \frac{m_\delta m^{k+1}}{t}\right)^n, \quad (48)$$

by induction on $n = 0, \dots, mt - \left(\frac{m^{k+1}-1}{m-1} - 1\right)$. For $n = 0$ the above clearly holds. Now assume that the above holds for $0 \leq n < mt - \left(\frac{m^{k+1}-1}{m-1} - 1\right)$, then

$$\begin{aligned} & \mathbb{P}\left[\bigcap_{i=1}^{n+1} \bar{\mathcal{E}}_i(\mathcal{T}) \mid \mathcal{T} \subseteq \text{PA}_{m,\delta}(2t)\right] \\ &= \mathbb{P}\left[\bar{\mathcal{E}}_{n+1}(\mathcal{T}) \mid \bigcap_{i=1}^n \bar{\mathcal{E}}_i(\mathcal{T}) \cap \{\mathcal{T} \subseteq \text{PA}_{m,\delta}(2t)\}\right] \mathbb{P}\left[\bigcap_{i=1}^n \bar{\mathcal{E}}_i(\mathcal{T}) \mid \mathcal{T} \subseteq \text{PA}_{m,\delta}(2t)\right] \\ &\geq \left(1 - \mathbb{P}\left[\mathcal{E}_{n+1}(\mathcal{T}) \mid \bigcap_{i=1}^n \bar{\mathcal{E}}_i(\mathcal{T}) \cap \{\mathcal{T} \subseteq \text{PA}_{m,\delta}(2t)\}\right]\right) \cdot \left(1 - \frac{m_\delta m^{k+1}}{t}\right)^n. \end{aligned} \quad (49)$$

Since it is known that at the time that the $(n+1)$ -th outside edge after time t is added, no other outside edge has connected to a vertex in the tree, we know that the degree of all vertices in the tree at that moment is at most $m+1$. Further, since this edge is added after time t , the denominator of (1), (2) and (3) will be at least t . Thus, the right hand side of (49) is at least

$$\begin{aligned} & \left(1 - \sum_{i \in \mathcal{T}} \frac{m+1+\delta}{t}\right) \cdot \left(1 - \frac{m_\delta m^{k+1}}{t}\right)^n \geq \left(1 - \frac{m_\delta m^{k+1}}{t}\right) \cdot \left(1 - \frac{m_\delta m^{k+1}}{t}\right)^n \\ &= \left(1 - \frac{m_\delta m^{k+1}}{t}\right)^{n+1}, \end{aligned} \quad (50)$$

where the inequality holds because there are less than m^{k+1} vertices in the tree. Applying the above to $n = mt - \left(\frac{m^{k+1}-1}{m-1} - 1\right)$, we obtain that

$$\mathbb{P}[\mathcal{B}_\mathcal{T} | \mathcal{T} \subseteq \text{PA}_{m,\delta}(2t)] \geq \left(1 - \frac{m_\delta m^{k+1}}{t}\right)^{mt - \left(\frac{m^{k+1}-1}{m-1} - 1\right)} \geq \left(1 - \frac{m_\delta m^{k+1}}{t}\right)^{mt}. \quad (51)$$

□

We finally give a lower bound on the number of possible proper k -exploration trees that can be formed. It should be noted that when a vertex i connects to a vertex j , we will always have that $i > j$. So when exploring a vertex i in the exploration tree, all m vertices this vertex connects to have a smaller label than i .

Lemma 12. *Fix $m \geq 2$ and $0 \leq k \leq \frac{\log \log t}{\log m}$. Then, for t sufficiently large, the number of possible proper k -exploration trees at time $2t$ is at least*

$$\left(\frac{t}{m^{k+1}} \right)^{\frac{m^{k+1}-1}{m-1}}. \quad (52)$$

Proof. For t sufficiently large and $k \leq \frac{\log \log t}{\log m}$, $m^{k+1} \leq m \log t \leq t$. Since the k -exploration tree of a vertex i has to be proper, there are no collisions, so the number of vertices in the tree equals

$$|\mathcal{T}_i^{(k)}| = \frac{m^{k+1} - 1}{m - 1}. \quad (53)$$

For any subset $X \subseteq [2t] \setminus [t]$ with $|X| = \frac{m^{k+1}-1}{m-1}$ there exists at least one possible proper k -exploration tree. To see this, first order the vertex labels in descending order. Let the first vertex, i.e. the vertex with the largest label, be the root of the tree. Then let the next m vertices be the vertices at distance 1 from the root, the next m^2 vertices be the vertices at distance 2 from the root, etcetera, until the last m^k vertices which will be at distance k from the root. This way, all vertices will connect to m vertices with a smaller label, i.e., vertices that were already in the graph when the vertex was added, so this is a possible proper k -exploration tree with all vertices in X .

The number of subsets of $[2t] \setminus [t]$ of size $\frac{m^{k+1}-1}{m-1}$ is $\binom{t}{\frac{m^{k+1}-1}{m-1}}$ which is at least

$$\left(\frac{t}{\frac{m^{k+1}-1}{m-1}} \right)^{\frac{m^{k+1}-1}{m-1}} \geq \left(\frac{t}{m^{k+1}} \right)^{\frac{m^{k+1}-1}{m-1}}. \quad (54)$$

Here we used that for $1 \leq b \leq a$ we have that $(a-i)b \geq (b-i)a$ for all $0 \leq i < b$, so that

$$\binom{a}{b} = \prod_{i=0}^{b-1} \frac{a-i}{b-i} \geq \left(\frac{a}{b} \right)^b. \quad (55)$$

□

Note that the number of possible proper k -exploration trees with vertices X is in fact much larger. For instance, all vertices at the same distance from the root can be permuted in any order. This gives an extra factor $\prod_{i=1}^k (m^i)!$. This, however, is of a much smaller order than the above, and thus is not necessary for our proof. We will include the exact computation of the number of possible proper k -exploration trees for completeness.

For given m and k let $A(m, k)$ be the number of possible proper trees given a set $X \subseteq [2t] \setminus [t]$ of size $\frac{m^{k+1}-1}{m-1}$. Note that this number does not depend on the set X , since only the order in which the vertices in the tree are added matters. Thus

$$|\mathfrak{T}_m^k(2t)| = \binom{t}{\frac{m^{k+1}-1}{m-1}} A(m, k). \quad (56)$$

As noted above, the vertex with the largest label has to be the root of the tree. We can then divide the remaining vertices in m groups, all of size $\frac{m^k-1}{m-1}$. Each such a group will form the tree of depth $k-1$ at one of the edges of the root of the tree. This can be done in $A(m, k-1)$ ways. We thus get the following recursive formula:

$$A(m, k) = \frac{\left(\frac{m^{k+1}-1}{m-1} - 1 \right)!}{\left(\left(\frac{m^k-1}{m-1} \right)! \right)^m} \cdot (A(m, k-1))^m, \quad (57)$$

where the first fraction is the number of ways the remaining vertices can be split in m groups of equal size. So:

$$\begin{aligned}
A(m, k) &= \frac{\binom{m^{k+1}-1}{m-1}!}{\left(\binom{m^k-1}{m-1}\right)^m} \cdot (A(m, k-1))^m \\
&= \frac{\binom{m^{k+1}-1}{m-1}!}{\left(\binom{m^k-1}{m-1}\right)^m} \frac{\left(\binom{m^k-1}{m-1}\right)^m}{\left(\binom{m^{k-1}-1}{m-1}\right)^{m^2}} \cdots \frac{\left(\binom{m^2-1}{m-1}\right)^{m^{k-1}}}{\left(\binom{m^1-1}{m-1}\right)^{m^k}} \frac{\left(\binom{m^1-1}{m-1}\right)^{m^k}}{\left(\binom{m^0-1}{m-1}\right)^{m^{k+1}}} \\
&= \frac{\binom{m^{k+1}-1}{m-1}!}{\binom{m^k-1}{m-1}^m \binom{m^{k-1}-1}{m-1}^{m^2} \cdots \binom{m^1-1}{m-1}^{m^k} \left(\binom{m^0-1}{m-1}\right)^{m^{k+1}}}. \tag{58}
\end{aligned}$$

We can now combine the three bounds above to get a lower bound on the expected number of proper k -exploration trees.

Corollary 13. Fix $m \geq 2$, $\delta > -m$ and $0 \leq k \leq \frac{\log \log t}{\log m}$. Then, for t sufficiently large,

$$\mathbb{E} [Z_{m,\delta}^k(2t)] \geq \frac{t}{a_{m,\delta}} \left(\frac{a_{m,\delta}}{m^{k+1}}\right)^{m^{k+1}} \left(1 - \frac{m_\delta m^{k+1}}{t}\right)^{mt}. \tag{59}$$

Proof. Using the bounds from Lemmas 10, 11 and 12 we get that

$$\begin{aligned}
\mathbb{E} [Z_{m,\delta}^k(2t)] &= \sum_{\mathcal{T} \in \mathfrak{T}_m^k(2t)} \mathbb{P}[\mathcal{B}_{\mathcal{T}} | \mathcal{T} \subseteq \text{PA}_{m,\delta}(2t)] \cdot \mathbb{P}[\mathcal{T} \subseteq \text{PA}_{m,\delta}(2t)] \\
&\geq \sum_{\mathcal{T} \in \mathfrak{T}_m^k(2t)} \left(1 - \frac{m_\delta m^{k+1}}{t}\right)^{mt} \left(\frac{a_{m,\delta}}{t}\right)^{\binom{m^{k+1}-1}{m-1}-1} \\
&= |\mathfrak{T}_m^k(2t)| \left(1 - \frac{m_\delta m^{k+1}}{t}\right)^{mt} \left(\frac{a_{m,\delta}}{t}\right)^{\binom{m^{k+1}-1}{m-1}-1} \\
&\geq \left(\frac{t}{m^{k+1}}\right)^{\binom{m^{k+1}-1}{m-1}} \left(1 - \frac{m_\delta m^{k+1}}{t}\right)^{mt} \left(\frac{a_{m,\delta}}{t}\right)^{\binom{m^{k+1}-1}{m-1}-1} \\
&= \frac{t}{a_{m,\delta}} \left(\frac{a_{m,\delta}}{m^{k+1}}\right)^{\binom{m^{k+1}-1}{m-1}} \left(1 - \frac{m_\delta m^{k+1}}{t}\right)^{mt} \\
&\geq \frac{t}{a_{m,\delta}} \left(\frac{a_{m,\delta}}{m^{k+1}}\right)^{m^{k+1}} \left(1 - \frac{m_\delta m^{k+1}}{t}\right)^{mt}. \tag{60}
\end{aligned}$$

□

The factor t in the corollary above turns out to be crucial for the remainder of the proof. This factor arises from the fact that there is exactly one edge less in a proper k -exploration tree than there are vertices.

We can now show that the expected number of k -exploration trees tends to infinity, for $k = \frac{\varepsilon}{\log m} \log \log t$, with $0 < \varepsilon < 1$.

Proof of Proposition 8. First note that for $k = \frac{\varepsilon}{\log m} \log \log t$, with $0 < \varepsilon < 1$, $m^k = (\log t)^\varepsilon$. We

can then use Corollary 13 to get that

$$\begin{aligned}
\lim_{t \rightarrow \infty} \mathbb{E} [Z_{m,\delta}^k(2t)] &\geq \lim_{t \rightarrow \infty} \frac{t}{a_{m,\delta}} \left(\frac{a_{m,\delta}}{m^{k+1}} \right)^{m^{k+1}} \left(1 - \frac{mm_\delta m^{k+1}}{mt} \right)^{mt} \\
&= \lim_{t \rightarrow \infty} \frac{t}{a_{m,\delta}} \left(\frac{a_{m,\delta}}{m(\log t)^\varepsilon} \right)^{m(\log t)^\varepsilon} \left(1 - \frac{m^2 m_\delta (\log t)^\varepsilon}{mt} \right)^{mt} \\
&= \lim_{t \rightarrow \infty} \frac{t}{a_{m,\delta}} \left(\frac{a_{m,\delta}}{m(\log t)^\varepsilon} \right)^{m(\log t)^\varepsilon} \frac{mt \log \left(1 - \frac{m^2 m_\delta (\log t)^\varepsilon}{mt} \right)}{e} \\
&= \lim_{t \rightarrow \infty} \frac{t}{a_{m,\delta}} \left(\frac{a_{m,\delta}}{m(\log t)^\varepsilon} \right)^{m(\log t)^\varepsilon} \frac{-m^2 m_\delta (\log t)^\varepsilon + \mathcal{O}\left(\frac{(\log t)^{2\varepsilon}}{t}\right)}{e} = \infty. \quad (61)
\end{aligned}$$

□

It is easy to see that the same argument can be applied to $k = \frac{\log \log t}{\log m} - \frac{\log \log \log t}{\log m} - 1$.

5.2 The second moment

In this section we will investigate the variance of $Z_{m,\delta}^k(2t)$. To shorten the notation, for a k -exploration tree $\mathcal{T} \in \mathfrak{T}_m^k(2t)$, let $\mathcal{F}_\mathcal{T}$ denote the event that $\mathcal{T} \subseteq \text{PA}_{m,\delta}(2t)$ and \mathcal{T} is proper. Then, the variance of the number of proper k -exploration trees in $\text{PA}_{m,\delta}(2t)$

$$\begin{aligned}
\text{Var} [Z_{m,\delta}^k(2t)] &= \text{Var} \left[\sum_{\mathcal{T} \in \mathfrak{T}_m^k(2t)} \mathbb{I}\{\mathcal{T} \subseteq \text{PA}_{m,\delta}(2t) \text{ and } \mathcal{T} \text{ is proper}\} \right] \\
&= \text{Var} \left[\sum_{\mathcal{T} \in \mathfrak{T}_m^k(2t)} \mathbb{I}\{\mathcal{F}_\mathcal{T}\} \right] \\
&= \sum_{\mathcal{T}, \mathcal{T}' \in \mathfrak{T}_m^k(2t)} \text{Cov} [\mathbb{I}\{\mathcal{F}_\mathcal{T}\}, \mathbb{I}\{\mathcal{F}_{\mathcal{T}'}\}] \\
&= \sum_{\substack{\mathcal{T}, \mathcal{T}' \in \mathfrak{T}_m^k(2t) \\ \mathcal{T} \neq \mathcal{T}'}} (\mathbb{P}[\mathcal{F}_\mathcal{T} \cap \mathcal{F}_{\mathcal{T}'}] - \mathbb{P}[\mathcal{F}_\mathcal{T}] \mathbb{P}[\mathcal{F}_{\mathcal{T}'}]) \\
&\quad + \sum_{\mathcal{T} \in \mathfrak{T}_m^k(2t)} \mathbb{P}[\mathcal{F}_\mathcal{T}] (1 - \mathbb{P}[\mathcal{F}_\mathcal{T}]). \quad (62)
\end{aligned}$$

We will first study the terms of the first sum in the following lemma.

Lemma 14. *Fix $m \geq 2$, $\delta > -m$ and $0 \leq k \leq \frac{\log \log t}{\log m}$. Let $\mathcal{T}, \mathcal{T}' \in \mathfrak{T}_m^k(2t)$ with $\mathcal{T} \neq \mathcal{T}'$. Then, for t sufficiently large,*

$$\mathbb{P}[\mathcal{F}_\mathcal{T} \cap \mathcal{F}_{\mathcal{T}'}] - \mathbb{P}[\mathcal{F}_\mathcal{T}] \mathbb{P}[\mathcal{F}_{\mathcal{T}'}] \leq \left(\left(1 + \frac{2m_\delta m \log t}{t} \right)^{2m \log t} - 1 \right) \mathbb{P}[\mathcal{F}_\mathcal{T}] \mathbb{P}[\mathcal{F}_{\mathcal{T}'}]. \quad (63)$$

Proof. When $\mathcal{T} \cap \mathcal{T}' \neq \emptyset$, at least one edge of one of the trees will connect to a vertex in the other tree, so the trees \mathcal{T} and \mathcal{T}' cannot both be proper. Thus, for $\mathcal{T} \cap \mathcal{T}' \neq \emptyset$, trivially (63) holds.

For $\mathcal{T} \cap \mathcal{T}' = \emptyset$, we have to take a closer look at the probabilities involved. All three probabilities in the lemma are a product over all edges of the probability that either the edge does not connect to any of the vertices in the tree(s) or the probability that the edge makes a prescribed connection in (one of) the tree(s). Let $\mathcal{E}_{j,s}(A)$ denote the event that the j -th edge of vertex s connects to a vertex in A , with $\mathcal{E}_{j,s}(i) = \mathcal{E}_{j,s}(\{i\})$. Let $\bar{\mathcal{E}}_{j,s}(A)$ be the complement of $\mathcal{E}_{j,s}(A)$. We have that

$$\mathbb{P}[\mathcal{E}_{j,s}(A)] = \sum_{i \in A} \mathbb{P}[\mathcal{E}_{j,s}(i)], \quad (64)$$

because the events on the righthand side are disjunct. These probabilities are given by the growth rules (1-3).

Suppose that the j -th edge, $1 \leq j \leq m$, of a vertex t_0 should not connect to a vertex in $\mathcal{T} \cup \mathcal{T}'$. Then in $\mathbb{P}[\mathcal{F}_{\mathcal{T}} \cap \mathcal{F}_{\mathcal{T}'}]$, there will be a factor

$$\mathbb{P}[\bar{\mathcal{E}}_{j,t_0}(\mathcal{T} \cup \mathcal{T}')] = 1 - \mathbb{P}[\mathcal{E}_{j,t_0}(\mathcal{T} \cup \mathcal{T}')] = 1 - \sum_{i \in \mathcal{T} \cup \mathcal{T}'} \mathbb{P}[\mathcal{E}_{j,t_0}(i)]. \quad (65)$$

In $\mathbb{P}[\mathcal{F}_{\mathcal{T}}] \mathbb{P}[\mathcal{F}_{\mathcal{T}'}]$, there will be a factor

$$\left(1 - \sum_{i \in \mathcal{T}} \mathbb{P}[\mathcal{E}_{j,t_0}(i)]\right) \left(1 - \sum_{i \in \mathcal{T}'} \mathbb{P}[\mathcal{E}_{j,t_0}(i)]\right). \quad (66)$$

It is easy to see that $1 - x - y \leq (1 - x)(1 - y)$ for $x, y \geq 0$, so (66) is at least as big as (65).

When the j -th edge, $1 \leq j \leq m$, of a vertex t_0 , $t + 1 \leq t_0 \leq 2t$, should connect to a vertex $h \in \mathcal{T}$, then in $\mathbb{P}[\mathcal{F}_{\mathcal{T}} \cap \mathcal{F}_{\mathcal{T}'}]$ there will only be a factor

$$\mathbb{P}[\mathcal{E}_{j,t_0}(h)], \quad (67)$$

since it will then automatically not connect to a vertex in \mathcal{T}' . In $\mathbb{P}[\mathcal{F}_{\mathcal{T}}] \mathbb{P}[\mathcal{F}_{\mathcal{T}'}]$, however, there will be a factor

$$\mathbb{P}[\mathcal{E}_{j,t_0}(h)] \left(1 - \sum_{i \in \mathcal{T}'} \mathbb{P}[\mathcal{E}_{j,t_0}(i)]\right). \quad (68)$$

When we multiply this by $(1 - \sum_{i \in \mathcal{T}'} \mathbb{P}[\mathcal{E}_{j,t_0}(i)])^{-1}$ it will be at least (67) again. By symmetry, the same holds when an edge should connect to a vertex in \mathcal{T}' . Since the degree of the vertices in the trees is at most $m + 1$, the edges of interest are added after time t and there are less than m^{k+1} vertices in the tree, we have that

$$\left(1 - \sum_{i \in \mathcal{T}'} \mathbb{P}[\mathcal{E}_{j,t_0}(i)]\right)^{-1} \leq \left(1 - \frac{m_{\delta} m^{k+1}}{t}\right)^{-1}. \quad (69)$$

Since there are less than m^{k+1} edges in both \mathcal{T} and \mathcal{T}' , for $\mathcal{T} \cap \mathcal{T}' = \emptyset$,

$$\begin{aligned} \mathbb{P}[\mathcal{F}_{\mathcal{T}} \cap \mathcal{F}_{\mathcal{T}'}] - \mathbb{P}[\mathcal{F}_{\mathcal{T}}] \mathbb{P}[\mathcal{F}_{\mathcal{T}'}] &\leq \left(\left(1 - \frac{m_{\delta} m^{k+1}}{t}\right)^{-2m^{k+1}} - 1 \right) \mathbb{P}[\mathcal{F}_{\mathcal{T}}] \mathbb{P}[\mathcal{F}_{\mathcal{T}'}] \\ &= \left(\left(1 + \frac{m_{\delta} m^{k+1}}{t - m_{\delta} m^{k+1}}\right)^{2m^{k+1}} - 1 \right) \mathbb{P}[\mathcal{F}_{\mathcal{T}}] \mathbb{P}[\mathcal{F}_{\mathcal{T}'}] \\ &\leq \left(\left(1 + \frac{m_{\delta} m \log t}{t - m_{\delta} m \log t}\right)^{2m \log t} - 1 \right) \mathbb{P}[\mathcal{F}_{\mathcal{T}}] \mathbb{P}[\mathcal{F}_{\mathcal{T}'}] \\ &\leq \left(\left(1 + \frac{2m_{\delta} m \log t}{t}\right)^{2m \log t} - 1 \right) \mathbb{P}[\mathcal{F}_{\mathcal{T}}] \mathbb{P}[\mathcal{F}_{\mathcal{T}'}]. \end{aligned} \quad (70)$$

□

We will now give an upper bound on the factor in front of the probabilities.

Lemma 15. *Fix $m \geq 2$ and $\delta > -m$. Then, for t sufficiently large,*

$$\left(1 + \frac{2m_{\delta} m \log t}{t}\right)^{2m \log t} - 1 \leq c_{m,\delta} \frac{(\log t)^2}{t}, \quad (71)$$

where $c_{m,\delta} = 8m_{\delta} m^2$.

Proof. In [22], it is shown that

$$(1+x)^n \leq 1+2nx, \quad (72)$$

whenever $0 \leq (n-1)x < \frac{1}{2}$ and $n \geq 2$. Since

$$\lim_{t \rightarrow \infty} (2m \log t - 1) \frac{2m_\delta m \log t}{t} \leq \lim_{t \rightarrow \infty} \frac{c_{m,\delta}}{2} \frac{(\log t)^2}{t} = 0, \quad (73)$$

we indeed have that, for t sufficiently large, $(2m \log t - 1) \frac{2m_\delta m \log t}{t} < \frac{1}{2}$. We thus have that,

$$\left(1 + \frac{2m_\delta m \log t}{t}\right)^{2m \log t} - 1 \leq c_{m,\delta} \frac{(\log t)^2}{t}. \quad (74)$$

□

We can now use the two lemmas above to give an upper bound on the variance of $Z_{m,\delta}^k(2t)$ in terms of the expectation of $Z_{m,\delta}^k(2t)$.

Proof of Proposition 9. Let $c_{m,\delta} = 8m_\delta m^2$. Then, using Lemmas 14 and 15, we have that

$$\begin{aligned} \text{Var} [Z_{m,\delta}^k(2t)] &= \sum_{\substack{\mathcal{T}, \mathcal{T}' \in \mathfrak{T}_m^k(2t) \\ \mathcal{T} \neq \mathcal{T}'}} (\mathbb{P}[\mathcal{F}_\mathcal{T} \cap \mathcal{F}_{\mathcal{T}'}] - \mathbb{P}[\mathcal{F}_\mathcal{T}] \mathbb{P}[\mathcal{F}_{\mathcal{T}'}]) + \sum_{\mathcal{T} \in \mathfrak{T}_m^k(2t)} \mathbb{P}[\mathcal{F}_\mathcal{T}] (1 - \mathbb{P}[\mathcal{F}_\mathcal{T}]) \\ &\leq \sum_{\substack{\mathcal{T}, \mathcal{T}' \in \mathfrak{T}_m^k(2t) \\ \mathcal{T} \neq \mathcal{T}'}} \left(\left(1 + \frac{2m_\delta m \log t}{t}\right)^{2m \log t} - 1 \right) \mathbb{P}[\mathcal{F}_\mathcal{T}] \mathbb{P}[\mathcal{F}_{\mathcal{T}'}] + \sum_{\mathcal{T} \in \mathfrak{T}_m^k(2t)} \mathbb{P}[\mathcal{F}_\mathcal{T}] \\ &\leq c_{m,\delta} \frac{(\log t)^2}{t} \sum_{\substack{\mathcal{T}, \mathcal{T}' \in \mathfrak{T}_m^k(2t) \\ \mathcal{T} \neq \mathcal{T}'}} \mathbb{P}[\mathcal{F}_\mathcal{T}] \mathbb{P}[\mathcal{F}_{\mathcal{T}'}] + \mathbb{E} [Z_{m,\delta}^k(2t)] \\ &\leq c_{m,\delta} \frac{(\log t)^2}{t} \sum_{\mathcal{T}, \mathcal{T}' \in \mathfrak{T}_m^k(2t)} \mathbb{P}[\mathcal{F}_\mathcal{T}] \mathbb{P}[\mathcal{F}_{\mathcal{T}'}] + \mathbb{E} [Z_{m,\delta}^k(2t)] \\ &= c_{m,\delta} \frac{(\log t)^2}{t} \mathbb{E} [Z_{m,\delta}^k(2t)]^2 + \mathbb{E} [Z_{m,\delta}^k(2t)]. \end{aligned} \quad (75)$$

□

6 Late vertices have small degrees in PA models

Define the set of vertices with large degree at time t as

$$\text{Core}_t = \{i \in [t] \mid D_i(t) \geq (\log t)^\sigma\}, \quad (76)$$

for $\sigma > 1$. In the following theorem we will prove that, for models (a-c), all vertices with large degree will be early vertices. We need this result to prove Theorem 6.

Theorem 16. *Fix $m \geq 2, \delta > -m$ and $\sigma > 1$. Then, **whp**,*

$$\text{Core}_{2t} \subseteq [t]. \quad (77)$$

Proof. Note that

$$\begin{aligned} \mathbb{P}[\text{Core}_{2t} \subseteq [t]] &\geq 1 - \sum_{i=t+1}^{2t} \mathbb{P}[D_i(2t) \geq (\log 2t)^\sigma] \\ &\geq 1 - \sum_{i=t+1}^{2t} \mathbb{P}[D_t(2t) \geq (\log 2t)^\sigma], \end{aligned} \quad (78)$$

because vertex t is more likely to have a large degree than vertices added after time t . In Lemma 17 we will show that $\mathbb{P}[D_t(2t) \geq (\log 2t)^\sigma] = o\left(\frac{1}{t}\right)$, so that

$$\mathbb{P}[\text{Core}_{2t} \subseteq [t]] \geq 1 - o(1). \quad (79)$$

□

Lemma 17. *Fix $m \geq 2, \delta > -m$ and $\sigma > 1$. Then,*

$$\mathbb{P}[D_t(2t) \geq (\log 2t)^\sigma] = o\left(\frac{1}{t}\right). \quad (80)$$

Proof. As noted in Section 2.2, $\text{PA}_{m,\delta}(2t)$ can be constructed from $\text{PA}_{1,\delta'}(2mt)$, with $\delta' = \delta/m$. Let us label the vertices of $\text{PA}_{1,\delta'}(2mt)$ by v_1, \dots, v_{2mt} to avoid confusion. Thus identify, for $i \in [2t]$, vertices $v_{(i-1)m+1}, \dots, v_{im}$ with vertex i . So (80) is equivalent to

$$\mathbb{P}[D_{v_{(t-1)m+1}}(2mt) + \dots + D_{v_{tm}}(2mt) \geq (\log 2t)^\sigma] = o\left(\frac{1}{t}\right). \quad (81)$$

We will now color the vertices and edges in the following way. Color the vertices $v_1, \dots, v_{(t-1)m}$ and all edges between these vertices blue and color the vertices $v_{(t-1)m+1}, \dots, v_{tm}$ and the m edges that are attached to them at time mt red. When a vertex, that was added after time mt , connects to a blue (red) vertex, also color that vertex and its edge blue (red). Color vertices with a self-loop and its edge blue. Then, at time $2mt$ the total degree of vertices $v_{(t-1)m+1}, \dots, v_{tm}$ is at most equal to the number of red edges plus m , because no blue edges are connected to these red vertices, and all red edges are connected with at most one endpoint to these vertices. The only exception are the first m red edges, which might connect with both endpoints to these vertices, hence we have to add m to the number of red edges. Thus

$$\mathbb{P}[D_{v_{(t-1)m+1}}(2mt) + \dots + D_{v_{tm}}(2mt) \geq (\log 2t)^\sigma] \leq \mathbb{P}[\#\{\text{red edges}\} + m \geq (\log 2t)^\sigma]. \quad (82)$$

Since we will bound the righthand side of the formula above, it is allowed to increase the probability of attaching to a red vertex, or, equivalently, to decrease the probability of attaching to a blue vertex. It is also allowed to increase the total degree of the red vertices, or to decrease the total degree of the blue vertices. All this will only increase the probability of the number of red edges being large.

Therefore, we are allowed to assume that the first m red edges are all self-loops. Further, we will not allow for self-loops after time t , which will increase the probability of attaching to a red vertex in models (a) and (b), in model (c) nothing changes. When we look at model (c), we see that the degrees should only be updated after each m -th vertex has been added. For $j \geq mt$, no more than m edges and vertices can be added before updating the degrees, so

$$\begin{aligned} \mathbb{P}\left[v_{j+1} \text{ connects to a red vertex} \mid \text{PA}_{1,\delta'}^{(c)}(j)\right] &= \frac{\sum_{v \text{ red}} (D_v(m \lfloor j/m \rfloor) + \delta')}{m \lfloor j/m \rfloor (2 + \delta')} \\ &\leq \frac{\sum_{v \text{ red}} (D_v(j) + \delta')}{j(2 + \delta') - m(2 + \delta')}. \end{aligned} \quad (83)$$

Thus, we are allowed to update the degrees after adding each vertex, but then we have to lower the total weight that blue vertices and edges contribute to the connecting probabilities by $m(2 + \delta')$. The above bound on the connecting probabilities also holds for models (a) and (b).

Since we are only interested in the number of red and blue vertices and edges, the problem reduces to the following Pólya urn scheme. Let there be an urn with, at time s , $S_1(s)$ red balls, corresponding to the total weight that red vertices and edges contribute to the connecting probabilities, and $S_2(s)$ blue balls, corresponding to the lowered total weight that blue vertices and edges contribute to the connecting probabilities. At time $s = 0$ we will start with $S_1(0) = m(2 + \delta')$ and $S_2(0) = m(t - 1)(2 + \delta') - m(2 + \delta')$. We then successively take one ball proportional to the

number of balls of a certain color, and replace it together with another $2 + \delta'$ balls of the same color. This corresponds to attaching a new vertex to a vertex of that color.

So $\frac{S_1(mt)}{2+\delta'}$ has the same distribution as the number of red edges at time $2mt$. So

$$\mathbb{P}[D_t(2t) \geq (\log 2t)^\sigma] \leq \mathbb{P}\left[\frac{S_1(mt)}{2+\delta'} + m \geq (\log 2t)^\sigma\right]. \quad (84)$$

To analyse the probability on the righthand side, we need De Finetti's Theorem. This theorem states that for an infinite sequence of exchangeable random variables $\{X_i\}_{i=1}^\infty$, $X_i \in \{0, 1\}$, there exists a random variable U with $\mathbb{P}[U \in [0, 1]] = 1$, such that for all $1 \leq k \leq n$,

$$\mathbb{P}[X_1 = \dots = X_k = 1, X_{k+1} = \dots = X_n] = \mathbb{E}[U^k(1-U)^{n-k}]. \quad (85)$$

The random variable U can be computed explicitly. Note that this implies that

$$\mathbb{P}\left[\sum_{i=1}^n X_i = k\right] = \mathbb{E}[\mathbb{P}[\text{BIN}(n, U) = k]]. \quad (86)$$

Let X_i denote the indicator that the i -th ball drawn in the Pólya urn scheme described above is red. As shown in Section 11.1 of [15], $\{X_i\}_{i=1}^\infty$ is an infinite exchangeable sequence. Note that

$$S_1(s) = (2 + \delta')m + (2 + \delta') \sum_{i=1}^s X_i. \quad (87)$$

So,

$$\mathbb{P}\left[\frac{S_1(mt)}{2+\delta'} + m \geq (\log 2t)^\sigma\right] = \mathbb{E}\left[\mathbb{P}[\text{BIN}(mt, U) \geq (\log 2t)^\sigma - 2m]\right], \quad (88)$$

where U turns out to have a Beta-distribution with parameters $\alpha = m$ and $\beta = m(t-2)$.

We can rewrite (88) as

$$\begin{aligned} & \mathbb{E}\left[\mathbb{P}[\text{BIN}(mt, U) \geq (\log 2t)^\sigma - 2m] \mid U \leq g(t)\right] \mathbb{P}[U \leq g(t)] \\ & \quad + \mathbb{E}\left[\mathbb{P}[\text{BIN}(mt, U) \geq (\log 2t)^\sigma - 2m] \mid U > g(t)\right] \mathbb{P}[U > g(t)] \\ & \leq \mathbb{E}\left[\mathbb{P}[\text{BIN}(mt, U) \geq (\log 2t)^\sigma - 2m] \mid U = g(t)\right] + \mathbb{P}[U > g(t)], \end{aligned} \quad (89)$$

where $g(t) = \frac{(\log 2t)^\sigma - 2m}{7(mt)}$. We then have that

$$(\log 2t)^\sigma - 2m \geq 7\mathbb{E}[\text{BIN}(mt, g(t))], \quad (90)$$

so that we may apply Corollary 2.4 of [20], which states that

$$\mathbb{P}[X \geq x] \leq e^{-x}, \quad x \geq 7\mathbb{E}[X], \quad (91)$$

where $X \sim \text{BIN}(n, p)$. So (89) is at most

$$e^{-(\log 2t)^\sigma + 2m} + \mathbb{P}[U > g(t)] = o\left(\frac{1}{t}\right) + \mathbb{P}[U > g(t)]. \quad (92)$$

It remains to show that also $\mathbb{P}[U > g(t)] = o\left(\frac{1}{t}\right)$. Since $\alpha, \beta > 1$, we have that the probability density function of U is unimodal, with its turning point at $t = \frac{\alpha-1}{\alpha+\beta-2}$ ([26]). It is easy to verify that $g(t) \geq \frac{\alpha-1}{\alpha+\beta-2}$, for t sufficiently large, so that

$$\begin{aligned} \mathbb{P}[U > g(t)] & \leq (1-g(t)) \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} (g(t))^{\alpha-1} (1-g(t))^{\beta-1} \\ & \leq \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} (1-g(t))^\beta. \end{aligned} \quad (93)$$

Using Stirling's formula, one can show that (see e.g. [1]) there exists a constant $C > 0$, such that (93) is at most

$$\begin{aligned} C \frac{\beta^\alpha}{\Gamma(\alpha)} (1 - g(t))^\beta &\leq C(mt)^m \left(1 - \frac{(\log 2t)^\sigma}{8m(t-2)}\right)^{m(t-2)} \\ &= C(mt)^m e^{m(t-2) \log\left(1 - \frac{(\log 2t)^\sigma}{8m(t-2)}\right)} \\ &= C(mt)^m e^{-\frac{1}{8}(\log 2t)^\sigma + \mathcal{O}\left(\frac{(\log 2t)^{2\sigma}}{t}\right)} = o\left(\frac{1}{t}\right), \end{aligned} \quad (94)$$

because $\sigma > 1$.

Note that we in fact proved that $\mathbb{P}[D_t(2t) \geq (\log 2t)^\sigma] = o(t^{-\gamma})$, for any constant γ . \square

7 Average distances for $\delta = 0$ in PA models

In [5] Bollobás and Riordan study model (a) for $\delta = 0$ and $m \geq 2$. They show that in this case, **whp**, $\text{PA}_{m,0}(t)$ has a diameter satisfying

$$\frac{\log t}{\log(3Cm^2 \log t)} \leq \text{diam}(\text{PA}_{m,0}(t)) \leq (1 + \varepsilon) \frac{\log t}{\log \log t}, \quad (95)$$

for some constant $C > 0$ and any constant $\varepsilon > 0$.

We will show in this section that also the average distance between vertices grows like $\log t / \log \log t$, by extending the proof of the above lower bound in a similar way as was done in [16] for $\delta > 0$. The upper bound on average distances follows trivially from (95).

Lemma 18. *Fix $m \geq 1$ and $\delta = 0$. Let $H_t = \text{dist}_{\text{PA}_{m,\delta}(t)}(A_1, A_2)$ be the distance between two vertices A_1 and A_2 , chosen uniformly at random, at time t . Then, for some constant $C > 0$, **whp**,*

$$H_t \geq \frac{\log t}{\log(2Cm^2 \log t)}. \quad (96)$$

Proof. Let $L = \frac{\log t}{\log(2Cm^2 \log t)}$ and define

$$B_t \equiv \#\{i, j \in [t], i < j : \text{dist}_{\text{PA}_{m,\delta}(t)}(i, j) \leq L\}. \quad (97)$$

In [5] it was shown that, for model (a), the expected number of paths of length l between vertices i and j in $\text{PA}_{m,\delta}(t)$ is bounded from above by $\frac{(Cm^2)^l}{\sqrt{ij}} (2 \log t)^{l-1}$, for some constant $C > 0$. This can be extended to models (b) and (c) as was done in [16]. So

$$\mathbb{P}[\text{dist}_{\text{PA}_{m,\delta}(t)}(i, j) \leq L] \leq \sum_{l=1}^L \frac{(Cm^2)^l}{\sqrt{ij}} (2 \log t)^{l-1} = \frac{Cm^2 (2Cm^2 \log t)^L - 1}{\sqrt{ij} (2Cm^2 \log t - 1)} \leq \frac{1}{\sqrt{ij}} \frac{t}{\log t}. \quad (98)$$

Since $\sum_{i=1}^{j-1} \frac{1}{\sqrt{i}} \leq 2\sqrt{j}$, we have that

$$\mathbb{E}[B_t] \leq \sum_{1 \leq i < j \leq t} \frac{1}{\sqrt{ij}} \frac{t}{\log t} \leq 2 \frac{t^2}{\log t}. \quad (99)$$

Thus,

$$\mathbb{P}[H_t \leq L] = \mathbb{E}[\mathbb{I}[\text{dist}_{\text{PA}_{m,\delta}(t)}(A_1, A_2) \leq L]] \leq \frac{2\mathbb{E}[B_t] + t}{t^2} = o(1). \quad (100)$$

\square

8 Distances for $\tau = 3$ in the configuration model

In Section 7 we saw that in preferential attachment models with $\tau = 3$ the average distance grows like $\log n / \log \log n$. A similar result is not known for the configuration model, but we will prove that in this case, **whp**, the average distance is larger than $\log n / \log \log n$. More precisely:

Theorem 19. *Let $\{D_i\}_{i=1}^n$ be a sequence of i.i.d. random variables, with $\mathbb{P}[D_1 > x] = cx^{-2}(1 + o(1))$ for $x = 1, 2, \dots$. Then, for any $\varepsilon > 0$,*

$$\mathbb{P}[H_n \leq (1 - \varepsilon) \frac{\log n}{\log \log n}] = o(1). \quad (101)$$

Proof. When $H_n = l$ then there should at least be a path of length l between vertices 1 and 2. Let $i \rightarrow j$ denote the event that vertex i is adjacent to vertex j in $\text{CM}_n(\{D_i\}_{i=1}^n)$ and recall that $L_n = \sum_{i=1}^n D_i$. Then, for $l \leq (1 - \varepsilon) \frac{\log n}{\log \log n}$,

$$\begin{aligned} \mathbb{P}[H_n = l] &\leq \sum_{1 \leq i_1, \dots, i_{l-1} \leq n} \mathbb{P}[1 \rightarrow i_1 \rightarrow \dots \rightarrow i_{l-1} \rightarrow 2] \\ &\leq \mathbb{E} \left[\sum_{1 \leq i_1, \dots, i_{l-1} \leq n} \frac{D_1 D_{i_1}}{L_n - 1} \frac{D_{i_1} D_{i_2}}{L_n - 3} \dots \frac{D_{i_{l-2}} D_{i_{l-1}}}{L_n - (2l - 3)} \frac{D_{i_{l-1}} D_2}{L_n - (2l - 1)} \right] \end{aligned} \quad (102)$$

Note that **whp**, $D_1 D_2 \leq \log n$, since the degrees do not depend on n , and $L_n \geq n$, because $D_i \geq 1$ for all i with probability 1. In Lemma 20 we will show that $\nu_n / \log n$ converges in probability to c/μ , so **whp** $\nu_n \leq (1 + \varepsilon)c/\mu \log n$. Denote the event that these three bounds hold by \mathcal{A}_n . Since \mathcal{A}_n holds **whp**, (102) is **whp** equal to

$$\begin{aligned} &\mathbb{E} \left[\sum_{1 \leq i_1, \dots, i_{l-1} \leq n} \frac{D_1 D_{i_1}}{L_n - 1} \frac{D_{i_1} D_{i_2}}{L_n - 3} \dots \frac{D_{i_{l-2}} D_{i_{l-1}}}{L_n - (2l - 3)} \frac{D_{i_{l-1}} D_2}{L_n - (2l - 1)} \mathbb{I}\{\mathcal{A}_t\} \right] \\ &\leq \mathbb{E} \left[\frac{D_1 D_2}{\left(\frac{1}{2} L_n\right)^l} \left(\sum_{i=1}^n D_i^2 \right)^{l-1} \mathbb{I}\{\mathcal{A}_t\} \right] = \mathbb{E} \left[\frac{2^l D_1 D_2}{L_n} \nu_n^{l-1} \mathbb{I}\{\mathcal{A}_t\} \right] \\ &\leq \frac{2^l \log n}{n} \left((1 + \varepsilon) \frac{c}{\mu} \log n \right)^{l-1} \leq 2 \left((1 + \varepsilon) \frac{2c}{\mu} \right)^{l-1} n^{-\varepsilon} = o\left(\frac{1}{k}\right). \end{aligned} \quad (103)$$

□

Lemma 20. *Let $\{D_i\}_{i=1}^n$ be a sequence of i.i.d. random variables, with $\mathbb{P}[D_1 > x] = cx^{-2}(1 + o(1))$ for $x = 1, 2, \dots$. Let $\mu = \mathbb{E}[D_1]$. Define*

$$\nu_n = \frac{\sum_{i=1}^n D_i^2}{\sum_{i=1}^n D_i}. \quad (104)$$

Then

$$\frac{\nu_n}{\log n} \xrightarrow{\mathbb{P}} \frac{c}{\mu}, \quad \text{for } n \rightarrow \infty. \quad (105)$$

Proof. First, note that, by the strong law of large numbers,

$$\frac{1}{n} \sum_{i=1}^n D_i \xrightarrow{\mathbb{P}} \mu, \quad (106)$$

where $\mu > 0$. It remains to show that

$$\frac{1}{n} \frac{\sum_{i=1}^n D_i^2}{\log n} \xrightarrow{\mathbb{P}} c, \quad (107)$$

because then (105) holds by Slutsky's theorem ([14], p. 318).

Let $a_n = \sqrt{n \log n}$. Then

$$\mathbb{P} \left[\max_{i=1}^n D_i > a_n \right] \leq n \mathbb{P}[D_1 > a_n] = \frac{nc}{a_n^2} (1 + o(1)) = o(1), \quad (108)$$

so, **whp**, $\max_{i=1}^n D_i \leq a_n$. Thus, also **whp**,

$$\frac{1}{n} \frac{\sum_{i=1}^n D_i^2}{\log n} = \frac{1}{n} \frac{\sum_{i=1}^n D_i^2 \mathbb{I}\{D_i \leq a_n\}}{\log n} \equiv X, \quad (109)$$

say.

Note that for a positive and integer valued random variable M ,

$$\begin{aligned} \mathbb{E}[M^2] &= \sum_{m=1}^{\infty} m^2 \mathbb{P}[M = m] = \sum_{m=1}^{\infty} \sum_{x=1}^m (2x-1) \mathbb{P}[M = m] \\ &= \sum_{x=1}^{\infty} (2x-1) \sum_{m=x}^{\infty} \mathbb{P}[M = m] = \sum_{x=1}^{\infty} (2x-1) \mathbb{P}[M \geq m], \end{aligned} \quad (110)$$

and similarly,

$$\begin{aligned} \mathbb{E}[M^4] &= \sum_{m=1}^{\infty} m^4 \mathbb{P}[M = m] = \sum_{m=1}^{\infty} \sum_{x=1}^m (4x^3 - 6x^2 + 4x - 1) \mathbb{P}[M = m] \\ &= \sum_{x=1}^{\infty} (4x^3 - 6x^2 + 4x - 1) \sum_{m=x}^{\infty} \mathbb{P}[M = m] = \sum_{x=1}^{\infty} (4x^3 - 6x^2 + 4x - 1) \mathbb{P}[M \geq m]. \end{aligned} \quad (111)$$

So, when we study the variance of X ,

$$\begin{aligned} \text{Var}[X] &= \frac{1}{n(\log n)^2} \text{Var} [D_1^2 \mathbb{I}\{D_1 \leq a_n\}] \leq \frac{1}{n(\log n)^2} \mathbb{E} [D_1^4 \mathbb{I}\{D_1 \leq a_n\}] \\ &= \frac{1}{n(\log n)^2} \sum_{x=1}^{a_n} (4x^3 - 6x^2 + 4x - 1) \mathbb{P}[D_1 \mathbb{I}\{D_1 \leq a_n\} \geq x] \\ &\leq \frac{1}{n(\log n)^2} \sum_{x=1}^{a_n} 4x^3 \mathbb{P}[D_1 \geq x] = \frac{1}{n(\log n)^2} \sum_{x=1}^{a_n} 4cx(1 + o(1)) \\ &= \frac{2ca_n^2}{n(\log n)^2} (1 + o(1)) = \frac{2c}{\log n} (1 + o(1)). \end{aligned} \quad (112)$$

Let $\varepsilon > \frac{1}{(\log n)^{1/4}}$. Then, by the Chebychev inequality,

$$\mathbb{P}[|X - \mathbb{E}[X]| > \varepsilon] \leq \frac{\text{Var}[X]}{\varepsilon^2} = o(1). \quad (113)$$

So, X converges in probability to $\mathbb{E}[X]$, which converges to c :

$$\begin{aligned} \mathbb{E}[X] &= \frac{1}{\log n} \sum_{x=1}^{a_n} (2x-1) \mathbb{P}[D_1 \mathbb{I}\{D_1 \leq a_n\} \geq x] = \frac{1}{\log n} \sum_{x=1}^{a_n} (2x-1) (\mathbb{P}[D_1 \geq x] - \mathbb{P}[D_1 > a_n]) \\ &= \frac{1}{\log n} \sum_{x=1}^{a_n} \left(\frac{2c}{x} (1 + o(1)) \right) - \frac{a_n}{\log n} ca_n^{-2} (1 + o(1)) = 2c \frac{\log a_n}{\log n} (1 + o(1)) = c(1 + o(1)). \end{aligned} \quad (114)$$

□

9 Conclusion

In this thesis we studied two random graph models which have a power law degree sequence, namely the configuration model and the preferential attachment model. An overview is given of results on distances in these graphs and some new results have been presented. These results indeed show the *small world* property observed in many real-world networks. They also show that distances in these two models behave similarly, i.e., for equal power-law exponent τ , the distances seem to be of the same order.

This overview, however, is not complete. For example, a log log lower bound on average distances in the preferential model for $\tau \in (2, 3)$ is missing. Also a log / log log upper bound on distances in the configuration model for $\tau = 3$ is not known, although this behavior is expected from the preferential attachment model.

We would also like to see how distances in the various other preferential attachment models studied in the literature behave. Ideally, we would like to find a general approach which can be applied for many of these models as Bhamidi has found for $m = 1$ ([4]).

It would also be interesting to study random processes on these random graph models. This can model, for example, the spread of diseases in a social network or the spread of viruses in a computer network.

References

- [1] M. Abramowitz and I.A. Stegun (eds.), *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, US Government Printing Office, Washington, D.C., 1964.
- [2] A.-L. Barabási, *Linked: The New Science of Networks*, Perseus, Cambridge, 2002.
- [3] A.-L. Barabási and R. Albert, *Emergence of Scaling in Random Networks*, Science, vol. 286, pp. 509-512, 1999.
- [4] S. Bhamidi, *Universal Techniques to Analyze Preferential Attachment Trees: Global and Local Analysis*, Preprint, 2007.
- [5] B. Bollobás and O. Riordan, *The Diameter of a Scale-Free Random Graph*, Combinatorica, vol. 24, pp. 5-34, 2004.
- [6] B. Bollobás, O. Riordan, J. Spencer and G. Tusnády, *The Degree Sequence of a Scale-Free Random Graph Process*, Random Structures and Algorithms, vol. 18, pp. 279-290, 2001.
- [7] T. Britton, M. Deijfen and A. Martin-Löf, *Generating Simple Random Graphs with Prescribed Degree Distribution*, Journal of Statistical Physics, vol. 124, pp. 1377-1397, 2006.
- [8] P.L. Davies, *The Simple Branching Process: A Note on Convergence when the Mean is Infinite*, Journal of Applied Probability, vol. 15, pp. 466-480, 1978.
- [9] M. Deijfen, H. van den Esker, R. van der Hofstad and G. Hooghiemstra, *A Preferential Attachment Model with Random Initial Degrees*, Preprint, 2007.
- [10] P. Erdős and A. Rényi, *On the Evolution of Random Graphs*, Publications of the Mathematical Institute of the Hungarian Academy of Science, vol. 5, pp. 17-61, 1960.
- [11] H. van den Esker, R. van der Hofstad, G. Hooghiemstra and D. Znamenski, *Distances in Random Graphs with Infinite Mean Degrees*, Extremes, vol. 8, pp. 111-141, 2006.
- [12] M. Faloutsos, P. Faloutsos and C. Faloutsos, *On Power-Law Relationships of the Internet Topology*, Computer Communications Review, vol. 29, pp. 251-262, 1999.
- [13] W. Feller, *An Introduction to Probability Theory and Its Applications*, Volume II, 2nd Edition, John Wiley and Sons, New York, 1971.

- [14] G.R. Grimmett and D. Stirzaker, *Probability and Random Processes*, 2nd Edition, Oxford, Clarendon, 1992.
- [15] R. van der Hofstad, *Random Graphs and Complex Networks*, Lecture Notes, 2008.
- [16] R. van der Hofstad and G. Hooghiemstra, *Diameters in Preferential Attachment Models*, Preprint, 2008.
- [17] R. van der Hofstad, G. Hooghiemstra and P. van Mieghem, *Distances in Random Graphs with Finite Variance Degrees*, Random Structures and Algorithms, vol. 27, pp. 76-123, 2005.
- [18] R. van der Hofstad, G. Hooghiemstra and D. Znamenski, *Distances in Random Graphs with Finite Mean and Infinite Variance Degrees*, Electronic Journal of Probability, vol 12., pp. 703-766, 2007.
- [19] R. van der Hofstad, G. Hooghiemstra and D. Znamenski, *A Phase Transition for the Diameter of the Configuration Model*, Preprint, 2008.
- [20] S. Janson, T. Łuczak and A. Ruciński, *Random Graphs*, John Wiley & Sons, Inc., New York, 2000.
- [21] S. Milgram, *The Small World Problem*, Physiology Today, vol. 2, pp. 60-67, 1967.
- [22] D.S. Mitrinović, *Analytic Inequalities*, Springer-Verlag, Berlin, 1970.
- [23] M.E.J. Newman, *The Structure and Function of Complex Networks*, SIAM Review, vol. 45, pp. 167-256, 2003.
- [24] M.E.J. Newman, S.H. Strogatz and D.J. Watts, *Random Graphs with Arbitrary Degree Distributions and Their Applications*, Physical Review E, vol. 64, 026118, 2001.
- [25] B. Pittel, *Note on the Heights of Random Recursive Trees and Random m-ary Search Trees*, Random Structures and Algorithms, vol. 5, pp. 337-347, 1994.
- [26] G.P. Wadsworth and J.G. Bryan, *Introduction to Probability and Random Variables*, McGraw-Hill Book Company, Inc., New York, 1960.