Diameters in preferential attachment models

Remco van der Hofstad* Gerard Hooghiemstra†

May 25, 2007

Abstract

In this paper, we investigate the diameter in preferential attachment (PA-) models, thus quantifying the statement that these models are small worlds. There is a substantial amount of literature proving that, in quite generality, PA-graphs possess power-law degree sequences with exponent $\tau > 2$. The models studied here are such that edges are attached to older vertices proportional to the degree plus a constant, i.e., we consider linear PA-models. We prove that the diameter is bounded by a constant times $\log t$, where t is the size of the graph. When the power-law exponent τ exceeds 3, then we also prove a lower bound of the form $\frac{\log t}{\log \log t}$, while when $\tau \in (2,3)$, we improve the upper bound to a constant times $\log \log t$. These bounds are consistent with predictions by physicists that the distances in PA-graphs are similar to the ones in other scale-free random graphs, where distances have been shown to be of order $\log \log t$, when $\tau \in (2,3)$, and of order $\log t$ when $\tau > 3$.

1 Introduction

In the past decade, many examples have been found of real world complex networks that are *small worlds* and *scale-free*. The small-world phenomenon states that distances in many networks are small. The scale-free phenomenon states that the degree sequences in many networks satisfy a power law. See [2, 19, 30] for reviews on complex networks, and [4] for a more expository account. As a result, these complex networks are not at all like classical random graphs (see [3, 7, 29] and the references therein), particularly since the classical models do not have power-law degrees. As a result, these empirical findings have ignited enormous research on adaptations of the classical random graph that do obey power-law degree sequences. See [9] for the most general models, as well as a review of the models under investigation.

While these models have power-law degree sequences, they do not explain *why* many complex networks are scale-free. A possible and convincing explanation was given by Barabási and Albert [5] by a phenomenon called *preferential attachment* (PA). Preferential attachment models the growth of the network in such a way that new vertices are more likely to add their edges to already present vertices having a high degree. For example, in a social network, a newcomer is more likely to get to know a person who is socially active, and, therefore, already has a high degree. Interestingly, PA-models have power-law degree sequences, and, therefore, preferential attachment offers a convincing explanation why many real world networks have power-law degree sequences. As a result, many papers appeared that study such models. See e.g. [1, 8, 10, 11, 12, 13, 15, 17] and the references therein. The literature primarily focusses on three main questions. The first is to prove that such random graphs are indeed scale-free [1, 8, 10, 11, 15, 17]. The second is to show that the resulting models are small worlds by investigating the distances in them.

*Department of Mathematics and Computer Science, Eindhoven University of Technology, P.O. Box 513, 5600 MB Eindhoven, The Netherlands. E-mail: rhofstad@win.tue.nl

†Delft University of Technology, Electrical Engineering, Mathematics and Computer Science, P.O. Box 5031, 2600 GA Delft, The Netherlands. E-mail: G.Hooghiemstra@ewi.tudelft.nl

See for example [13] for a result on the diameter. In non-rigorous work, it is often suggested that many of the scale-free models, such as the configuration model, the models in [9] and the PA-models, have similar properties for their distances. Distances in the configuration model have been shown to depend on the number of finite moments of the degree distribution. The natural question is therefore whether the same applies to preferential attachment models. A partial result is in [13], and this question will be taken up again here. A third key question for PA-models is their vulnerability, for example to deliberate attack [11] or to the spread of a disease [6]. The most complete discussion of scale-free random graphs and processes living on them is given in [20].

In this paper, we investigate the diameter in some PA-models. The models that we investigate produce a *graph sequence* or *graph process* $\{G_m(t)\}$, which, for fixed $t \geq 1$ or $t \geq 2$ yields a graph with t vertices and mt edges for some given integer $m \geq 1$. We shall consider three slight variations of the model, which we denote by model (a), (b) and (c).

(a) The first model is an extension of the Barabási-Albert model, formulated rigorously in [15]. We start with $G_1(1)$ consisting of a single vertex with a single self-loop. We denote the vertices of the graph by $1, 2, \ldots$, so that the vertices of $G_1(t)$ are equal to $\{1, 2, \ldots, t\}$. We denote the degree of node i by $d_i(t)$, where, in the degree, a self-loop increases the degree by 2.
Then, for $m = 1$, and conditionally on $G_1(t)$, the growth rule to obtain $G_1(t+1)$ is as follows. We add a single vertex $t+1$ having a single edge. This edge is connected to a second end point, which is equal to $t+1$ with probability proportional to $1+\delta$, and to a vertex $i \in G_1(t)$ with probability proportional to $d_i(t)+\delta$, where $\delta \geq -1$ is a parameter of the model. Thus,

$$\mathbb{P}\big(t+1 \to i \big| G_1(t)\big) = \begin{cases} \frac{1+\delta}{t(2+\delta)+(1+\delta)}, & \text{for } i = t+1, \\ \frac{d_i(t)+\delta}{t(2+\delta)+(1+\delta)}, & \text{for } i = 1, 2, \ldots, t. \end{cases} \tag{1.1}$$

The model with integer $m > 1$, is defined in terms of the model for $m = 1$ as follows. We start with $G_1(mt)$, with $\delta' = \frac{\delta}{m} \geq -1$, and denote the vertices in $G_1(mt)$ by $1', \ldots, (mt)'$. Then we identify the vertices $1', 2' \ldots, m'$ in $G_1(mt)$ to be vertex 1 in $G_m(t)$, and for $1 < j \leq t$, the vertices $((j-1)m+1)', \ldots, (jm)'$ in $G_1(mt)$ to be vertex j in $G_m(t)$; in particular the degree $d_j(t)$of vertex j in $G_m(t)$ is equal to the sum of the degrees of the vertices $((j-1)m+1)', \ldots, (jm)'$ in $G_1(mt)$. This defines the model for integer $m \geq 1$. Observe that the range of δ is $[-m, \infty)$.
The resulting graph $G_m(t)$ has precisely mt edges and t vertices at time t, but is not necessarily connected. For $\delta = 0$ we obtain the model studied in [15].

(b) The second model is identical to the one above, apart from the fact that no self-loops are allowed. We start again with the definition for $m = 1$. To prevent a self loop in the first step, we let $G_1(1)$ undefined, and start from $G_1(2)$, which is defined by the vertices 1 and 2 joined together by 2 edges. Then, for $t > 2$, we define, conditionally on $G_1(t)$, the growth rule to obtain $G_1(t+1)$ as follows. For $\delta \geq -1$,

$$\mathbb{P}\big(t+1 \to i \big| G_1(t)\big) = \frac{d_i(t)+\delta}{t(2+\delta)}, \qquad \text{for } i = 1, \ldots, t. \tag{1.2}$$

The model with $m > 1$ is again defined in terms of the model for $m = 1$, in precisely the same way as in model (a).

(c) In the third model, and conditionally on $G_m(t)$, the end points of each of the m edges of vertex $t+1$ are chosen *independently*, and are equal to a vertex $i \in G_m(t)$, with probability proportionally to $d_i(t)+\delta$, where $\delta \geq -m$. We start again from $G_m(2)$, with the nodes 1 and 2 joined together by $2m$, $m \geq 1$, edges. Since the end point of the edges are chosen independently we can give the definition of $\{G_m(t)\}_{t\geq 2}$, for $m \geq 1$, in one step. For $1 \leq j \leq m$,

$$\mathbb{P}(j^{\text{th}} \text{ edge of } t+1 \text{ is connected to } i|G_m(t)) = \frac{d_i(t)+\delta}{t(2m+\delta)}, \qquad \text{for } i = 1, \ldots, t. \tag{1.3}$$

In this model, as is the case in model (b), the graph $G_m(t)$ is a connected random graph with precisely t vertices and mt edges.

Remark 1.1. *In models (a) and (b) for $m > 1$, the choice of $\delta' = \frac{\delta}{m}$ is such that in the resulting graph $G_m(t)$, where m vertices in $G_1(mt)$ are grouped together to a single vertex in $G_m(t)$, the end points of the added edges are chosen according to the degree plus the constant δ.*

Remark 1.2. *For $m = 1$, the models (b) and (c) are the same. This fact shall be used later on.*

One would expect the models (a)–(c) to behave quite similarly. In [18], it was proved that for model (c), the degree sequence is close to a power law with exponent $\tau = 3 + \frac{\delta}{m}$. For model (a) and $\delta = 0$, this was proved in [15], while in [17], power-law degree sequences are proved in rather large generality.

The goal in this paper is to study the diameter in the above models, as a first step towards the study of distances in PA-models. In non-rigorous work, it is often suggested that the distances are similarly behaved in the various scale-free random graph models, such as the configuration model or various models with conditional independence of edges as in [9]. The results on distances are most complete for the configuration model, see e.g. [22, 23, 26, 27, 28, 33]. In the configuration model, there are various cases depending on the tails of the degree distribution. When the degrees have infinite mean, then distances are bounded [22], when they have finite mean but infinite variance, distances grow like $\log\log t$ [27, 33], where t is the size of the graph, while, for finite variance degrees, the distances grow proportionally to $\log t$ [26]. Similar results for models with conditionally independent edges exist, see e.g. [9, 16, 21, 31]. Thus, for these classes of models, distances are quite well understood. If the distances in PA-models are similar to the ones in e.g. the configuration model, then we should have that the distances are of order $\log t$ when $\tau > 3$, i.e., $\delta > 0$, while they should be of order $\log\log t$ when $\tau \in (2, 3)$, i.e., for $\delta < 0$. In PA-models with linear growth of the number of edges, infinite mean degrees have not been observed, so this case does not arise. An attempt in this direction is in [18], where a preferential attachment is presented in which a *random* number of edges per new vertex is added. In this model, it is shown that the degrees again obey a power law with exponent equal to $\tau = \min\{3 + \frac{\delta}{\mu}, \tau_w\}$, where τ_w is the power-law exponent for the number of edges added and $\mu \leq \infty$ the expected number of added edges per vertex. Thus, when $\tau_w \in (1, 2)$, infinite mean degrees can arise.

There are few results on distances in PA-models. In [13], it was proved that in model (a) and for $\delta = 0$, for which $\tau = 3$, the diameter of the resulting graph is equal to $\frac{\log t}{\log\log t}(1+o(1))$. Unfortunately, the matching result for the configuration model has not been proved, so that this does not allow us to verify whether the models have similar distances. In this paper, we take a first step towards the verification of the heuristic, by investigating the diameter of the preferential attachment graph both for $\delta > 0$ and for $\delta < 0$. In the following section, we describe our precise results.

1.1 The diameter in preferential attachment models

In this section, we present the diameter results for the PA-models (a)-(c). We prove that for model (b) and (c) and for all $\delta > -m$, the diameter of $G_m(t)$ is bounded by a large constant times $\log t$. This result does not hold for model (a), since the graph is not necessarily connected. When $\delta \geq 0$, we adapt the argument in [13] to prove that for all three models considered here the diameter is bounded from below by $(1-\varepsilon)\frac{\log t}{\log\log t}$, while, for $\delta < 0$, we prove that the diameter is bounded above by a large constant times $\log\log t$. This establishes a phase transition for the diameter of PA-models when δ changes sign. We now state the precise results. In the results below, for a sequence of events $\{E_t\}_{t\geq 1}$, we write that E_t occurs *with high probability* (**whp**) when $\lim_{t\to\infty} \mathbb{P}(E_t) = 1$.

Theorem 1.3 (A $\log t$ upper bound for the diameter)**.** *Fix $m \geq 1$ and $\delta > -m$ in models (b) and (c). Then, there exists a constant $C = C(m, \delta) > 0$ such that*

$$\mathbb{P}\Big(\mathrm{diam}(G_m(t)) > C \log t\Big) = o(1), \quad t \to \infty, \tag{1.4}$$

i.e., **whp**, *the diameter of $G_m(t)$ is at most $C \log t$.*

When $m = 1$, so that the graphs are in fact trees, there is a sharper result proved by Pittel [32]. In this case, Pittel shows that the height of the tree, which is equal to the maximal graph distance between vertex 1 and any of the other vertices, grows like $\frac{1+\delta}{\gamma(2+\delta)} \log t(1 + o(1))$, where γ solves the equation

$$\gamma + (1 + \delta)(1 + \log \gamma) = 0. \tag{1.5}$$

This immediately proves that the diameter is at least as large, and suggests that the diameter has size $2\frac{1+\delta}{\gamma(2+\delta)} \log t(1 + o(1))$. Scale-free trees have received substantial attention in the literature, we refer to [14, 32] and the references therein. It is not hard to see that a similar result as proved in [32] also follows for model (a). This is also proved when $\delta = 0$ in [14], where it is shown that the diameter in model (a) has size $\gamma^{-1} \log t$, where γ is the solution of (1.5) when $\delta = 0$. Thus, we see that the $\log t$ upper bound in Theorem 1.3 is sharp, at least for $m = 1$. To see the result for model (a), we note that N_t, the number of connected components of $G_1(t)$ in model (a), has distribution $N_t = 1 + I_2 + \cdots + I_t$, where I_i is the indicator that the i^{th} edge connects to itself, so that $\{I_i\}_{i=2}^t$ are independent indicator variables with

$$\mathbb{P}(I_i = 1) = \frac{1 + \delta}{(2 + \delta)(i - 1) + 1 + \delta}. \tag{1.6}$$

As a result, it is not hard to see that $N_t/\log t$ converges in probability to $(1+\delta)/(2+\delta) < 1$, so that **whp** there exists a largest connected component of size at least $t/\log t$. The law of any connected component of size s_t in model (a) is *equal* in distribution to the law of the graph $G_1(s_t + 1)$ in model (b), apart from the fact that the vertices 1 and 2 in $G_1(s_t + 1)$ are identified (thus creating a unique self-loop). This close connection between the two models allows one to transfer the results for model (b) to model (a).

Theorem 1.4 (A lower bound for the diameter)**.** *Fix $m \geq 1$ and $\delta \geq 0$ in models (a)–(c). Then, for every $\varepsilon > 0$,*

$$\mathbb{P}\Big(\mathrm{diam}(G_m(t)) < (1 - \varepsilon)\frac{\log t}{\log \log t}\Big) = o(1), \tag{1.7}$$

i.e., **whp**, *the diameter of $G_m(t)$ is at least $(1 - \varepsilon)\frac{\log t}{\log \log t}$.*

We conjecture that for $\delta > 0$, the above lower bound is *not* sharp:

Conjecture 1.5. *Fix $m \geq 1$ and $\delta > 0$ in models (a)–(c). Then, there exists a constant $C = C(m, \delta) > 0$ such that*

$$\mathbb{P}\Big(\mathrm{diam}(G_m(t)) = C \log t(1 + o(1))\Big) = o(1), \tag{1.8}$$

i.e., $\mathrm{diam}(G_m(t))/\log t$ *converges in probability to a positive constant.*

Theorems 1.3, 1.4 and Conjecture 1.5 indicate that distances in PA-models are similar to the ones in other scale-free models for $\tau > 3$. We now turn to the case where $\delta < 0$ and hence $\tau = 3 + \delta/m \in (2, 3)$.

Theorem 1.6 (A $\log\log t$ upper bound on the diameter for $\delta < 0$)**.** *Fix $m \geq 2$ and assume that $\delta \in (-m, 0)$ in models (a)–(c). Then, for every $\sigma > \frac{1}{3-\tau}$ and with*

$$C_G = \frac{4}{|\log(\tau-2)|} + \frac{4\sigma}{\log m}. \tag{1.9}$$

the diameter of $G_m(t)$ is, **whp***, bounded above by $C_G \log\log t$, as $t \to \infty$.*

In the last result, we do not obtain a sharp result in terms of the constant. However, the proof suggests that for most pairs of vertices the distance should be equal to $\frac{4}{|\log(\tau-2)|}\log\log t(1+o(1))$.

The results stated above are consistent with the predictions by physicists that the distances in preferential attachment graphs should be similar to the ones in other scale-free random graphs. The only two missing bounds for a complete picture of the diameter in these PA-models are a $\log t$ lower bound for $\delta > 0$ and a $\log\log t$ lower bound for $\delta < 0$.

1.2 Organization of the paper

This paper is organized as follows. In Section 2, we prove the $\log t$ upper bound for the diameter stated in Theorem 1.3. In Section 3, we prove the $\log t/\log\log t$ lower bound for the diameter stated in Theorem 1.4, and in Section 4, we prove the $\log\log t$ upper bound on the diameter for $\delta < 0$ of Theorem 1.6.

2 An upper bound on the diameter: Proof of Theorem 1.3

For model (c), and with $s_i > s_j$ fixed, we write $s_i \longrightarrow s_j$ when the *first* edge of s_i is connected to vertex s_j. In case of model (b), we write $s_i \longrightarrow s_j$ when in $\{G_1(t)\}_{t\geq 1}$ vertex $(s_i-1)m+1$ is connected to one of the vertices $(s_j-1)m+1, \ldots, s_j m$. In model (a) self-loops are possible, so in this case the proof breaks down, which is understandable as model (a) is not necessarily connected.

For $s_1 = s > s_2 > \cdots > s_k = 1$, and denoting $\vec{s}_k = (s_1, s_2, \ldots, s_k)$, we write

$$E_{\vec{s}_k} = \bigcap_{i=1}^{k-1} \{s_i \longrightarrow s_{i+1}\}. \tag{2.1}$$

For a configuration of $G_m(t)$, we let $\mathrm{dist}(s,1)$ denote the unique value of k such that $s = s_1 \longrightarrow s_2 \longrightarrow \cdots \longrightarrow s_{k-1} \longrightarrow s_k = 1$. Then, clearly,

$$\mathrm{diam}(G_m(t)) < 2 \max_{s=1,\ldots,t} \mathrm{dist}(s,1), \tag{2.2}$$

because the distance between any two vertices is smaller than the right side of (2.2). We will show that there exists a constant C' such that for each $1 \leq s \leq t$, and with $\varepsilon > 0$,

$$\mathbb{P}(\mathrm{dist}(s,1) > C' \log s) \leq s^{-(1+\varepsilon)}. \tag{2.3}$$

Using first (2.2) and consecutively that $\mathrm{dist}(s,1) \leq s$, the result (2.3) implies Theorem 1.3 with $C = 2C'$, because

$$\begin{aligned}
\mathbb{P}(\mathrm{diam}(G_m(t)) > C\log t) &\leq \mathbb{P}\left(\max_{1\leq s\leq t} \mathrm{dist}(s,1) > C'\log t\right) \\
&= \mathbb{P}\left(\max_{C'\log t \leq s\leq t} \mathrm{dist}(s,1) > C'\log t\right) \leq \mathbb{P}\left(\max_{C'\log t \leq s\leq t} \mathrm{dist}(s,1) > C'\log s\right) \\
&\leq \sum_{C'\log t\leq s\leq t} \mathbb{P}\left(\mathrm{dist}(s,1) > C'\log s\right) \leq \sum_{C'\log t\leq s\leq t} s^{-(1+\varepsilon)} = O\left((\log t)^{-\varepsilon}\right) = o(1).
\end{aligned} \tag{2.4}$$

To see (2.3), we note from Boole's inequality, that

$$\mathbb{P}(\mathrm{dist}(s,1)>k) \quad \le \quad \sum_{l>k}\sum_{\vec{s}_l}\mathbb{P}\Big(\bigcap_{i=1}^{l-1}\{s_i\longrightarrow s_{i+1}\}\Big), \tag{2.5}$$

where, the sum is over all *ordered* vectors $\vec{s}_l$ of length l, for which $s_1=s$ and $s_l=1$. We claim that

$$\mathbb{P}\Big(\bigcap_{i=1}^{l-1}\{s_i\longrightarrow s_{i+1}\}\Big)=\prod_{i=1}^{l-1}\mathbb{P}(s_i\longrightarrow s_{i+1}). \tag{2.6}$$

We prove the independence by induction. For simplicity, we assume that we are in model (c), the analysis in model (b) is quite similar, and will be completed later. First note that

$$\begin{aligned}\mathbb{P}\Big(\bigcap_{i=1}^{l-1}\{s_i\longrightarrow s_{i+1}\}\Big) \quad &= \quad \mathbb{E}\Big[\mathbb{P}\Big(\bigcap_{i=1}^{l-1}\{s_i\longrightarrow s_{i+1}\}\Big|G_m(s_1-1)\Big)\Big]\\ &= \quad \mathbb{E}\Big[I\big[\bigcap_{i=2}^{l-1}\{s_i\longrightarrow s_{i+1}\}\big]\mathbb{P}(s_1\longrightarrow s_2|G_m(s_1-1))\Big],\end{aligned} \tag{2.7}$$

since the event $\bigcap_{i=2}^{l-1}\{s_i\longrightarrow s_{i+1}\}$ is measurable with respect to $G_m(s_1-1)$, and where we write $I[A]$ to denote the indicator of the event A. Furthermore, from (1.3),

$$\mathbb{P}(s_1\longrightarrow s_2|G_m(s_1-1))=\frac{d_{s_2}(s_1-1)+\delta}{(2m+\delta)(s_1-1)}. \tag{2.8}$$

In particular, we have that

$$\mathbb{P}(s_1\longrightarrow s_2)=\mathbb{E}\Big[\frac{d_{s_2}(s_1-1)+\delta}{(2m+\delta)(s_1-1)}\Big]. \tag{2.9}$$

Therefore,

$$\begin{aligned}\mathbb{P}\Big(\bigcap_{i=1}^{l-1}\{s_i\longrightarrow s_{i+1}\}\Big) \quad &= \quad \mathbb{E}\Big[I\big[\bigcap_{i=2}^{l-1}\{s_i\longrightarrow s_{i+1}\}\big]\frac{d_{s_2}(s_1-1)+\delta}{(2m+\delta)(s_1-1)}\Big]\\ &= \quad \mathbb{P}\Big(\bigcap_{i=2}^{l-1}\{s_i\longrightarrow s_{i+1}\}\Big)\mathbb{E}\Big[\frac{d_{s_2}(s_1-1)+\delta}{(2m+\delta)(s_1-1)}\Big],\end{aligned} \tag{2.10}$$

since the random variable $d_{s_2}(s_1-1)$ only depends on how many edges are connected to s_2 after time s_2, which is independent of the event $\bigcap_{i=2}^{l-1}\{s_i\longrightarrow s_{i+1}\}$, which depends on the attachment of the edges *up to and including* time s_2 only. We conclude that

$$\mathbb{P}\Big(\bigcap_{i=1}^{l-1}\{s_i\longrightarrow s_{i+1}\}\Big)=\mathbb{P}(\{s_1\longrightarrow s_2\})\mathbb{P}\Big(\bigcap_{i=2}^{l-1}\{s_i\longrightarrow s_{i+1}\}\Big). \tag{2.11}$$

Iteration leads to the independence claim in (2.6).

In model (b), we have $\{s_1\longrightarrow s_2\}$ if and only if vertex $m(s_1-1)+1$ is connected to one of the vertices $m(s_2-1)+1,\ldots,ms_2$, and by (1.2) this probability equals,

$$\frac{\sum_{k=m(s_2-1)+1}^{ms_2}d_k(m(s_1-1))+\delta}{(2+\delta)(m(s_1-1))}.$$

We observe that this probability depends only on the attachment in $\{G_1(t)\}$ of edges after time ms_2 and is independent of the event $\bigcap_{i=2}^{l-1}\{s_i\longrightarrow s_{i+1}\}$, which depends on the attachment in $\{G_1(t)\}$ *up to and including* time ms_2 only.

We proceed with the proof by investigating $\mathbb{P}(s_i\longrightarrow s_{i+1})$ in the following lemma:

Lemma 2.1. *For model (c) we have for all $s > t \geq 1$, and with $a = \frac{m}{2m+\delta} \in (0,1)$,*

$$\mathbb{P}(s \longrightarrow t) \leq 2^a t^{-a}(s-1)^{a-1}. \tag{2.12}$$

Similarly, for model (b), and again with $a = \frac{m}{2m+\delta} =\in (0,1)$,

$$\mathbb{P}(s \longrightarrow t) \leq 2^a \left(t-1+1/m\right)^{-a} (s-1)^{a-1}. \tag{2.13}$$

Proof. We start with the proof for model (c), the proof for model (b) follows at the end. By definition $\mathbb{P}(2 \longrightarrow 1) = 1$, which is bounded by the right side of (2.12). For $s > 2$, we note from (2.9) that

$$\mathbb{P}(s \longrightarrow t) \quad = \quad \mathbb{E}\Big[\frac{d_t(s-1)+\delta}{(2m+\delta)(s-1)}\Big]. \tag{2.14}$$

Since for $s > 2$, and conditionally on $G_m(s-1)$,

$$X_{s,t} = d_t(s) - d_t(s-1),$$

has a binomial distribution with parameters m and success probability $(d_t(s-1)+\delta)/((s-1)(2m+\delta))$, we find

$$\mathbb{E}[d_t(s)+\delta] = \mathbb{E}[d_t(s-1)+\delta]\Big(1+\frac{m}{(2m+\delta)(s-1)}\Big). \tag{2.15}$$

We now prove by induction on s that

$$\mathbb{E}[d_t(s)+\delta] \leq 2^a(2m+\delta)\left(\frac{s-1}{t}\right)^a. \tag{2.16}$$

For $s = t$, the left side is at most $2m+\delta$, so that we can start the induction. From (2.15) and the induction hypothesis, we obtain

$$\mathbb{E}[d_t(s+1)+\delta] \leq 2^a(2m+\delta)\left(\frac{s-1}{t}\right)^a\left(1+\frac{a}{s}\right) = 2^a(2m+\delta)\left(\frac{s}{t}\right)^a\left(1-\frac{1}{s}\right)^a\left(1+\frac{a}{s}\right). \tag{2.17}$$

Thus, it suffices to prove that for all $x = s^{-1} \in [0,1]$ and $a \in (0,1)$, we have that

$$(1-x)^a(1+ax) \leq 1.$$

The proof of this elementary inequality is left to the reader as an exercise. As a result of (2.14) and (2.16), we obtain (2.12) for all $s > t \geq 1$,

We now turn to a proof of the lemma for model (b). We start with $m = 1$ and note from (1.2) that

$$\mathbb{P}(s \longrightarrow t) \quad = \quad \mathbb{E}\Big[\frac{d_t(s-1)+\delta}{(2+\delta)(s-1)}\Big]. \tag{2.18}$$

and so

$$\mathbb{E}[d_t(s)+\delta] = \mathbb{E}[d_t(s-1)+\delta]\Big(1+\frac{1}{(2+\delta)(s-1)}\Big). \tag{2.19}$$

This yields the inequality

$$\mathbb{E}[d_t(s)+\delta] \leq 2^a(2+\delta)\left(\frac{s-1}{t}\right)^a, \quad s \geq 2, \tag{2.20}$$

where $m = 1$ and $a = \frac{1}{2+\delta}$. Hence (2.12) is valid in model (b), for $m = 1$. For $m > 1$, we have $\{s \longrightarrow t\}$ if and only if vertex $m(s-1)+1$ is connected to one of the vertices $m(t-1)+1, \ldots, mt$ in $\{G_1(t)\}$, so that with $a' = 1/(2+\delta')$, and $\delta' = \delta/m$,

$$\mathbb{P}(s \longrightarrow t) \leq 2^{a'} \sum_{j=m(t-1)+1}^{mt} j^{-a'} \left(m(s-1)\right)^{a'-1} \leq 2^{a'} (s-1)^{a'-1} (t-1+1/m)^{-a'}. \tag{2.21}$$

This gives the result (2.13) because $a' = 1/(2+\delta') = m/(2m+\delta)$ for $\delta' = \delta/m$. □

We now finish the proof of (2.3) for model (c). The proof for model (b) proceeds similarly, and is omitted. Consider $\vec{s}_l$ with $s_1 = s$ and $s_l = 1$. We obtain from (2.12), using at the end $s_l = 1$

$$\begin{aligned}\mathbb{P}\Big(\bigcap_{i=1}^{l-1}\{s_i \longrightarrow s_{i+1}\}\Big) &= \prod_{i=1}^{l-1} \mathbb{P}(s_i \longrightarrow s_{i+1}) \leq \prod_{i=1}^{l-1} 2^a s_{i+1}^{-a}(s_i-1)^{a-1} \\ &= \prod_{i=1}^{l-1} 2^a \left(\frac{s_i-1}{s_{i+1}}\right)^a \prod_{i=1}^{l-1} \frac{1}{s_i-1} \leq (s_1-1)^{a-1}\, 2^{a(l-1)} \prod_{i=2}^{l-1} \frac{1}{s_i-1}.\end{aligned} \tag{2.22}$$

Therefore, we arrive from $s_1 = s$ and $s_l = 1$ at

$$\mathbb{P}(\mathrm{dist}(s,1) > k) \quad \leq \quad s^{a-1} \sum_{l>k} 2^{a(l-1)} \sum_{\vec{s}_l} \prod_{i=2}^{l-1} \frac{1}{s_i-1}. \tag{2.23}$$

We recall that the sum over $\vec{s}_l$ is over *ordered* vectors $s_2, \ldots, s_{l-1}$. When we turn the sum into a sum over vectors $\vec{s}_l$ with $s_1 = s$ and $s_l = 1$ with only *distinct* coordinates, we need to divide by a factor $(l-2)!$. Denoting the *unordered* vector by $\vec{t}_l$, we then obtain

$$\mathbb{P}(\mathrm{dist}(s,1) > k) \quad \leq \quad 2^{1-a} s^{a-1} \sum_{l>k} \frac{2^{a(l-1)}}{(l-2)!} \sum_{\vec{t}_l} \prod_{i=2}^{l-1} \frac{1}{t_i-1}. \tag{2.24}$$

If we were to sum over all vectors $\vec{t}_l$, rather than the vectors with different coordinates, we would obtain

$$\sum_{\vec{t}_l} \prod_{i=2}^{l-1} \frac{1}{t_i-1} = \Big(\sum_{u=1}^{s-1} \frac{1}{u}\Big)^{l-2}. \tag{2.25}$$

Clearly, this is an upper bound on the sum, so that

$$\mathbb{P}(\mathrm{dist}(s,1) > k) \quad \leq \quad 2^{1-a} s^{a-1} \sum_{l>k} \frac{2^{a(l-1)}}{(l-2)!} \Big(\sum_{u=1}^{s-1} \frac{1}{u}\Big)^{l-2}. \tag{2.26}$$

Using that $\sum_{u=1}^{s-1} \frac{1}{u} \leq 1 + \log s$, we arrive at

$$\mathbb{P}(\mathrm{dist}(s,1) > k) \quad \leq \quad 2s^{a-1} \sum_{l>k} \frac{\left(2^a(1+\log s)\right)^{l-2}}{(l-2)!} \leq c s^{2^a+a-1} \mathbb{P}(X \geq k-1), \tag{2.27}$$

where $c = 2\exp(2^a)$, and X is a Poisson random variable with mean $2^a(1+\log s)$.

By [29, (2.5) and Remark 2.6], for any Poisson random variable Y with mean λ, we have

$$\mathbb{P}(Y > 5\lambda) = \mathbb{P}(Y > \mathbb{E}[Y] + 4\lambda) \leq \exp\left(-\frac{8\lambda^2}{(\lambda + 4\lambda/3)}\right) = e^{-24\lambda/7}. \tag{2.28}$$

Switching back to (2.27), taking $\lambda = 2^a(1+\log s)$, we obtain that

$$\mathbb{P}(\mathrm{dist}(s,1) > 5\cdot 2^a(1+\log s)+1) \le cs^{1+a}\mathbb{P}(X \ge 5\lambda) \le cs^{1+a}\exp\{-24\lambda/7\} \le cs^{-(1+\varepsilon)}, \quad (2.29)$$

where we can take $\varepsilon = 3/7$ and where we used that $0 < a < 1$. This completes the proof. □

Remark. It is immediate from (1.1) and (1.2) that the upper bound (2.13) of the lemma also holds for model (a).

Remark. When $m = 1$, we see that $\mathrm{dist}(s,1)$ is *equal* to the graph distance between vertex 1 and s. This can be used to prove a $\log t$ lower bound on the diameter of $G_1(t)$ in the case of models (b) and (c). We refrain from working this argument out, as sharper results follow from [32].

3 A $\log$ lower bound on the diameter for $\delta \ge 0$: Proof of Theorem 1.4

In this section, we prove Theorem 1.4 by extending the proof in [13] from $\delta = 0$ to $\delta \ge 0$.

Denote in model (c) by

$$\{g(t,j) = s\}, \quad 1 \le j \le m, \quad (3.1)$$

the event that at time t the j^{th} edge of vertex t is attached to the earlier vertex s. For the models (a) and (b), this event means that in $\{G_1(mt)\}$ the edge from vertex $m(t-1)+j$ is attached to one of the vertices $m(s-1)+1,\ldots,ms$. It is a direct consequence of the definition of PA-models that the event (3.1) increases the preference for vertex s, and hence decreases (in a relative way) the preference for the vertices u, $1 \le u \le t$, $u \ne s$. It should be intuitively clear that another way of expressing this effect is to say that, for different $s_1 \ne s_2$, the events $\{g(t_1,j_1) = s_1\}$ and $\{g(t_2,j_2) = s_2\}$ are negatively correlated. In order to state such a result, we introduce some notation. For integer $n_s \ge 1$, we denote by

$$E_s = \bigcap_{i=1}^{n_s} \{g(t_i^{(s)}, j_i^{(s)}) = s\}, \quad (3.2)$$

the event that at time t_i the j_i^{th} edge of vertex t_i is attached to the earlier vertex s. We will start by proving that for each $k \ge 1$ and all possible choices of $t_i^{(s)}, j_i^{(s)}$, the events E_s, for different s, are negatively correlated:

Lemma 3.1. *For distinct $s_1, s_2, \ldots, s_k$,*

$$\mathbb{P}\Big(\bigcap_{i=1}^{k} E_{s_i}\Big) \le \prod_{i=1}^{k} \mathbb{P}(E_{s_i}). \quad (3.3)$$

Proof. We will use induction on the largest edge number present in the events E_s. Here, for an event $\{g(t,j) = s\}$, we let the edge number be $m(t-1)+j$, which is the order of the edge when we consider the edges as being attached in sequence. The induction hypothesis is that (3.3) holds for all k and all choices of $t_i^{(s)}, j_i^{(s)}$ such that $\max_{i,s} m(t_i^{(s)}-1)+j_i^{(s)} \le e$, where induction is performed with respect to e. We now complete the induction argument. To initialize the induction, we note that for $e = 1$, the induction hypothesis holds trivially, since $\bigcap_{i=1}^{k} E_{s_i}$ can be empty or consist of exactly one event, and in the latter case there is nothing to prove. This initializes the induction.

To advance the induction, we assume that (3.3) holds for all k and all choices of $t_i^{(s)}, j_i^{(s)}$ such that $\max_{i,s} m(t_i^{(s)}-1)+j_i^{(s)} \le e-1$, and we extend it to all k and all choices of $t_i^{(s)}, j_i^{(s)}$ such that $\max_{i,s} m(t_i^{(s)}-1)+j_i^{(s)} \le e$. Clearly, for k and $t_i^{(s)}, j_i^{(s)}$ such that $\max_{i,s} m(t_i^{(s)}-1)+j_i^{(s)} \le e-1$, the bound follows from the induction hypothesis, so we may restrict attention to the case that $\max_{i,s} m(t_i^{(s)}-1)+j_i^{(s)} = e$. We note that there is a unique choice of t,j such that $m(t-1)+j = e$. In this case, there are again two possibilities. Either there is exactly one choice of s and $t_i^{(s)}, j_i^{(s)}$ such that $t_i^{(s)} = t, j_i^{(s)} = j$, or there are at least *two* of such choices. In the latter case, we immediately have that $\bigcap_{s=1}^k E_s = \varnothing$, since the e^{th} edge can only be connected to a *unique* vertex. Hence, there is nothing to prove. Thus, we are left to investigate the case where there exists unique s and $t_i^{(s)}, j_i^{(s)}$ such that $t_i^{(s)} = t, j_i^{(s)} = j$. Denote by

$$E_s' = \bigcap_{i=1:(t_i^{(s)}, j_i^{(s)}) \neq (t,j)}^{n_s} \{g(t_i^{(s)}, j_i^{(s)}) = s\}, \tag{3.4}$$

the restriction of E_s to the *other* edges. Then we can write

$$\bigcap_{i=1}^k E_{s_i} = \{g(t,j) = s\} \cap E_s' \cap \bigcap_{i=1:s_i \neq s}^k E_{s_i}. \tag{3.5}$$

By construction, all the edge numbers of the events in $E_s' \cap \bigcap_{i=1:s_i \neq s}^k E_{s_i}$ are at most $e-1$. Thus, we obtain

$$\mathbb{P}\Big(\bigcap_{i=1}^k E_{s_i}\Big) \leq \mathbb{E}\Big[I[E_s' \cap \bigcap_{i=1:s_i \neq s}^k E_{s_i}]\mathbb{P}_{e-1}(g(t,j) = s)\Big], \tag{3.6}$$

where $\mathbb{P}_{e-1}$ denotes the conditional probability given the edge attachments up to the $(e-1)^{\text{st}}$ edge connection.

We now first go to model (c), for which we have that

$$\mathbb{P}_{e-1}(g(t,j) = s) = \frac{d_s(t-1) + \delta}{(2m+\delta)(t-1)}. \tag{3.7}$$

We wish to use the induction hypothesis. For this, we note that

$$d_s(t-1) = m + \sum_{(t',j'):t' \leq t-1} I[g(t',j') = s]. \tag{3.8}$$

We note that each of the terms in (3.8) has edge number strictly smaller than e and occurs with a non-negative multiplicative constant. As a result, we may use the induction hypothesis for each of these terms. Thus, we obtain, using also $m + \delta \geq 0$, that,

$$\begin{aligned}(2m+\delta)(t-1)\mathbb{P}\Big(\bigcap_{i=1}^k E_{s_i}\Big) &\leq (m+\delta)\mathbb{P}(E_s') \prod_{i=1:s_i \neq s}^k \mathbb{P}(E_{s_i}) \\ &+ \sum_{(t',j'):t' \leq t-1} \mathbb{P}(E_s' \cap \{g(t',j') = s\}) \prod_{i=1:s_i \neq s}^k \mathbb{P}(E_{s_i}).\end{aligned} \tag{3.9}$$

We can recombine to obtain

$$\mathbb{P}\Big(\bigcap_{i=1}^k E_{s_i}\Big) \leq \mathbb{E}\Big[I[E_s'] \frac{d_s(t-1)+\delta}{(2m+\delta)(t-1)}\Big] \prod_{i=1:s_i \neq s}^k \mathbb{P}(E_{s_i}), \tag{3.10}$$

and the advancement is completed when we note that

$$\mathbb{E}\Big[I[E_s'] \frac{d_s(t-1)+\delta}{(2m+\delta)(t-1)}\Big] = \mathbb{P}(E_s). \tag{3.11}$$

The proofs for models (a) and (b) are somewhat simpler, since the events E_{s_i} can be reformulated in terms of the graph process $\{G_1(t)\}_{t \geq 1}$. □

We next study the probabilities of E_s when $n_s \leq 2$:

Lemma 3.2. *There exist absolute constants M_1, M_2, such that (i) for each $1 \leq j \leq m$, and $t > s$,*

$$\mathbb{P}(g(t,j) = s) \leq \frac{M_1}{t^{1-a}s^a}, \tag{3.12}$$

and (ii) for $t_2 > t_1 > s$, and any $1 \leq j_1, j_2 \leq m$,

$$\mathbb{P}\Big(g(t_1,j_1) = s, g(t_2,j_2) = s\Big) \leq \frac{M_2}{(t_1t_2)^{1-a}s^{2a}}, \tag{3.13}$$

where, as before, $a = \frac{m}{2m+\delta}$.

Proof. For model (c), we find from (2.12),

$$\mathbb{P}(g(t,j) = s) = \mathbb{P}(g(t,1) = s) = \mathbb{P}(t \longrightarrow s) \leq 2^a s^{-a}(t-1)^{a-1} \leq M_1 s^{-a} t^{a-1}, \tag{3.14}$$

provided that $M_1 \geq 2^a \sup_{t\geq 2}(\frac{t}{t-1})^{1-a} = 2$. For models (a) and (b), we need

$$2^a\Big(s-1+\frac{1}{m}\Big)^{-a}\Big(t-1+\frac{j-1}{m}\Big)^{a-1} \leq M_1 s^{-a} t^{a-1},$$

which holds when we choose $M_1 \geq 2m^a$.

We proceed with the proof of (3.13). We show (3.13) for model (c), the proof for the other models being similar. For some constant M_3,

$$\begin{aligned}
\mathbb{P}\Big(g(t_1,j_1) = s, g(t_2,j_2) = s\Big) &= \mathbb{E}\Big[\mathbb{P}(g(t_1,j_1) = s, g(t_2,j_2) = s | G_m(t_2-1))\Big] \\
&= \mathbb{E}\left[I[g(t_1,j_1) = s]\left(\frac{d_s(t_2-1)+\delta}{(t_2-1)(2m+\delta)}\right)\right] \\
&= \frac{1}{(t_2-1)(2m+\delta)}\frac{\Gamma(t_2-1+a)\Gamma(t_1)}{\Gamma(t_2-1)\Gamma(t_1+a)}\mathbb{E}\left[I[g(t_1,j_1) = s]\left(d_s(t_1)+\delta\right)\right] \\
&\leq \frac{M_3}{t_1^a t_2^{1-a}}\mathbb{E}\left[I[g(t_1,j_1) = s]\left(d_s(t_1)+\delta\right)\right],
\end{aligned} \tag{3.15}$$

using for $t_1 < u \leq t_2 - 1$, the iteration (compare (2.15)),

$$\mathbb{E}\left[I[g(t_1,j_1) = s]\left(d_s(u)+\delta\right)\right] = \Big(1+\frac{a}{u-1}\Big)\mathbb{E}\left[I[g(t_1,j_1) = s]\left(d_s(u-1)+\delta\right)\right],$$

and the asymptotic identities $\frac{\Gamma(t_1)}{\Gamma(t_1+a)} \sim t_1^{-a}$, $\frac{\Gamma(t_2-1+a)}{(t_2-1)\Gamma(t_2-1)} \sim t_2^{a-1}$.

We are lead to compute $\mathbb{E}\left[I[g(t_1,j_1) = s]\left(d_s(t_1)+\delta\right)\right]$. We do so by recursion:

$$\begin{aligned}
&\mathbb{E}\left[I[g(t_1,j_1) = s]\left(d_s(t_1)+\delta\right)|G_m(t_1-1)\right] \\
&\quad= \mathbb{E}\left[I[g(t_1,j_1) = s]\left(d_s(t_1)-d_s(t_1-1)\right)|G_m(t_1-1)\right] + \mathbb{E}\left[I[g(t_1,j_1) = s]\left(d_s(t_1-1)+\delta\right)|G_m(t_1-1)\right] \\
&\quad= \sum_{j=1}^m \mathbb{E}\left[I[g(t_1,j_1) = s]I[g(t_1,j) = s]\big|G_m(t_1-1)\right] + \left(d_s(t_1-1)+\delta\right)\mathbb{E}\left[I[g(t_1,j_1) = s]\big|G_m(t_1-1)\right] \\
&\quad= \frac{d_s(t_1-1)+\delta}{(t_1-1)(2m+\delta)} + (m-1)\left(\frac{d_s(t_1-1)+\delta}{(t_1-1)(2m+\delta)}\right)^2 + \frac{(d_s(t_1-1)+\delta)^2}{(t_1-1)(2m+\delta)} \\
&\quad\leq \frac{d_s(t_1-1)+\delta}{(t_1-1)(2m+\delta)} + m\frac{(d_s(t_1-1)+\delta)^2}{(t_1-1)(2m+\delta)}.
\end{aligned} \tag{3.16}$$

We prove in Lemma 5.1 of the appendix that, for some constant M_4,

$$\mathbb{E}[(d_s(t)+\delta)^2] \leq M_4(t/s)^{2a}. \tag{3.17}$$

Combining (3.15), (3.16), (3.17) and (2.16), we find, possibly after enlarging M_4,

$$\mathbb{P}(g(t_1,j_1)=s, g(t_2,j_2)=s) \leq \left(\frac{M_3}{t_1^a t_2^{1-a}}\right)\left[\frac{mM_4}{(2m+\delta)(t_1-1)}\Big(\frac{t_1-1}{s}\Big)^{2a}\right] \leq \frac{M_2}{(t_1t_2)^{1-a}s^{2a}}, \tag{3.18}$$

for some M_2. □

We combine the results of Lemmas 3.1 and 3.2 into the following corollary, yielding an upper bound for the probability of the existence of a path. In its statement, we call a path $\Gamma=(s_0,s_1,\ldots,s_l)$ self-avoiding when $s_i \neq s_j$ for all $1 \leq i < j \leq l$.

Corollary 3.3. *Let $\Gamma=(s_0,s_1,\ldots,s_l)$ be a self-avoiding path of length l consisting of the $l+1$ unordered vertices $s_0,s_1,\ldots,s_l$, then*

$$\mathbb{P}(\Gamma \in G_m(t)) \leq (m^2C)^l \prod_{i=0}^{l-1} \frac{1}{(s_i \wedge s_{i+1})^a (s_i \vee s_{i+1})^{1-a}}. \tag{3.19}$$

Proof. Since Γ is self-avoiding, we can write $\{\Gamma \in G\} = \cap_{i=1}^k E_{s_i}$, where either

$$E_s = \{g(t,j)=s\} \tag{3.20}$$

for some $t > s$ and some $1 \leq j \leq m$, or

$$E_s = \{g(t_1,j_1)=g(t_2,j_2)=s\}, \tag{3.21}$$

for some $t_1,t_2 > s$ and some $1 \leq j_1, j_2 \leq m$. In the first case we have according to (3.12),

$$\mathbb{P}(E_s) = \mathbb{P}(g(t,j)=s) \leq \frac{M_1}{t^{1-a}s^a}, \tag{3.22}$$

whereas in the second case, according to (3.13),

$$\mathbb{P}(E_s) = \mathbb{P}(g(t_1,j_1)=s, g(t_2,j_2)=s) \leq \frac{M_2}{(t_1t_2)^{1-a}s^{2a}} = \frac{M_2}{t_1^{1-a}s^a t_2^{1-a}s^a}. \tag{3.23}$$

In both cases M_i, $i=1,2$, is an absolute constant. Lemma 3.1 then yields (3.19), where the factor m^{2l} originates from the choices of $j \in \{1,2,\ldots,m\}$. □

3.1 Application to the lower bound on the diameter when $\delta \geq 0$

Observe that for integers $j > i$ we have

$$\left(\frac{j}{i}\right)^{a-\frac{1}{2}} \leq 1, \tag{3.24}$$

if and only if $a \leq \frac{1}{2}$, which happens if and only if $\delta \geq 0$, because $a = \frac{m}{2m+\delta}$. Hence, $j^{a-1}i^{-a} \leq (ij)^{-1/2}$ precisely when $\delta \geq 0$. It now follows from (3.19) and the above inequality that for $\delta \geq 0$,

$$\mathbb{P}(\Gamma \in G_m(t)) \leq (m^2C)^l \prod_{i=0}^{l-1} \frac{1}{\sqrt{s_is_{i+1}}}. \tag{3.25}$$

The further proof that (3.25) implies that for $\delta \geq 0$,

$$L = \frac{\log(t-1)}{\log(3Cm^2 \log t)}, \tag{3.26}$$

is a lower bound for the diameter of $G_m(t)$, is identical to the proof of [13, Theorem 5, p. 14], with n replaced by t. □

4 A $\log\log$ upper bound on the diameter: Proof of Theorem 1.6

The proof of Theorem 1.6 is divided into two key steps. In the first, in Theorem 4.1, we give a bound on the diameter of the *core* which consists of the vertices with degree at least a certain power of $\log t$. This argument is close in spirit to the argument in [33] used to prove bounds on the average distance for the configuration model, but substantial adaptations are necessary to deal with preferential attachment. After this, in Theorem 4.7, we derive a bound on the distance between vertices with a small degree and the core. We start by defining and investigating the core of the preferential attachment model. In the sequel, it will be convenient to prove Theorem 1.6 for $2t$ rather than for t. Clearly, this does not make any difference for the results.

4.1 The diameter of the core

We recall that

$$\tau = 3 + \frac{\delta}{m}, \tag{4.1}$$

so that $-m < \delta < 0$ corresponds to $\tau \in (2,3)$. Throughout this section, we fix $m \geq 2$.

We take $\sigma > \frac{1}{3-\tau} = -\frac{m}{\delta} > 1$ and define the *core* Core_t of the PA-model $G_m(2t)$ to be

$$\mathrm{Core}_t = \big\{ i \in \{1, 2, \ldots, t\} : d_i(t) \geq (\log t)^{\sigma} \big\}, \tag{4.2}$$

i.e., all the vertices which at time t have degree at least $(\log t)^{\sigma}$.

For a graph G with vertex set $\{1, 2, \ldots, t\}$ and a given edge set, we write $d_G(i,j)$ for the shortest-path distance between i and j in the graph G. Also, for $A \subseteq \{1, 2, \ldots, t\}$, we write

$$\mathrm{diam}_t(A) = \max_{i,j \in A} d_{G_m(t)}(i,j). \tag{4.3}$$

Then, the diameter of the core in the graph $G_m(2t)$, which we denote by $\mathrm{diam}_{2t}(\mathrm{Core}_t)$, is bounded in the following theorem:

Theorem 4.1 (The diameter of the core)**.** *Fix $m \geq 2$. For every $\sigma > \frac{1}{3-\tau}$,* **whp**,

$$\mathrm{diam}_{2t}(\mathrm{Core}_t) \leq (1 + o(1)) \frac{4 \log\log t}{|\log(\tau - 2)|}. \tag{4.4}$$

The proof of Theorem 4.1 is divided into several smaller steps. We start by proving that the diameter of the *inner core* Inner_t, which is defined by

$$\mathrm{Inner}_t = \big\{ i \in \{1, 2, \ldots, t\} : d_i(t) \geq t^{\frac{1}{2(\tau-1)}} (\log t)^{-\frac{1}{2}} \big\}, \tag{4.5}$$

is, **whp**, at most 10. After this, we will show that the distance from the *outer core*, which is defined to be equal to $\mathrm{Outer}_t = \mathrm{Core}_t \backslash \mathrm{Inner}_t$, to the inner core can be bounded by a fixed constant times $\log\log t$. This also shows that the diameter of the outer core is bounded by a different constant times $\log\log t$. We now give the details.

Proposition 4.2 (The diameter of the inner core)**.** *Fix $m \geq 2$, then* **whp***,*

$$\mathrm{diam}_{2t}(\mathrm{Inner}_t) \leq 10. \tag{4.6}$$

Proof. We first introduce the important notion of a t-connector between a vertex $i \in \{1, 2, \dots, t\}$ and a set of vertices $A \subseteq \{1, 2, \dots, t\}$, which plays a crucial role throughout the proof. Fix a set of vertices A and a vertex i. We say that the vertex $j \in \{t+1, t+2, \dots, 2t\}$ is a *t-connector* between i and A if one of the edges incident to j connects to i and another edge incident to j connects to a vertex in A. Thus, when there exists a t-connector between i and A, the distance between i and A in $G_m(2t)$ is at most 2.

We continue the analysis by first considering model (c). We note that for a set of vertices A and a vertex i with degree at time t equal to $d_i(t)$, we have that, conditionally on $G_m(t)$, the probability that $j \in \{t+1, t+2, \dots, 2t\}$ is a t-connector for i and A is at least

$$\frac{(d_A(t) + \delta|A|)(d_i(t) + \delta)}{[2t(2m + \delta)]^2}, \tag{4.7}$$

independently of the fact whether the other vertices are t-connectors or not, and where, for any $A \subseteq \{1, 2, \dots, t\}$, we write

$$d_A(t) = \sum_{i \in A} d_i(t). \tag{4.8}$$

Since $d_i(t) + \delta \geq m + \delta > 0$ for every $i \leq t$, and $\delta < 0$, we have that

$$d_i(t) + \delta = d_i(t)\Big(1 + \frac{\delta}{d_i(t)}\Big) \geq d_i(t)(1 + \frac{\delta}{m}) = d_i(t)\frac{m + \delta}{m}, \tag{4.9}$$

and, thus, also $d_A(t) + \delta|A| \geq d_A(t)\frac{m+\delta}{m}$. As a result, for $\eta = (m+\delta)^2/(2m(2m+\delta))^2 > 0$, the probability that $j \in \{t+1, t+2, \dots, 2t\}$ is a t-connector for i and A is at least $\frac{\eta d_A(t) d_i(t)}{t^2}$, independently of the fact whether the other vertices are t-connectors or not. Therefore, the probability that there is no t-connector for i and A is, conditionally on $G_m(t)$, bounded above by

$$\Big(1 - \frac{\eta d_A(t) d_i(t)}{t^2}\Big)^t \leq \exp\Big\{-\frac{\eta d_A(t) d_i(t)}{t}\Big\}. \tag{4.10}$$

We shall make use of (4.10) in several places throughout the proof.

For model (a) and (b), the right hand side of (4.10) also serves as an upper bound on the probability of non-existence of a t-connector between A and i, conditionally on $G_m(t)$, with trivial adaptations, and a slightly different value of $\eta > 0$.

In the course of the proof we will make use of the following lemma:

Lemma 4.3 (The maximal degree)**.** *Fix $m \geq 1$. With high probability, for $\delta \leq 0$,*

$$\max_{i \leq \log t} d_i(t) \geq t^{\frac{1}{\tau-1}} (\log t)^{-1}. \tag{4.11}$$

We defer the proof of Lemma 4.3 to Section 5.3 of the appendix.

We start by proving the claim of Proposition 4.2 for $\tau \in (\frac{5}{2}, 3)$. Let $i^* \leq \log t$ be such that $d_{i^*}(t) = \max_{i \leq \log t} d_i(t)$. Further, for $i \in \mathrm{Inner}_t$, we have that $d_i(t) \geq t^{\frac{1}{2(\tau-1)}} (\log t)^{-\frac{1}{2}}$. Thus, by (4.10), the probability that there is no t-connector between $i \in \mathrm{Inner}_t$ and i^* is bounded by

$$\exp\{-\frac{\eta d_{i^*}(t) d_i(t)}{t}\} \leq \exp\Big\{-\eta \frac{t^{\frac{3}{2(\tau-1)}}}{t (\log t)^{3/2}}\Big\}, \tag{4.12}$$

which converges to 0 since, for $\tau > 5/2$, we have $\frac{3}{2(\tau-1)} > 1$. Thus, the distance between any $i \in \text{Inner}_t$ and i^* is, **whp**, bounded by 2. This implies that, **whp**, $\text{diam}_{2t}(\text{Inner}_t) \leq 4$.

We next extend the result to $\tau \in (2, \frac{5}{2}]$. Observe from Lemma 5.2 of the appendix that, **whp**, Inner_t contains at least $\sqrt{t}$ vertices and denote the first $\sqrt{t}$ vertices of Inner_t by I. Observe that for $\tau > 2$ we have $t^{(\tau-1)^{-1}-1} \downarrow 0$ so that, for any $i, j \in I$, the probability that there exists a t-connector for i and j is bounded below by,

$$1 - \exp\{-\eta t^{\frac{1}{\tau-1}-1}(\log t)^{-1}\} \geq p_t \equiv t^{\frac{1}{\tau-1}-1}(\log t)^{-2}, \tag{4.13}$$

for t sufficiently large.

We wish to couple Inner_t with an Erdős-Rényi random graph with $n_t = \sqrt{t}$ vertices and edge probability p_t, which we denote by $G(n_t, p_t)$. For this, for $i, j \in \{1, 2, \ldots, n_t\}$, we say that an edge between i and j is present when there exists a t-connector connecting the i^{th} and j^{th} vertex in I. We now prove that this graph is bounded below by $G(n_t, p_t)$. Note that (4.13) does not guarantee this coupling, instead we should prove that the lower bound holds uniformly, when i and j belong to I.

For this, we order the $n_t(n_t-1)/2$ edges in an arbitrary way, and bound the conditional probability that the l^{th} edge is present conditionally on the previous edges from below by p_t, for every l. This would prove the claimed stochastic domination by $G(n_t, p_t)$.

Indeed, the l^{th} edge is present precisely when there exists a t-connector connecting the corresponding vertices which we call i and j in I. Moreover, we shall not make use of the first vertices which were used to t-connect the previous edges. This removes at most $n_t(n_t-1)/2 \leq t/2$ possible t-connectors, after which at least another $t/2$ remain. The probability that one of them is a t-connector for the i^{th} and j^{th} vertex in I is bounded below by, for t sufficiently large,

$$1 - \exp\{-\eta t^{\frac{1}{\tau-1}-2}(\log t)^{-1} t/2\} = 1 - \exp\{-\eta t^{\frac{1}{\tau-1}-1}(\log t)^{-1}/2\} \geq p_t \equiv t^{\frac{1}{\tau-1}-1}(\log t)^{-2}, \tag{4.14}$$

using $1 - e^{-x} \geq x/2$ for $x \in [0, 1]$ and $\eta/2 \geq \log t^{-1}$ for t sufficiently large.

This proves the claimed stochastic domination of the random graph on the vertices I and $G(n_t, p_t)$. Next, we show that $\text{diam}(G(n_t, p_t))$ is, **whp**, bounded by 3. For this we use the result in [7, Corollary 10.12], which gives sharp bounds on the diameter of an Erdős-Rényi random graph. Indeed, this result implies that if $p^d n^{d-1} - 2\log n \to \infty$, while $p^{d-1}n^{d-2} - 2\log n \to -\infty$, then $\text{diam}(G(n,p)) = d$, **whp**. In our case, $n = n_t = t^{1/2}$ and $p = p_t = t^{\frac{1}{\tau-1}-1}(\log t)^{-2}$, which implies that, **whp**, $\frac{\tau-1}{3-\tau} < d \leq \frac{\tau-1}{3-\tau} + 1$. Since, for $\tau \in (2, 5/2]$, we have $\frac{\tau-1}{3-\tau} < 3$, we obtain that the diameter of I in $G_m(2t)$ is **whp** bounded by 6 in this case.

We finally show that for any $i \in \text{Inner}_t \setminus I$, the probability that there does not exist a t-connector connecting i and I is small. Indeed, this probability is, since $d_I(t) \geq \sqrt{t}t^{\frac{1}{\tau-1}}(\log t)^{-1/2}$, and $d_i(t) \geq t^{\frac{1}{2(\tau-1)}}(\log t)^{-1/2}$, the probability of there not existing a t-connector is bounded above by $e^{-\eta t^{1/(\tau-1)-1/2}(\log t)^{-1}}$, which is tiny since $\tau < 3$. This proves that **whp**the distance between any vertex $i \in \text{Inner}_t \setminus I$ and I is bounded by 2, and, together with the fact that $\text{diam}_{2t}(I) \leq 6$ thus implies that $\text{diam}_{2t}(\text{Inner}_t) \leq 10$. □

Proposition 4.4 (The distance between the outer and the inner core)**.** *Fix $m \geq 2$. With high probability, the inner core* Inner_t *can be reached from any vertex in the outer core* Outer_t *using no more than* $\frac{2\log\log t}{|\log(\tau-2)|}$ *edges in $G_m(2t)$. More precisely,* **whp**

$$\max_{i \in \text{Outer}_t} \min_{j \in \text{Inner}_t} d_{G_m(2t)}(i,j) \leq \frac{2\log\log t}{|\log(\tau-2)|}. \tag{4.15}$$

Proof. Recall that

$$\text{Outer}_t = \text{Core}_t \setminus \text{Inner}_t. \tag{4.16}$$

and define

$$\mathcal{N}^{(1)} = \text{Inner}_t = \{i : d_i(t) \geq u_1\}, \tag{4.17}$$

where

$$u_1 = l_t = t^{\frac{1}{2(\tau-1)}}(\log t)^{-\frac{1}{2}}. \tag{4.18}$$

We now recursively define a sequence u_k, for $k \geq 2$, so that for any vertex $i \in \{1, 2, \dots, t\}$ with degree at least u_k, the probability that there is no t-connector for the vertex i and the set

$$\mathcal{N}^{(k-1)} = \{j : d_j(t) \geq u_{k-1}\}, \tag{4.19}$$

conditionally on $G_m(t)$ is tiny. According to (4.10) and (5.6) in the appendix, this probability is at most

$$\exp\Big\{ -\frac{\eta B t [u_{k-1}]^{2-\tau} u_k}{t} \Big\} = o(t^{-1}), \tag{4.20}$$

when we define

$$u_k = D \log t (u_{k-1})^{\tau-2}, \tag{4.21}$$

with D exceeding $(\eta B)^{-1}$. The following lemma identifies u_k:

Lemma 4.5 (Identification of u_k)**.** *For each $k \in \mathbb{N}$,*

$$u_k = D^{a_k} (\log t)^{b_k} t^{c_k}, \tag{4.22}$$

where

$$c_k = \frac{(\tau-2)^{k-1}}{2(\tau-1)}, \qquad b_k = \frac{1-(\tau-2)^{k-1}}{3-\tau} - \frac{1}{2}(\tau-2)^{k-1}, \qquad a_k = \frac{1-(\tau-2)^{k-1}}{3-\tau}. \tag{4.23}$$

Proof. We identify a_k, b_k and c_k recursively. We note that $c_1 = \frac{1}{2(\tau-1)}, b_1 = -\frac{1}{2}, a_1 = 0$. By (4.21), we can relate c_k, b_k and a_k to c_{k-1}, b_{k-1} and a_{k-1} as follows:

$$c_k = (\tau-2)c_{k-1}, \qquad b_k = 1 + (\tau-2)b_{k-1}, \qquad a_k = 1 + (\tau-2)a_{k-1}. \tag{4.24}$$

Iterating (4.24) we obtain the expressions for c_k, b_k and a_k in (4.23). □

Then, the key step in the proof of Proposition 4.4 is the following lemma:

Lemma 4.6 (Connectivity between $\mathcal{N}^{(k-1)}$ and $\mathcal{N}^{(k)}$)**.** *Fix $m, k \geq 2$. Then the probability that there exists an $i \in \mathcal{N}^{(k)}$ that is not at distance two from $\mathcal{N}^{(k-1)}$ in $G_m(2t)$ is $o(t^{-1})$.*

Proof. We note that, by Lemma 5.2, that with probability exceeding $1 - o(t^{-1})$,

$$\sum_{i \in \mathcal{N}^{(k-1)}} d_i(t) \geq Bt[u_{k-1}]^{2-\tau}. \tag{4.25}$$

On the event that the bounds in (4.25) hold, we obtain by (4.10) that the conditional probability, given $G_m(t)$, that there exists an $i \in \mathcal{N}^{(k)}$ such that there is no t-connector between i and $\mathcal{N}^{(k-1)}$ is bounded, using Boole's inequality, by

$$te^{-\eta B [u_{k-1}]^{2-\tau} u_k} = te^{-\eta B D \log t} = o(t^{-1}), \tag{4.26}$$

where we have used (4.21) and we have taken $D > 2(\eta B)^{-1}$. □

We now complete the proof of Proposition 4.4. Fix

$$k^* = \Big\lfloor \frac{\log\log t}{|\log(\tau-2)|} \Big\rfloor. \tag{4.27}$$

As a result of Lemma 4.6, we have that the distance between $\mathcal{N}^{(k^*)}$ and Inner_t is at most $2k^*$. Therefore, we are done when we can show that

$$\mathrm{Outer}_t \subseteq \{i : d_i(t) \geq (\log t)^\sigma\} \subseteq \mathcal{N}^{(k^*)} = \{i : d_i(t) \geq u_{k^*}\}, \tag{4.28}$$

so that it suffices to prove that $(\log t)^\sigma \geq u_{k^*}$, for any $\sigma > \frac{1}{3-\tau}$. For this, we note that, by Lemma 4.5, we have that

$$u_{k^*} = D^{a_{k^*}} (\log t)^{b_{k^*}} t^{c_{k^*}}. \tag{4.29}$$

We have that $t^{c_{k^*}} = O(1) = (\log t)^{o(1)}$, $(\log t)^{b_{k^*}} = (\log t)^{\frac{1}{3-\tau}+o(1)}$, and $D^{a_{k^*}} = (\log t)^{o(1)}$. Thus,

$$u_{k^*} = (\log t)^{\frac{1}{3-\tau}+o(1)}. \tag{4.30}$$

Thus, by picking t sufficiently large, we can make $\sigma \geq \frac{1}{3-\tau} + o(1)$. This completes the proof of Proposition 4.4. □

Proof of Theorem 4.1. We note that **whp** $\mathrm{diam}_{2t}(\mathrm{Core}_t) \leq 10 + 2k^*$, where k^* is the upper bound on $\max_{i\in\mathrm{Outer}_t} \min_{j\in\mathrm{Inner}_t} d_{G_m(2t)}(i,j)$ in Proposition 4.4, and we have made use of Proposition 4.2. This proves Theorem 4.1. □

4.2 Connecting the periphery to the core

In this section, we extend the results of the previous section and, in particular, study the distance between the vertices not in the core Core_t and the core. The main result in this section is the following theorem:

Theorem 4.7 (Connecting the periphery to the core)**.** *Fix $m \geq 2$. For every $\sigma > \frac{1}{3-\tau}$,* **whp***, the maximal distance between any vertex and* Core_t *in $G_m(2t)$ is bounded from above by $2\sigma \log\log t/\log m$.*

Together with Theorem 4.1, Theorem 4.7 proves the main result in Theorem 1.6.

The proof of Theorem 4.7 consists of two key steps. The first in Proposition 4.8 states that for any two vertices $i, j \in \{1, 2, \ldots, t\}$ the distance $d_{G_m(2t)}(i,j)$ is bounded by a constant times $\log\log t$, i.e., $\mathrm{diam}_{2t}(G_m(t))$ is bounded by some constant times $\log\log t$. The second in Proposition 4.11 shows that the distance between any vertex in $\{t+1, t+2, \ldots, 2t\}$ and $\{1, 2, \ldots, t\}$ is bounded by another constant times $\log\log t$.

Proposition 4.8 (Connecting half of the periphery to the core)**.** *Fix $m \geq 2$. For every $\sigma > \frac{1}{3-\tau}$,* **whp***, the maximal distance between any vertex in $\{1, 2, \ldots, t\}$ and the core* Core_t *in $G_m(2t)$ is bounded from above by $\sigma \log\log t/\log m$.*

Proof. We start from a vertex $i \in \{1, 2, \ldots, t\}$ and will show that the probability that the distance between i and Core_t is at least $C \log\log t$ is $o(t^{-1})$ where $C = \sigma/\log m$. This proves the claim. For this, we explore the neighborhood of i as follows. From i, we connect its $m \geq 2$ edges. Then, successively, we connect the m edges from each of the at most m vertices that i has connected to and have not yet been explored. We continue in the same fashion. We call the arising process when we have explored up to distance k from the initial vertex i the *k-exploration tree.*

When we never connect two edges to the same vertex, then the number of vertices we can reach within k steps is *precisely* equal to m^k. We call an event where an edge connects to a vertex which already was in the exploration tree a *collision*. When k increases, the probability of a collision increases. However, the probability that there exists a vertex for which *more than 2* collisions occur in the k-exploration tree before it hits the core is small, as we prove now:

Lemma 4.9 (A bound on the probability of multiple collisions)**.** *Fix $m \geq 2$ and $\delta \in (-m, 0)$. Fix $C = \sigma/\log m$, $l \geq 1$, $b \in (0,1]$ and take $k \leq C \log\log t$. Then, for every vertex $i \in \{1, 2, \ldots, t\}$, the probability that its k-exploration tree has at least l collisions before it hits* $\mathrm{Core}_t \cup \{j : j \leq t^b\}$ *is bounded above by*

$$\left((\log t)^d t^{-b}\right)^l = (\log t)^{dl} t^{-bl},$$

for some $d > 0$.

Proof. Take $i \in \{\lceil t^b \rceil + 1, \lceil t^b \rceil + 2, \ldots, t\}$ and consider its k-exploration tree $\mathcal{T}_i^{(k)}$. Since we add edges after time t^b the denominator in (1.1)-(1.3) is at least t^b. Moreover, before hitting the core, any vertex in the k-exploration tree has degree at most $(\log t)^\sigma$. Hence, for $l = 1$, the probability mentioned in the statement of the lemma is at most

$$\sum_{v \in \mathcal{T}_i^{(k)}} \frac{d_v(t) + \delta}{t^b} \leq \sum_{v \in \mathcal{T}_i^{(k)}} \frac{(\log t)^\sigma}{t^b} \leq \frac{m^k (\log t)^\sigma}{t^b} \tag{4.31}$$

where the bound follows from $\delta < 0$ and $|\mathcal{T}_i^{(k)}| \leq m^k$. For general l this upper bound becomes:

$$\left(\frac{m^k (\log t)^\sigma}{t^b}\right)^l$$

When $k = C \log\log t$ with $C = \sigma/\log m$, we have that $m^{lk} = (\log t)^{l\sigma}$. Therefore, the claim in Lemma 4.9 holds with $d = 2\sigma$. □

We next prove that, **whp**, $\{j : j \leq t^b\}$ is a subset of the core:

Lemma 4.10 (Early vertices have large degrees **whp**)**.** *Fix $m \geq 1$. There exists a $b > 0$ such that,* **whp**, $\min_{j \leq t^b} d_j(t) \geq (\log t)^\sigma$, *for some $\sigma > \frac{1}{3-\tau}$. As a result,* **whp**, $\{j : j \leq t^b\} \subseteq \mathrm{Core}_t$.

We defer the proof of Lemma 4.10 to Section 5.3 of the appendix.

Now we are ready to complete the proof of Proposition 4.8:

Proof of Proposition 4.8. By combining Lemmas 4.9 and 4.10, the probability that there exists an $i \in \{1, 2, \ldots, t\}$ for which the exploration tree $\mathcal{T}_i^{(k)}$ has at least l collisions before hitting the core is $o(1)$, whenever $l > 1/b$, since, by Boole's inequality, it is bounded by

$$\sum_{i=1}^{t} (\log t)^{dl} t^{-bl} = (\log t)^{2\sigma l} t^{-bl+1} = o(1), \tag{4.32}$$

precisely when $l > \frac{1}{b}$. When the k-exploration tree hits the core, then we are done by Theorem 4.1. When the k-exploration tree from a vertex i does not hit the core, but has less than l collisions, then there are at least m^{k-l} vertices in k-exploration tree. Indeed, when there are at most l collisions, the minimal size of the tree is obtained by identifying at most l vertices and their complete offspring, and the size of the pruned tree has size at least m^{k-l}.

When $k = C \log\log t$ with $C = \sigma/\log m$, this number is at least equal to $(\log t)^{\sigma+o(1)}$. The total weight of the core is, by (5.6) in the appendix, at least

$$\sum_{i\in \mathrm{Core}_t} (d_i(t)+\delta) \geq Bt(\log t)^{-(\tau-2)\sigma}. \tag{4.33}$$

The probability that there does not exist a t-connector between the k-exploration tree and the core is, by (4.10) bounded above by

$$\exp\left\{-\frac{\eta Bt(\log t)^{-(\tau-2)\sigma}(\log t)^{\sigma+o(1)}}{t}\right\} = o(t^{-1}), \tag{4.34}$$

by picking B sufficiently large, since $\sigma > 1/(3-\tau)$. This completes the proof. □

Proposition 4.11. *Fix $m \geq 2$. For every $\sigma > \frac{1}{3-\tau}$,* **whp**, *the maximal distance between any vertex and $\mathrm{Core}_t \cup \{1,2,\ldots,t\}$ in $G_m(2t)$ is bounded from above by $\frac{\sigma\log\log t}{\log m}$.*

Proof. Denote $k = \lfloor \frac{\sigma \log\log t}{\log m}\rfloor - 1$. We again grow the k-exploration trees from the vertices $i \in \{t+1, t+2,\ldots,2t\}$.

By Lemma 4.9 for $b = 1$, the probability that there exists a vertex whose k-exploration tree contains at least two collisions before hitting the vertex set $\mathrm{Core}_t \cup \{1,2,\ldots,t\}$ is bounded above by $t^{-2}(\log t)^{d_1}$ for some d_1 sufficiently large. When the k-exploration tree contains a vertex in $\mathrm{Core}_t \cup \{1,2,\ldots,t\}$, then we are done by Proposition 4.8 and Theorem 4.1. If not, and there are at most 2 collisions, then there are at least $m_k = (m-1)m^{k-1}$ vertices in $\{t+1,t+2,\ldots,2t\}$ at distance precisely equal to k from the original vertex. Denote these vertices by $i_1,\ldots,i_{m_k}$, and denote the k-exploration tree of vertex $i \in \{t+1,t+2,\ldots,2t\}$ by $\mathcal{T}_i^{(k)}$. We write

$$\begin{aligned}&\mathbb{P}(\nexists j \in \{1,2,\ldots,m_k\} \text{ such that } i_j \longrightarrow \{1,2,\ldots,t\}|\mathcal{T}_i^{(k)}) \\ &\quad= \prod_{j=1}^{m_k} \mathbb{P}(i_j \not\longrightarrow \{1,2,\ldots,t\}|i_s \not\longrightarrow \{1,2,\ldots,t\}\forall s<j, \mathcal{T}_i^{(k)}).\end{aligned} \tag{4.35}$$

Now we note that, uniformly in the way all edges in $G_m(2t)$ are formed, we have that for every $s \in \{t+1,t+2,\ldots,2t\}$,

$$\frac{\sum_{i=1}^t (d_i(s)+\delta)}{(2m+\delta)s} \geq \frac{1}{2}. \tag{4.36}$$

Thus, for any vertex i_j in the boundary of $\mathcal{T}_i^{(k)}$, the probability that it will be directly connected to $\{1,2,\ldots,t\}$ is at least $1/2$. As a result, we have that, uniformly in t, i and j,

$$\mathbb{P}(i_j \not\longrightarrow \{1,2,\ldots,t\}|i_s \not\longrightarrow \{1,2,\ldots,t\}\forall s<j, \mathcal{T}_i^{(k)}) \leq \frac{(2m+\delta)t}{(2m+\delta)(2t)} = \frac{1}{2}. \tag{4.37}$$

Therefore, we obtain that

$$\mathbb{P}(\nexists j = 1,\ldots,m_k \text{ such that } i_j \longleftrightarrow \{1,2,\ldots,t\}|\mathcal{T}_i^{(k)}) \leq 2^{-m_k}. \tag{4.38}$$

Since $m_k = \frac{m-1}{m}(\log t)^\sigma$, with $\sigma = \frac{1}{3-\tau} > 1$, we have that $2^{-m_k} = o(t^{-1})$. Therefore, any vertex $i \in \{t+1,t+2,\ldots,2t\}$ is, **whp**, within distance $k+1$ from $\{1,2,\ldots,t\}$. □

Proof of Theorem 4.7. Proposition 4.11 states that **whp** every vertex in $G_m(2t)$ is within distance $k+1 = \lfloor \frac{\sigma\log\log t}{\log m}\rfloor$ of $\mathrm{Core}_t \cup \{1,2,\ldots,t\}$. Proposition 4.8 states that **whp** every vertex in $\{1,2,\ldots,t\}$

is at most distance $k+1$ from the core Core_t. This shows that every vertex in $G_m(2t)$ is **whp** within distance $2(k+1)$ from the core. □

Proof of Theorem 1.6. Theorem 4.7 states that every vertex in $G_m(2t)$ is within distance $\frac{2\sigma \log\log t}{\log m}$ of the core Core_t. Theorem 4.1 states that the diameter of the core is at most $\frac{4\log\log t}{|\log(\tau-2)|}(1+o(1))$, so that the diameter of $G_m(2t)$ is at most $C_G \log\log t$, where C_G is given in (1.9). This completes the proof of Theorem 1.6. □

5 Appendix

5.1 The second moment of the degree sequence

Lemma 5.1. *For some constant $M_4 > 0$,*

$$\mathbb{E}[(d_s(t)+\delta)^2] \leq M_4 (t/s)^{2a}. \tag{5.1}$$

Proof. We start with model (c). We can compute $\mathbb{E}[(d_s(t_1+1)+\delta)^2]$ recursively by:

$$\begin{aligned}
&\mathbb{E}[(d_s(t_1+1)+\delta)^2|G_m(t_1)] = \mathbb{E}[(d_s(t_1+1)-d_s(t_1)+d_s(t_1)+\delta)^2|G_m(t_1)] \\
&\quad = \mathbb{E}[(d_s(t_1+1)-d_s(t_1))^2|G_m(t_1)] + 2(d_s(t_1)+\delta)\mathbb{E}[d_s(t_1+1)-d_s(t_1)|G_m(t_1)] + (d_s(t_1)+\delta)^2 \\
&\quad = \mathbb{E}\Big[\Big(\sum_{j=1}^m I[g(t_1,j)=s]\Big)^2\Big|G_m(t_1)\Big] + \frac{2m(d_s(t_1)+\delta)^2}{(2m+\delta)t_1} + (d_s(t_1)+\delta)^2 \\
&\quad = m\left(\frac{d_s(t_1)+\delta}{(2m+\delta)t_1}\right) - m\left(\frac{d_s(t_1)+\delta}{(2m+\delta)t_1}\right)^2 + m^2\left(\frac{d_s(t_1)+\delta}{(2m+\delta)t_1}\right)^2 + \frac{2m(d_s(t_1)+\delta)^2}{(2m+\delta)t_1} + (d_s(t_1)+\delta)^2,
\end{aligned}$$

from which we obtain

$$\mathbb{E}[(d_s(t_1+1)+\delta)^2] = \left(1+\frac{m^2-m}{(2m+\delta)^2 t_1^2}+\frac{2m}{(2m+\delta)t_1}\right)\mathbb{E}[(d_s(t_1)+\delta)^2] + \frac{m}{(2m+\delta)t_1}\mathbb{E}[d_s(t_1)+\delta].$$

Define

$$q_s(t_1) = \mathbb{E}[(d_s(t_1)+\delta)^2], \qquad e_s(t_1) = \mathbb{E}[d_s(t_1)+\delta], \qquad a = \frac{m}{2m+\delta}, \qquad b = \frac{1}{2m+\delta}.$$

Then from (2.15) and the above, respectively,

$$\begin{aligned}
e_s(t_1+1) &= (1+\frac{a}{t_1})e_s(t_1) \\
q_s(t_1+1) &= \Big(1+\frac{a(a-b)}{t_1^2}+\frac{2a}{t_1}\Big)q_s(t_1) + \frac{a}{t_1}e_s(t_1).
\end{aligned}$$

We write $q_s(t+1) = (1+a(t))q_s(t) + r_s(t)$ with $a(k) = \frac{a(a-b)+2ak}{k^2}$ and $r_s(t) = \frac{a}{t}e_s(t)$, then

$$\begin{aligned}
q_s(t) &= (1+a(t-1))q_s(t-1) + r_s(t-1) \\
&= q_s(s)\prod_{k=s}^{t-1}(1+a(k)) + \sum_{k=s}^{t-1} r_s(k) \prod_{l=k+1}^{t-1}(1+a(l)).
\end{aligned} \tag{5.2}$$

Since $a(k) = \frac{a(a-b)+2ak}{k^2} \leq \frac{2a}{k}$, and $a \geq b > 0$,

$$\prod_{k=s}^{t-1}(1+a(k)) \leq \exp\{\sum_{k=s}^{t-1} a(k)\} \leq \exp\{\sum_{k=s}^{t-1}\frac{2a}{k}\} \leq \exp\{\int_{s-1}^{t-1}\frac{2a}{x}dx\} = \Big(\frac{t-1}{s-1}\Big)^{2a}. \tag{5.3}$$

Substitution of $r_s(t) = \frac{a}{t} e_s(t)$ and using that by (2.16),

$$e_s(t) \le M\Big(\frac{t}{s}\Big)^a, \tag{5.4}$$

we find, for some constant $M_5 > 0$,

$$\begin{aligned}\sum_{k=s}^{t-1} r_s(k) \prod_{l=k+1}^{t-1} (1+a(l)) &\le (t-1)^{2a} \sum_{k=s}^{t-1} \frac{r_s(k)}{k^{2a}} \le a(t-1)^{2a} \sum_{k=s}^{t-1} \frac{e_s(k)}{k^{1+2a}} \\ &\le aM \frac{t^{2a}}{s^a} \sum_{k=s}^{t-1} \frac{1}{k^{1+a}} \le aM \frac{t^{2a}}{s^a} \sum_{k=s}^{\infty} \frac{1}{k^{1+a}} \le M_5 \frac{t^{2a}}{s^{2a}}. \end{aligned} \tag{5.5}$$

From (5.2) and (5.5) together, we obtain (5.1). □

5.2 The total degree of an event

Lemma 5.2. *Assume that $l = l(t) \to \infty$, as $t \to \infty$ and that $l(t) \le u_1$, then there exists a constant $B > 0$ such that with probability exceeding $1 - o(t^{-1})$,*

$$\sum_{i: d_i(t) \ge l} d_i(t) \ge B t l^{2-\tau}. \tag{5.6}$$

Proof. We note that

$$\sum_{i: d_i(t) \ge l} d_i(t) \ge l P_{\ge l}(t), \tag{5.7}$$

where $P_{\ge l}(t) = \#\{i \le t : d_i(t) \ge l\}$ is the number of vertices with degree at least l.

In [18], detailed asymptotics for $P_{\ge l}(t)$ were proved for model (c) that we will survey now. These asymptotics play a key role throughout the proof. We shall comment on the adaptations of the proofs for models (a) and (b) below.

Firstly, it is shown that there exists a B_1 such that uniformly for all l,

$$\mathbb{P}\Big(|P_{\ge l}(t) - \mathbb{E}[P_{\ge l}(t)]| \ge B_1 \sqrt{t \log t}\Big) = o(t^{-1}). \tag{5.8}$$

This proves a *concentration bound* on the number of vertices with at least a given degree. The proof of this result follows the argument in [15], and holds for any of the models (a)-(c).

Secondly, it is shown that with

$$P_l(t) = \#\{i \le t : d_i(t) = l\} \tag{5.9}$$

equal to the total number of vertices of degree equal to l, and with p_k defined by

$$p_k = \frac{\theta \Gamma(k+\delta)\Gamma(m+\delta+\theta)}{\Gamma(m+\delta)\Gamma(k+1+\delta+\theta)}, \quad k \ge m, \tag{5.10}$$

so that $p_k \sim k^{-\tau}$, and where $\theta = 2 + \delta/m$, there exists a constant B_2 such that

$$\sup_{l \ge 1} |\mathbb{E}[P_l(t)] - t p_l| \le B_2. \tag{5.11}$$

For model (c), this is shown in [18], for model (a) this is shown in [25, Chapter 8]. This latter proof can easily be adapted to deal with model (b) as well. In rather generality, results of this kind (with the sharp bound in (5.11)) are proved in [24].

Therefore, we obtain that, with probability exceeding $1 - o(t^{-1})$,

$$P_{\geq l_t}(t) \geq \mathbb{E}[P_{\geq l_t}(t)] - B_1\sqrt{t\log t} \geq \mathbb{E}[P_{\geq l_t}(t)] - \mathbb{E}[P_{\geq 2l_t}(t)] - B_1\sqrt{t\log t}$$
$$\geq \sum_{l=l_t}^{2l_t-1}[tp_l - B_2] - B_1\sqrt{t\log t} \geq B_3 t l_t^{1-\tau} - B_2 l_t - B_1\sqrt{t\log t}. \tag{5.12}$$

We now wish to pick l_t such that

$$t l_t^{1-\tau} \gg l_t, \tag{5.13}$$

or $l_t \ll t^{\frac{1}{\tau}}$, and

$$t l_t^{1-\tau} \gg \sqrt{t\log t}, \tag{5.14}$$

or $l_t \ll t^{\frac{1}{2(\tau-1)}}(\log t)^{-\frac{1}{2(\tau-1)}}$. Note that $\frac{1}{\tau} \geq \frac{1}{2(\tau-1)}$ for all $\tau > 2$, so we need

$$l_t \leq t^{\frac{1}{2(\tau-1)}}(\log t)^{-\frac{1}{2}} \equiv u_1. \tag{5.15}$$

Then, for this choice, we have with probability exceeding $1 - o(t^{-1})$,

$$\sum_{i:d_i(t)\geq l_t} d_i(t) \geq Bt l_t^{2-\tau}. \tag{5.16}$$

Also, for this choice, **whp**, $P_{\geq l_t}(t) \gg \sqrt{t}$. □

5.3 Proof of Lemmas 4.3 and 4.10

Proof of Lemma 4.3. We prove the bound for model (c), the proof for models (a) and (b) is similar. We first write the degree of vertex i as

$$d_i(t) = m + \sum_{s=i+1}^{t}\sum_{j=1}^{m} I[g(s,j) = i], \tag{5.17}$$

where we recall that $g(s,j) = i$ denotes that the j^{th} edge of vertex s is attached to vertex i. By Lemma 3.1, the indicator variables $I[g(s_1,j_1) = i_1]$ and $I[g(s_2,j_2) = i_2]$ are negatively corelated for any $i_1 \neq i_2$. As a result, we obtain that $d_{i_1}(t)$ and $d_{i_2}(t)$ are negatively correlated, so that

$$\text{Var}\Big(\sum_{i=1}^{\log t} d_i(t)\Big) \leq \sum_{i=1}^{\log t}\text{Var}(d_i(t)). \tag{5.18}$$

Also, by (2.15), it follows that $\mathbb{E}[d_i(t)] \geq c\Big(\frac{t}{i}\Big)^a$, so that, for some $\varepsilon > 0$, and noting that $a < 1$,

$$\mathbb{E}[\sum_{i=1}^{\log t} d_i(t)] \geq \varepsilon t^a(\log t)^{1-a}. \tag{5.19}$$

Furthermore, by (3.17), we have that

$$\text{Var}(d_i(t)) = \text{Var}(d_i(t) + m) \leq \mathbb{E}[(d_i(t) + \delta)^2] \leq M_4\Big(\frac{t}{i}\Big)^{2a}, \tag{5.20}$$

so that, for any $\varepsilon > 0$, and noting that $a > 1/2$ for $\delta < 0$,

$$\text{Var}\Big(\sum_{i=1}^{\log t} d_i(t)\Big) \leq \sum_{i=1}^{\log t} M_4\Big(\frac{t}{i}\Big)^{2a} = O(t^{2a}). \tag{5.21}$$

As a result, by the second moment method, we obtain that, noting that $a = \frac{1}{\tau-1} \in (\frac{1}{2}, 1)$,

$$\begin{aligned}
\mathbb{P}(\max_{i\leq \log t} d_i(t) \leq t^a(\log t)^{-1}) &\leq \mathbb{P}\Big(\sum_{i\leq \log t} d_i(t) \leq t^a\Big) \\
&\leq \mathbb{P}\Big(\sum_{i\leq \log t} |d_i(t) - \mathbb{E}[d_i(t)]| \geq \varepsilon t^a (\log t)^{1-a}\Big) \\
&\leq \frac{O(t^{2a})}{\big(\varepsilon t^a (\log t)^{1-a}\big)^2} = O\Big((\log t)^{-2(1-a)}\Big) = o(1). \qquad (5.22)
\end{aligned}$$

□

Proof of Lemma 4.10. We investigate the problem for model (a) first, the adaptation of the proof for model (b) is rather straightforward and will be treated immediately after the proof for model (a). The proof for model (c) is slightly more involved.

We first note that, for models (a) and (b), the model for general $m \geq 1$ is obtained from the model for $m = 1$ by taking $\delta' = \delta/m$ and identifying groups of m vertices. Thus, the degree of vertex i in $G_m(t)$ is bounded from below by the degree of vertex im in $G_1(mt)$. We now prove the statement for $m = 1$ and $\delta > -1$ fixed. We shall show by induction on j that, for $m = 1$, that for all $t \geq i$

$$\mathbb{P}(d_i(t) = j) \leq C_j \frac{\Gamma(t)\Gamma(i + \frac{1+\delta}{2+\delta})}{\Gamma(t + \frac{1+\delta}{2+\delta})\Gamma(i)}, \qquad (5.23)$$

where C_j will be determined in the course of the proof. Clearly, for every $t \geq i$, for model (a),

$$\mathbb{P}(d_i(t) = 1) = \prod_{s=i+1}^{t} \Big(1 - \frac{1+\delta}{(2+\delta)(s-1) + (1+\delta)}\Big) = \prod_{s=i+1}^{t} \Big(\frac{s-1}{s-1+\frac{1+\delta}{2+\delta}}\Big) = \frac{\Gamma(t)\Gamma(i+\frac{1+\delta}{2+\delta})}{\Gamma(t+\frac{1+\delta}{2+\delta})\Gamma(i)}, \qquad (5.24)$$

which initializes the induction hypothesis with $C_1 = 1$.

To advance the induction, we let $s \leq t$ be the last time at which a vertex is added to i. Then we have that

$$\mathbb{P}(d_i(t) = j) = \sum_{s=i+j-1}^{t} \mathbb{P}(d_i(s-1) = j-1) \frac{j-1+\delta}{(2+\delta)(s-1)+1+\delta} \mathbb{P}(d_i(t) = j | d_i(s) = j). \qquad (5.25)$$

By the induction hypothesis, we have that

$$\mathbb{P}(d_i(s-1) = j-1) \leq C_{j-1} \frac{\Gamma(s-1)\Gamma(i+\frac{1+\delta}{2+\delta})}{\Gamma(s-1+\frac{1+\delta}{2+\delta})\Gamma(i)}. \qquad (5.26)$$

Moreover, analogously to (5.24), we have that

$$\begin{aligned}
\mathbb{P}(d_i(t) = j | d_i(s) = j) &= \prod_{q=s+1}^{t} \Big(1 - \frac{j+\delta}{(2+\delta)(q-1)+(1+\delta)}\Big) \qquad (5.27) \\
&= \prod_{q=s+1}^{t} \Big(\frac{q-1-\frac{j-1}{2+\delta}}{q-1+\frac{1+\delta}{2+\delta}}\Big) = \frac{\Gamma(t-\frac{j-1}{2+\delta})\Gamma(s+\frac{1+\delta}{2+\delta})}{\Gamma(t+\frac{1+\delta}{2+\delta})\Gamma(s-\frac{j-1}{2+\delta})}.
\end{aligned}$$

Combining (5.26) and (5.27), we arrive at

$$\mathbb{P}(d_i(t)=j) \le \sum_{s=i+j-1}^{t} \Big(C_{j-1}\frac{\Gamma(s-1)\Gamma(i+\frac{1+\delta}{2+\delta})}{\Gamma(s-1+\frac{1+\delta}{2+\delta})\Gamma(i)}\Big)\Big(\frac{j-1+\delta}{(2+\delta)(s-1)+(1+\delta)}\Big)\Big(\frac{\Gamma(t-\frac{j-1}{2+\delta})\Gamma(s+\frac{1+\delta}{2+\delta})}{\Gamma(t+\frac{1+\delta}{2+\delta})\Gamma(s-\frac{j-1}{2+\delta})}\Big). \tag{5.28}$$

We next use that

$$\Gamma(s-1+\frac{1+\delta}{2+\delta})((2+\delta)(s-1)+(1+\delta)) = (2+\delta)\Gamma(s+\frac{1+\delta}{2+\delta}), \tag{5.29}$$

to arrive at

$$\mathbb{P}(d_i(t)=j) \le C_{j-1}\frac{j-1+\delta}{2+\delta}\frac{\Gamma(i+\frac{1+\delta}{2+\delta})}{\Gamma(i)}\frac{\Gamma(t-\frac{j-1}{2+\delta})}{\Gamma(t+\frac{1+\delta}{2+\delta})}\sum_{s=i+j-1}^{t}\frac{\Gamma(s-1)}{\Gamma(s-\frac{j-1}{2+\delta})}. \tag{5.30}$$

We note that, whenever $l+b, l+1+a>0$ and $a-b+1>0$,

$$\sum_{s=l}^{t}\frac{\Gamma(s+a)}{\Gamma(s+b)} = \frac{1}{a-b+1}\Big[\frac{\Gamma(t+1+a)}{\Gamma(t+b)} - \frac{\Gamma(l+1+a)}{\Gamma(l+b)}\Big] \le \frac{1}{a-b+1}\frac{\Gamma(t+1+a)}{\Gamma(t+b)}. \tag{5.31}$$

Application of (5.31) for $a=-1, b=-\frac{j-1}{2+\delta}, l=i+j-1$, so that $a-b+1=\frac{j-1}{2+\delta}>0$ when $j>1$, leads to

$$\begin{aligned}\mathbb{P}(d_i(t)=j) &\le C_{j-1}\frac{j-1+\delta}{2+\delta}\frac{\Gamma(i+\frac{1+\delta}{2+\delta})}{\Gamma(i)}\frac{\Gamma(t-\frac{j-1}{2+\delta})}{\Gamma(t+\frac{1+\delta}{2+\delta})}\frac{1}{\frac{j-1}{2+\delta}}\frac{\Gamma(t)}{\Gamma(t-\frac{j-1}{2+\delta})} \\ &= C_{j-1}\frac{j-1+\delta}{j-1}\frac{\Gamma(i+\frac{1+\delta}{2+\delta})}{\Gamma(i)}\frac{\Gamma(t)}{\Gamma(t+\frac{1+\delta}{2+\delta})}.\end{aligned} \tag{5.32}$$

Equation (5.32) advances the induction when we define

$$C_j = \frac{j-1+\delta}{j-1}C_{j-1}. \tag{5.33}$$

This completes the investigation of the probability that $\mathbb{P}(d_i(t)=j)$ for model (a).

For model (b), the argument is quite similar. Indeed, we now use as an induction hypothesis that

$$\mathbb{P}(d_i(t)=j) \le C_j\frac{\Gamma(t-\frac{1+\delta}{2+\delta})\Gamma(i)}{\Gamma(t)\Gamma(i-\frac{1+\delta}{2+\delta})}, \tag{5.34}$$

where again C_j will be determined in the course of the proof. Clearly, for every $t\ge i$, in model (b),

$$\mathbb{P}(d_i(t)=1) = \prod_{s=i+1}^{t}\Big(1-\frac{1+\delta}{(2+\delta)(s-1)}\Big) = \prod_{s=i+1}^{t}\Big(\frac{s-1-\frac{1+\delta}{2+\delta}}{s-1}\Big) = \frac{\Gamma(t-\frac{1+\delta}{2+\delta})\Gamma(i)}{\Gamma(t)\Gamma(i-\frac{1+\delta}{2+\delta})}, \tag{5.35}$$

which again initializes the induction hypothesis, and

$$\mathbb{P}(d_i(t)=j|d_i(s)=j) = \prod_{q=s+1}^{t}\Big(1-\frac{j+\delta}{(2+\delta)(q-1)}\Big) = \frac{\Gamma(t-\frac{j+\delta}{2+\delta})\Gamma(s)}{\Gamma(t)\Gamma(s-\frac{j+\delta}{2+\delta})}. \tag{5.36}$$

Therefore, (5.32) is now replaced by

$$\mathbb{P}(d_i(t)=j) \le C_{j-1}\frac{j-1+\delta}{2+\delta}\frac{\Gamma(i)}{\Gamma(i-\frac{1+\delta}{2+\delta})}\frac{\Gamma(t-\frac{j+\delta}{2+\delta})}{\Gamma(t)}\sum_{s=i+j-1}^{t}\frac{\Gamma(s-1-\frac{1+\delta}{2+\delta})}{\Gamma(s-\frac{j+\delta}{2+\delta})} \tag{5.37}$$
$$= C_{j-1}\frac{j-1+\delta}{j-1}\frac{\Gamma(i)}{\Gamma(i-\frac{1+\delta}{2+\delta})}\frac{\Gamma(t-\frac{1+\delta}{2+\delta})}{\Gamma(t)}.$$

We next turn to model (c). In the proofs for models (a) and (b), we have made crucial use of the relation between $G_m(t)$ and $G_1(mt)$ for these models, so that it was sufficient to investigate the case where $m=1$. When $m=1$, at any time step, at most one edge can be added. This relation unfortunately fails for model (c), and we first adapt the argument. Recall that $d_i(t)$ is the degree of vertex i at time t. We shall define $e_i(t)$ such that $e_i(t) \le d_i(t)$ and $e_i(t)$ grows by at most one at each time step. The definition of $e_i(t)$ is recursive. We let $e_i(i)=d_i(i)=m$, and, assuming we have shown that $d_i(t)=e_i(t)+r_i(t)$, where $r_i(t)\ge 0$, we proceed to time $t+1$ as follows. We can increase $e_i(t)$ only when the *first* edge of vertex $t+1$ attaches to vertex i, and this is the case with probability $\frac{e_i(t)+\delta}{(2m+\delta)t}$. With probability $\frac{r_i(t)}{(2m+\delta)t}$, we keep $e_i(t+1)=e_i(t)$ and we increase $r_i(t)$ by one. For the other $m-1$ edges, we increase $r_i(t)$ by one with probability $\frac{d_i(t)+\delta}{(2m+\delta)t}$. Then we clearly have that $e_i(t+1)\le d_i(t+1)$ if $e_i(t)\le d_i(t)$, since the difference between $d_i(t)$ and $e_i(t)$ equals $r_i(t)$, which is monotonically increasing. Moreover, we have that $e_i(t+1)$ equals $e_i(t)$ or $e_i(t)+1$, and the latter occurs with conditional probability

$$\mathbb{P}(e_i(t+1)=j|e_i(t)=j-1)=\frac{j-1+\delta}{(2m+\delta)t}. \tag{5.38}$$

We now adapt the above argument for models (a)-(b) to the random variable $e_i(t)$. Indeed, we now use as an induction hypothesis that

$$\mathbb{P}(e_i(t)=j) \le C_j\frac{\Gamma(t-\frac{m+\delta}{2m+\delta})\Gamma(i)}{\Gamma(t)\Gamma(i-\frac{m+\delta}{2m+\delta})}, \tag{5.39}$$

where again C_j will be determined in the course of the proof. Clearly, for every $t\ge i$, in model (c),

$$\mathbb{P}(e_i(t)=m)=\prod_{s=i+1}^{t}\Big(1-\frac{m+\delta}{(2m+\delta)(s-1)}\Big)=\prod_{s=i+1}^{t}\Big(\frac{s-1-\frac{m+\delta}{2m+\delta}}{s-1}\Big)=\frac{\Gamma(t-\frac{m+\delta}{2m+\delta})\Gamma(i)}{\Gamma(t)\Gamma(i-\frac{m+\delta}{2m+\delta})}, \tag{5.40}$$

and

$$\mathbb{P}(e_i(t)=j|e_i(s)=j)=\prod_{q=s+1}^{t}\Big(1-\frac{j+\delta}{(2m+\delta)(q-1)}\Big)=\frac{\Gamma(t-\frac{j+\delta}{2m+\delta})\Gamma(s)}{\Gamma(t)\Gamma(s-\frac{j+\delta}{2m+\delta})}. \tag{5.41}$$

Therefore, for $j>m$, the recursion leading to (5.32) is now replaced by

$$\mathbb{P}(e_i(t)=j) \le \sum_{s=i+j-1}^{t} C_{j-1}\Big(\frac{\Gamma(s-1-\frac{m+\delta}{2m+\delta})\Gamma(i)}{\Gamma(s-1)\Gamma(i-\frac{m+\delta}{2m+\delta})}\Big)\Big(\frac{j-1+\delta}{(2m+\delta)(s-1)}\Big)\Big(\frac{\Gamma(t-\frac{j+\delta}{2m+\delta})\Gamma(s)}{\Gamma(t)\Gamma(s-\frac{j+\delta}{2m+\delta})}\Big) \tag{5.42}$$
$$= C_{j-1}\frac{j-1+\delta}{2m+\delta}\frac{\Gamma(i)}{\Gamma(i-\frac{m+\delta}{2m+\delta})}\frac{\Gamma(t-\frac{j+\delta}{2m+\delta})}{\Gamma(t)}\sum_{s=i+j-1}^{t}\frac{\Gamma(s-1-\frac{m+\delta}{2m+\delta})}{\Gamma(s-\frac{j+\delta}{2m+\delta})}$$
$$\le C_{j-1}\frac{j-1+\delta}{j-m}\frac{\Gamma(i)}{\Gamma(i-\frac{m+\delta}{2m+\delta})}\frac{\Gamma(t-\frac{m+\delta}{2m+\delta})}{\Gamma(t)},$$

so that C_j is now given by $C_m = 1$ and, for $j > m$,

$$C_j = \frac{j-1+\delta}{j-m} C_{j-1}. \tag{5.43}$$

We summarize the bounds in models (a)-(c) by the fact that we have the bound, for all $m \geq 1$,

$$\mathbb{P}(d_i(t) = j) \leq C_j \frac{\Gamma(t-a_1)\Gamma(i+a_2)}{\Gamma(t+a_2)\Gamma(i-a_1)}, \tag{5.44}$$

where $a_1 = 0$ for model (a), while $a_1 = \frac{m+\delta}{2m+\delta}$ for models (b)–(c), while $a_2 = \frac{m+\delta}{2m+\delta}$ for model (a), while $a_2 = 0$ for models (b)–(c), and, for all models, $C_j \leq j^{p-1}$ for some $p \geq 1$.

Below, we shall rely on the obvious consequence of (5.44) that

$$\mathbb{P}(d_i(t) \leq j) \leq j^p \frac{\Gamma(t-a_1)\Gamma(i+a_2)}{\Gamma(t+a_2)\Gamma(i-a_1)}. \tag{5.45}$$

Obviously, for t and i large, we have that

$$\mathbb{P}(d_i(t) \leq j) \leq j^p t^{-(a_1+a_2)} i^{a_1+a_2}(1+o(1)). \tag{5.46}$$

We finally use (5.45) to complete the proof of Lemma 4.10. Take $0 < b < \frac{a_1+a_2}{a_1+a_2+1} = \frac{m+\delta}{3m+2\delta}$, then, by Boole's inequality,

$$\begin{aligned} \mathbb{P}(\exists i \leq t^b : d_i(t) \leq (\log t)^\sigma) &\leq \sum_{i=1}^{t^b} \mathbb{P}(d_i(t) \leq (\log t)^\sigma) \leq (\log t)^{\sigma p} \frac{\Gamma(t-a_1)}{\Gamma(t+a_2)} \sum_{i=1}^{t^b} \frac{\Gamma(i+a_2)}{\Gamma(i-a_1)} \\ &\leq (\log t)^{\sigma p}(a_1+a_2+1)^{-1} \frac{\Gamma(t-a_1)}{\Gamma(t+a_2)} \frac{\Gamma(t^b+a_2+1)}{\Gamma(t^b-a_1)} = o(1), \end{aligned} \tag{5.47}$$

by a similar equality as in (5.46). This completes the proof of Lemma 4.10. □

Acknowledgements

The work of RvdH was supported in part by Netherlands Organisation for Scientific Research (NWO). We thank Mia Deijfen and Henri van den Esker for many useful conversations throughout the project.

References

[1] W. Aiello, F. Chung and L. Lu. Random evolution in massive graphs. In *Handbook of massive data sets*, volume **4** of *Massive Comput.*, pages 97–122. Kluwer Acad. Publ., Dordrecht, (2002).

[2] R. Albert and A.-L. Barabási. Statistical mechanics of complex networks. *Rev. Modern Phys.*, **74**(1):47–97, (2002).

[3] N. Alon and J.H. Spencer. *The probabilistic method.* John Wiley & Sons, New York, second edition, (2000).

[4] A.-L. Barabási. *Linked: The New Science of Networks.* Perseus Publishing, Cambridge, Massachusetts, (2002).

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

[6] N. Berger, B. Bollobás, C. Borgs, J. Chayes, and O. Riordan. Degree distribution of the FKP network model. In *Automata, languages and programming*, volume **2719** of *Lecture Notes in Comput. Sci.*, pages 725–738. Springer, Berlin, (2003).

[7] B. Bollobás. *Random graphs*, volume **73** of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, second edition, (2001).

[8] B. Bollobás, C. Borgs, J. Chayes, and O. Riordan. Directed scale-free graphs. In *Proceedings of the Fourteenth Annual ACM-SIAM Symposium on Discrete Algorithms (Baltimore, MD, 2003)*, pages 132–139, New York, (2003).

[9] B. Bollobás, S. Janson, and O. Riordan. The phase transition in inhomogeneous random graphs. Preprint (2006). Available on `http://arxiv.org/abs/math.PR/0504589`.

[10] B. Bollobás and O. Riordan. Mathematical results on scale-free random graphs. In *Handbook of graphs and networks*, pages 1–34. Wiley-VCH, Weinheim, (2003).

[11] B. Bollobás and O. Riordan. Robustness and vulnerability of scale-free random graphs. *Internet Math.*, **1**(1):1–35, (2003).

[12] B. Bollobás and O. Riordan. Coupling scale-free and classical random graphs. *Internet Math.*, **1**(2):215–225, (2004).

[13] B. Bollobás and O. Riordan. The diameter of a scale-free random graph. *Combinatorica*, **24**(1):5–34, (2004).

[14] B. Bollobás and O. Riordan. Shortest paths and load scaling in scale-free trees. *Phys. Rev. E.*, **69**:036114, (2004).

[15] B. Bollobás, O. Riordan, J. Spencer, and G. Tusnády. The degree sequence of a scale-free random graph process. *Random Structures Algorithms*, **18**(3):279–290, (2001).

[16] F. Chung and L. Lu. The average distances in random graphs with given expected degrees. *Proc. Natl. Acad. Sci. USA*, **99**(25):15879–15882 (electronic), (2002).

[17] C. Cooper and A. Frieze. A general model of web graphs. *Random Structures Algorithms*, **22**(3):311–335, (2003).

[18] M. Deijfen, H. van den Esker, R. van der Hofstad, and G. Hooghiemstra. A preferential attachment model with random initial degrees. Preprint (2007). Available on `http://ssor.twi.tudelft.nl/ gerardh/`

[19] S.N. Dorogovtsev and J.F.F. Mendes. Evolution of networks. *Advances in Physics*, **51**:1079–1187, (2002).

[20] R. Durrett. *Random graph dynamics*. Cambridge University Press, Cambridge, (2007).

[21] H. van den Esker, R. van der Hofstad, and G. Hooghiemstra. Universality for the distance in finite variance random graphs. Preprint, (2006). Available on `http://ssor.twi.tudelft.nl/ gerardh/`

[22] H. van den Esker, R. van der Hofstad, G. Hooghiemstra, and D. Znamenski. Distances in random graphs with infinite mean degrees. *Extremes*, **8**:111–140, (2006).

[23] D. Fernholz and V. Ramachandran. The diameter of sparse random graphs. Preprint (2004). Available on `http://www.cs.utexas.edu/ fernholz/`

[24] O. Hagberg and C. Wiuf. Convergence properties of the degree distribution of some growing network models. *Bull. Math. Biol.*, **68**:1275–1291, (2006).

[25] R. van der Hofstad. Random graphs and complex networks. In preparation, (2008). Available on `http://www.win.tue.nl/ rhofstad/`

[26] R. van der Hofstad, G. Hooghiemstra and P. Van Mieghem. Distances in random graphs with finite variance degrees. Random Structures Algorithms **26**, 76-123, 2005.

[27] R. van der Hofstad, G. Hooghiemstra, and D. Znamenski. Distances in random graphs with finite mean and infinite variance degrees. (2004), to appear in *Electr. Journal Probab.*

[28] R. van der Hofstad, G. Hooghiemstra, and D. Znamenski. Random graphs with arbitrary i.i.d. degrees. Preprint Mittag-Leffer Institute, (2004). Available on `http://arxiv.org/abs/math.PR/0502580`.

[29] S. Janson, T. Łuczak, and A. Ruciński. *Random graphs*. Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley-Interscience, New York, (2000).

[30] M. E. J. Newman. The structure and function of complex networks. *SIAM Rev.*, **45**(2):167–256 (electronic), (2003).

[31] I. Norros and H. Reittu. On a conditionally Poissonian graph process. *Adv. in Appl. Probab.*, **38**(1):59–75, (2006).

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

[33] H. Reittu and I. Norros. On the power law random graph model of massive data networks. *Performance Evaluation*, **55**(1-2):3–23, (2004).