Diameters in preferential attachment models

Sander Dommers and Remco van der Hofstad *
Gerard Hooghiemstra†

March 20, 2009

Abstract

In this paper, we investigate the diameter in preferential attachment (PA-) models, thus quantifying the statement that these models are small worlds. The models studied here are such that edges are attached to older vertices proportional to the degree plus a constant, i.e., we consider *affine* PA-models. There is a substantial amount of literature proving that, quite generally, PA-graphs possess power-law degree sequences with a power-law exponent $\tau > 2$.

We prove that the diameter of the PA-model is bounded above by a constant times $\log t$, where t is the size of the graph. When the power-law exponent τ exceeds 3, then we prove that $\log t$ is the right order for the diameter, by proving a lower bound of this order, both for the diameter as well as for the average distance. This shows that, for $\tau > 3$, distances are of the order $\log t$. For $\tau \in (2, 3)$, we improve the upper bound to a constant times $\log \log t$, and prove a lower bound of the same order for the diameter. Unfortunately, this proof does not extend to average distances. These results do show that the diameter is of order $\log \log t$.

These bounds partially prove predictions by physicists that the average distance in PA-graphs are similar to the ones in other scale-free random graphs, such as the configuration model and various inhomogeneous random graph models, where average 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 networks are small. The scale-free phenomenon states that the degree sequences in these networks satisfy a power law. See [3, 24, 39] for reviews on complex networks, and [5] for a more expository account. Thus, these complex networks are not at all like classical random graphs (see [4, 9, 35] 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 random graph models that do obey power-law degree sequences. See [11] for the most general inhomogeneous random graph models, as well as a review of the models under investigation. Extensive discussions of various scale-free random graph models are given in [21, 25].

While these models have power-law degree sequences, they do not explain *why* many complex networks are scale free. A possible explanation was given by Barabási and Albert [6] 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

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

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 number of acquaintances (high degree). Interestingly, PA-models with so-called *affine* PA rules have power-law degree sequences, and, therefore, preferential attachment offers a convincing explanation why many real-world networks possess this property. There is a large amount of literature studying such models. See e.g. [2, 10, 12, 13, 14, 15, 17, 22] and the references therein. The literature primarily focusses on three main questions. The first key question for PA-models is to prove that such random graphs are indeed *scale free* [2, 10, 12, 13, 17, 22], by proving that their degree sequence indeed obeys a power law with a certain power-law exponent $\tau > 2$. The second key question for PA-models is their *vulnerability*, for example to deliberate attack [13] or to the spread of a disease [7]. The third key question for PA-models is to show that the resulting models are *small worlds* by investigating the distances in them. See in particular [15] for a result on the diameter for a PA-model with power-law exponent $\tau = 3$. In non-rigorous work, it is often suggested that many of the scale-free models, such as the configuration model, the inhomogeneous random graph models in [11] 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. Similar results are true for the so-called rank-1 inhomogeneous random graph (see e.g. [18, 19, 26, 40]). The natural question is, therefore, whether the same applies to preferential attachment models. This is the main goal of the present paper, in which we investigate the diameter of scale-free PA-models.

The remainder of this section is organized as follows. We first introduce the models that we will investigate in this paper. Then we give the main results and conclude with a discussion of universality in power-law random graphs.

In this paper, we investigate the diameter in some PA-models. The models that we investigate produce a *graph sequence* or *graph process* $\{G_{m,\delta}(t)\}$, which, for fixed $t \geq 1$ or $t \geq 2$, yields a graph with t vertices and mt edges for some given integer $m \geq 1$. In the sequel, we shall denote the vertices of $G_{m,\delta}(t)$ by $1^{(m)}, \ldots, t^{(m)}$. When m is clear from the context, we will leave out the superscript and write $[t] \equiv \{1, 2, \ldots, t\}$. We shall consider three slight variations of the PA-model, which we shall denote by models (a), (b) and (c), respectively.

(a) The first model is an extension of the Barabási-Albert model formulated rigorously in [17]. We start with $G_{1,\delta}(1)$ consisting of a single vertex with a single self-loop. We denote the degree of vertex $i^{(1)}$ at time t by $D_{i^{(1)}}(t)$, where, a self-loop increases the degree by 2.
Then, for $m = 1$, and conditionally on $G_{1,\delta}(t)$, the growth rule to obtain $G_{1,\delta}(t+1)$ is as follows. We add a single vertex $(t+1)^{(1)}$ having a single edge. This edge is connected to a second end point, which is equal to $(t+1)^{(1)}$ with probability proportional to $1+\delta$, and to a vertex $i^{(1)} \in G_{1,\delta}(t)$ with probability proportional to $D_{i^{(1)}}(t) + \delta$, where $\delta \geq -1$ is a parameter of the model. Thus,

$$\mathbb{P}\Big((t+1)^{(1)} \to i^{(1)} | G_{1,\delta}(t)\Big) = \begin{cases} \frac{1+\delta}{t(2+\delta)+(1+\delta)}, & \text{for } i = t+1, \\ \frac{D_{i^{(1)}}(t)+\delta}{t(2+\delta)+(1+\delta)}, & \text{for } i \in [t]. \end{cases} \tag{1.1}$$

The model with integer $m > 1$, is defined in terms of the model for $m = 1$ as follows. We start with $G_{1,\delta'}(mt)$, with $\delta' = \delta/m \geq -1$. Then we identify the vertices $1^{(1)}, 2^{(1)} \ldots, m^{(1)}$ in $G_{1,\delta}(mt)$ to be vertex $1^{(m)}$ in $G_{m,\delta}(t)$, and for $1 < j \leq t$, the vertices $((j-1)m+1)^{(1)}, \ldots, (jm)^{(1)}$ in $G_{1,\delta'}(mt)$ to be vertex $j^{(m)}$ in $G_{m,\delta}(t)$; in particular the degree $D_{j^{(m)}}(t)$ of vertex $j^{(m)}$ in $G_{m,\delta}(t)$ is equal to the sum of the degrees of the vertices $((j-1)m+1)^{(1)}, \ldots, (jm)^{(1)}$ in $G_{1,\delta'}(mt)$. This defines the model for integer $m \geq 1$. Observe that the range of δ is $[-m, \infty)$.
The resulting graph $G_{m,\delta}(t)$ has precisely mt edges and t vertices at time t, but is not necessarily connected. For $\delta = 0$ we obtain the original model studied in [17], and further studied in [13, 14, 15]. The extension to $\delta \neq 0$ is crucial in our setting, as we shall explain in more detail below.

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

$$\mathbb{P}\Big((t+1)^{(1)} \to i^{(1)} | G_{1,\delta}(t)\Big) = \frac{D_{i^{(1)}}(t) + \delta}{t(2+\delta)}, \qquad \text{for } i \in [t]. \tag{1.2}$$

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). This model is studied in detail in [25], and the model with $m = 1$ corresponds to scale-free trees as studied in e.g. [16, 37, 38, 42].

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

$$\mathbb{P}\Big(j^{\text{th}} \text{ edge of } (t+1)^{(m)} \text{ is connected to } i^{(m)} | G_{m,\delta}(t)\Big) = \frac{D_{i^{(m)}}(t) + \delta}{t(2m+\delta)}, \qquad \text{for } i \in [t]. \tag{1.3}$$

In this model, as is the case in model (b), the graph $G_{m,\delta}(t)$ is a connected random graph with precisely t vertices and mt edges. This model was studied in [23, 36].

Remark 1.1. *In models (a) and (b) for $m > 1$, the choice of $\delta' = \delta/m$ is such that in the resulting graph $G_{m,\delta}(t)$, where m vertices in $G_1(mt)$ are grouped together to a single vertex in $G_{m,\delta}(t)$, 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 will be used later on.*

The growth rules in (1.1)–(1.3) are indeed such that vertices with high degree are more likely to attract edges of new vertices. One would expect the models (a)–(c) to behave quite similarly, as is known rigorously for the scale-free behaviour, where the asymptotic degree distribution is known to be equal in models (a)–(c). As it turns out, the affine PA mechanism in (1.1)–(1.3) gives rise to power-law degree sequences. Indeed, in [23], it was proved that for model (c), the degree sequence is close to a power law with exponent $\tau = 3 + \delta/m$. For model (a) and $\delta = 0$, this was proved in [17], while in [22], power-law degree sequences for PA-models with affine PA mechanisms are proved in rather large generality. We see that, by varying the parameters $m \geq 1, \delta > -m$, we can obtain any power-law exponent $\tau > 2$, which is the reason for introducing the parameter δ in (1.1)–(1.3). However, there is no intrinsic reason for the *affine* PA mechanism. For results on PA-models in the non-affine case, see e.g., [41, 44]. In general, such models do *not* produce power laws.

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 and the verification of the prediction that distances behave similarly in various scale-free random models (see also Section 1.2 below). In the following section, we describe our precise results.

1.1 Bounds on the diameter in preferential attachment models

In this section, we present the diameter results for the PA-models (a)–(c). The *diameter* of a graph G is defined as

$$\text{diam}(G) = \max_{i,j\in G} \{\text{dist}_G(i,j) | \text{dist}_G(i,j) < \infty\}, \tag{1.4}$$

where $\mathrm{dist}_G(i,j)$ denotes the graph distance between vertices $i,j \in G$. We prove that, for all $\delta > -m$, the diameter of $G_{m,\delta}(t)$ is bounded by a constant times $\log t$. When $\delta = 0$, we adapt the argument in [15] to prove that the diameter is bounded from below by $(1-\varepsilon)\frac{\log t}{\log\log t}$. For $\delta > 0$, this lower bound is improved to a constant times $\log t$, while, for $\delta < 0$, we prove that the diameter is bounded above and below by a 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, which shall each hold for each of the models (a)–(c) simultaneously. In the results below, for a sequence of events $\{A_t\}_{t\geq 1}$, we write that A_t occurs *with high probability* (**whp**) when $\lim_{t\to\infty}\mathbb{P}(A_t) = 1$.

Theorem 1.3 (A $\log t$ upper bound on the diameter)**.** *Fix $m \geq 1$ and $\delta > -m$. Then, there exists a constant $c_1 = c_1(m,\delta) > 0$ such that* **whp***, the diameter of $G_{m,\delta}(t)$ is at most $c_1 \log t$.*

When $m = 1$, so that the graphs are in fact trees, there is a sharper result proved by Pittel [42], which, in particular, implies Theorem 1.3 for model (b). 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 proves that the diameter is at least as large, and suggests that the diameter has size $\frac{2(1+\delta)}{\gamma(2+\delta)}\log t(1+o(1))$. Scale-free trees have received substantial attention in the literature, we refer to [16, 42] and the references therein. It is not hard to see that a similar result as proved in [42] also follows for models (a) and (c). This is proved when $\delta = 0$ in [16], 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 $m = 1$.

It is not hard to extend the upper bound to $m \geq 2$. In particular, for model (b), the upper bound for $m \geq 2$ immediately follows from the upper bound for $m = 1$. For models (a) and (c), the extension is not as trivial, but the proof is fairly straightforward, and will be omitted here. To see an implication of [42] for model (a), we note that C_t, the number of connected components of $G_{1,\delta}(t)$ in model (a), has distribution $C_t = 1 + I_2 + \cdots + I_t$, 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, $C_t/\log t$ 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 $G_{1,\delta}(s_t+1)$ in model (b), apart from the fact that the vertices 1 and 2 in $G_{1,\delta}(s_t+1)$ are identified (thus creating a unique self-loop). This close connection between the two models allows one to transfer results for model (b) to model (a) when $m = 1$.

Theorem 1.4 (A $\log t$ lower bound on the diameter for $\delta > 0$)**.** *Fix $m \geq 1$ and $\delta > 0$. Then, there exists $c_2 = c_2(m,\delta) > 0$, such that* **whp***, the diameter of $G_{m,\delta}(t)$ is at least $c_2 \log t$.*

Theorems 1.3–1.4 imply that, for $\delta > 0$ and **whp**, $\mathrm{diam}(G_{m,\delta}(t)) = \Theta(\log t)$. Theorems 1.3–1.4 indicate that distances in PA-models are similar to the ones in other scale-free models for $\tau > 3$. We shall discuss this analogy in more detail below. As we shall see in Section 2.2, the proof of Theorem 1.4 also reveals that, **whp**, the average distance in $G_{m,\delta}(t)$, which is the distance between two uniformly chosen connected vertices in the graph, is also bounded from below by $c_2 \log t$.

We conjecture that, for $\delta > 0$, a limit result holds for the constant in front of the $\log t$. In its statement, we write $\mathrm{dist}_G(v_1,v_2)$ for the graph distance in the graph G between two vertices $v_1, v_2 \in [t]$. Then, the

average distance in a graph G is defined by $\mathrm{dist}_G(V_1, V_2)$ where $V_1, V_2 \in [t]$ are two uniformly chosen connected vertices.

Conjecture 1.5 (Convergence in probability for $\delta > 0$)**.** *Fix $m \geq 1$ and $\delta > 0$. Then, the diameter $\mathrm{diam}(G_{m,\delta}(t))/\log t$ and the average distance $\mathrm{dist}_G(V_1, V_2)/\log t$ converge in probability to positive and different constants.*

We now turn to the case where $\delta \in (-m, 0)$ 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 $\delta \in (-m, 0)$. Then, for every $\sigma > 1/(3-\tau)$ and with*

$$C_G = \frac{4}{|\log(\tau - 2)|} + \frac{4\sigma}{\log m}, \tag{1.7}$$

the diameter of $G_{m,\delta}(t)$ is, **whp***, bounded above by $C_G \log\log t$, as $t \to \infty$.*

In this 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 $\frac{4}{|\log(\tau-2)|}\log\log t(1 + o(1))$. When $m = 1$, Theorem 1.6 does *not* hold (see the discussion below Theorem 1.3).

We next discuss the lower bound on the diameter for $\delta \in (-m, 0)$:

Theorem 1.7 (A $\log\log t$ lower bound on the diameter)**.** *Fix $m \geq 2$ and $\delta > -m$. Then, the diameter of $G_{m,\delta}(t)$ is,* **whp***, bounded below by $\frac{\varepsilon}{\log m}\log\log t$, for all $\varepsilon \in (0, 1)$.*

Unfortunately, the proof of Theorem 1.7 does not allow for an extension to average distances, and, thus, we have no matching lower bound for this. We finally conjecture that, for $\delta \in (-m, 0)$, a limit results holds for the constant in front of the $\log\log t$:

Conjecture 1.8 (Convergence in probability for $\delta < 0$)**.** *Fix $m \geq 2$ and $\delta \in (m, 0)$. Then, the diameter $\mathrm{diam}(G_{m,\delta}(t))/\log\log t$ and the average distance $\mathrm{dist}_G(V_1, V_2)/\log\log t$ converge in probability to positive and different constants.*

1.2 Discussion of universality of distances in power-law random graphs

Theorems 1.3–1.7 prove that the diameter in PA-models with a power-law degree sequence denoted by τ undergoes a phase transition as τ changes from $\tau \in (2, 3)$ to $\tau > 3$. The results identify the order of growth of the diameter of three related models of affine PA models as the size of the graph t tends to infinity. We do not obtain the right constants. For the average or typical distances, we obtain a similar phase transition, and again the results identify the correct asymptotics for $\tau > 3$, but, for $\tau \in (2, 3)$ we miss a matching lower bound.

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 [11]. For power-law random graphs, this informal statement can be made precise by conjecturing that distances have the same leading order growth in graphs with the same power-law degree exponent. This, however, is not correct for the *diameter* of such power-law random graphs, since the diameter depends sensitively on the details of the graph, such as the proportion of vertices with degrees 1 and 2. See [29] and [34] for results showing that for the configuration model with power-law degree exponent $\tau \in (2, 3)$, the diameter can be of order $\log t$ or of order $\log\log t$ depending on the proportion of vertices with degrees 1 and 2, where t is the size of the graph. Similarly, in inhomogeneous random graphs with power-law degree exponent $\tau \in (2, 3)$ the diameter is always of order $\log t$ (see e.g.

[11]), while the average distances can be of order $\log\log t$ (see e.g. [18, 19]). Thus, we shall interpret the physicists' prediction by conjecturing that the leading order growth of the *average* distances of various power-law random graphs depends only on the power-law degree exponent $\tau \in (2,3)$.

The results on distances are most complete for the configuration model (CM), see e.g. [27, 29, 32, 33, 43]. In the CM, there are various cases depending on the tails of the degree distribution. When the degrees have infinite mean, then average distances are bounded [27], when the degrees have finite mean but infinite variance, average distances grow proportionally to $\log\log t$ [33, 43], where t is the size of the graph, while, for finite variance degrees, the average distances grow proportionally to $\log t$ [32]. Similar results for models with conditionally independent edges exist, see e.g. [11, 18, 26, 40], but particularly in the regime $\tau \in (2,3)$, the results are not that strong. 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 CM, 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 a linear growth of the number of edges, infinite mean degrees cannot arise, which explains why $\tau > 2$ for PA-models. An attempt in the direction of creating PA-models with power-law exponent $\tau \in (1,2)$ can be found in [23], where a preferential attachment model 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. This model is further studied in [8], where a wealth of results for various PA-models can be found.

There are few results on distances in PA-models. In [15], it was proved that in model (a) and for $\delta = 0$, for which $\tau = 3$, the diameter of the graph of size t is equal to $\frac{\log t}{\log\log t}(1 + o(1))$. Unfortunately, the matching result for the CM has not been proved, so that this does not allow us to verify whether the models have similar distances. The results stated above substantiate the physicists' prediction, since, for $\delta > 0$ for which $\tau \in (3, \infty)$, the average distances are of order $\log t$, while, for $\delta < 0$, for which $\tau \in (2,3)$, they are bounded above by $\log\log t$. A related result on PA-models in the spirit of [22] can be found in [20], where a similar phase transition as in this paper is proved, in the case where the number of edges grows at least $(\log t)^{1+\varepsilon}$ times as fast as the number of vertices.

It would be of interest to improve the bounds presented in this paper up to the constant in front of the $\log t$ and $\log\log t$, respectively. Due to the dynamical nature of PA-models, this is more involved for PA-models than it is for static models such as the CM and inhomogeneous random graphs.

This paper is organized as follows. In Section 2, we prove the $\log t$ lower bound for the diameter stated in Theorem 1.4. In Section 3 and Section 4, we prove the $\log\log t$ upper bound and the $\log\log t$ lower bound, on the diameter for $\delta < 0$, of Theorem 1.6 and Theorem 1.7, respectively.

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

In this section, we prove Theorem 1.4 by extending the argument in [15] from $\delta = 0$ to $\delta > 0$. We shall also extend the lower bound for $\delta = 0$ to models (b) and (c).

For model (c), denote by

$$\{g(t,j) = s\}, \quad 1 \leq j \leq m, \tag{2.1}$$

the event that at time t the j^{th} edge of vertex t is attached to the earlier vertex $s < t$. For models (a) and (b), this event means that in $\{G_{1,\delta'}(mt)\}$ 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 (2.1) increases the preference for vertex s, and hence decreases (in a relative way) the preference for the vertices u, $1 \leq u \leq t$, $u \neq s$. It should be intuitively clear that another way of expressing this effect is to

say that, for different $s_1 \neq s_2$, 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 $n_s \geq 1$ and $i = 1, \ldots, n_s$, we denote by

$$E_s = \bigcap_{i=1}^{n_s} \{g(t_i, j_i) = s\}, \tag{2.2}$$

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 $k \geq 1$ and all possible choices of t_i, j_i, the events E_s, for different s, are negatively correlated:

Lemma 2.1 (Negative correlation of attachment events)**.** *For distinct $s_1, s_2, \ldots, s_k$,*

$$\mathbb{P}\Big(\bigcap_{i=1}^{k} E_{s_i}\Big) \leq \prod_{i=1}^{k} \mathbb{P}(E_{s_i}). \tag{2.3}$$

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 (2.3) holds for all k and all choices of t_i, j_i such that $\max_{i,s} m(t_i - 1) + j_i \leq e$, where induction is performed with respect to e. 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 (2.3) holds for all k and all choices of t_i, j_i such that $\max_{i,s} m(t_i - 1) + j_i \leq e - 1$. Clearly, for k and t_i, j_i such that $\max_{i,s} m(t_i - 1) + j_i \leq e - 1$, the bound follows from the induction hypothesis, so we may restrict attention to the case that $\max_{i,s} m(t_i - 1) + j_i = e$. 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, j_i such that $t_i = t, j_i = j$, 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, j_i such that $t_i = t, j_i = j$. Denote by

$$E_s' = \bigcap_{i=1:(t_i,j_i)\neq(t,j)}^{n_s} \{g(t_i, j_i) = s\}, \tag{2.4}$$

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

$$\bigcap_{i=1}^{k} E_{s_i} = \{g(t, j) = s\} \cap E_s' \cap \bigcap_{i=1:s_i\neq s}^{k} E_{s_i}. \tag{2.5}$$

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

$$\mathbb{P}\Big(\bigcap_{i=1}^{k} E_{s_i}\Big) \leq \mathbb{E}\Big[I[E_s' \cap \bigcap_{i=1:s_i\neq s}^{k} E_{s_i}]\mathbb{P}_{e-1}(g(t, j) = s)\Big], \tag{2.6}$$

where $\mathbb{P}_{e-1}$ denotes the conditional probability given the edge attachments up to the $(e-1)^{\text{st}}$ edge connection, and where, for an event A, $I[A]$ denotes the indicator of A.

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

$$\mathbb{P}_{e-1}(g(t, j) = s) = \frac{D_s(t-1) + \delta}{(2m + \delta)(t-1)}. \tag{2.7}$$

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

$$D_s(t-1) = m + \sum_{(t',j'):t'\le t-1} I[g(t',j') = s]. \tag{2.8}$$

We note that each of the terms in (2.8) 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 $m+\delta \ge 0$, that,

$$\begin{aligned}(2m+\delta)(t-1)\mathbb{P}\Big(\bigcap_{i=1}^{k} E_{s_i}\Big) \quad &\le (m+\delta)\mathbb{P}(E'_s)\prod_{i=1:s_i\neq s}^{k}\mathbb{P}(E_{s_i}) \\ &+ \sum_{(t',j'):t'\le t-1}\mathbb{P}(E'_s\cap\{g(t',j')=s\})\prod_{i=1:s_i\neq s}^{k}\mathbb{P}(E_{s_i}).\end{aligned} \tag{2.9}$$

We can recombine to obtain

$$\mathbb{P}\Big(\bigcap_{i=1}^{k} E_{s_i}\Big) \le \mathbb{E}\Big[I[E'_s]\frac{D_s(t-1)+\delta}{(2m+\delta)(t-1)}\Big]\prod_{i=1:s_i\neq s}^{k}\mathbb{P}(E_{s_i}), \tag{2.10}$$

and the advancement is completed when we note that

$$\mathbb{E}\Big[I[E'_s]\frac{D_s(t-1)+\delta}{(2m+\delta)(t-1)}\Big] = \mathbb{P}(E_s). \tag{2.11}$$

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 $\{G_{1,\delta'}(t)\}_{t\ge 1}$. □

We next give the probabilities of E_s when $n_s \le 2$; we omit the proof, since it is a simple adaptation to that in [15].

Lemma 2.2 (Connections in PA-models)**.** *There exist absolute constants M_1, M_2, such that (i) for each $1 \le j \le m$, and $t > s$,*

$$\mathbb{P}\Big(g(t,j) = s\Big) \le \frac{M_1}{t^{1-a}s^a}, \tag{2.12}$$

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

$$\mathbb{P}\Big(g(t_1,j_1) = s, g(t_2,j_2) = s\Big) \le \frac{M_2}{(t_1t_2)^{1-a}s^{2a}}, \tag{2.13}$$

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

We combine the results of Lemmas 2.1 and 2.2 into the following corollary, yielding an upper bound for the probability of the existence of a path. In its statement, we call a path $\Gamma = (s_0, s_1, \dots, s_l)$ self-avoiding when $s_i \neq s_j$ for all $0 \le i < j \le l$. We use the notation $x \wedge y = \min(x,y)$ and $x \vee y = \max(x,y)$. Again, we omit the proof (for details, see [15]).

Corollary 2.3 (Path probabilities in PA-models)**.** *Let $\Gamma = (s_0, s_1, \dots, s_l)$ be a self-avoiding path of length l consisting of the $l+1$ unordered vertices $s_0, s_1, \dots, s_l$, then there exists an absolute constant $C > 0$ such that*

$$\mathbb{P}\Big(\Gamma \in G_{m,\delta}(t)\Big) \le (m^2C)^l\prod_{i=0}^{l-1}\frac{1}{(s_i\wedge s_{i+1})^a(s_i\vee s_{i+1})^{1-a}}. \tag{2.14}$$

2.1 Lower bound on the diameter for $\delta = 0$

It follows from (2.14) that for $\delta = 0$,

$$\mathbb{P}\Big(\Gamma \in G_{m,\delta}(t)\Big) \leq (m^2 C)^l \prod_{i=0}^{l-1} \frac{1}{\sqrt{s_i s_{i+1}}}. \tag{2.15}$$

The further proof that (2.15) implies that for $\delta \geq 0$,

$$L = \frac{\log t}{\log(3Cm^2 \log t)}, \tag{2.16}$$

is a lower bound for the diameter of $G_{m,\delta}(t)$, is identical to the proof of [15, Theorem 5, p. 14], with n replaced by t. This extends the lower bound for $\delta = 0$ for model (a) in [15] to models (b)–(c). □

2.2 The lower bound on distances for $\delta > 0$

We next improve the bound in the previous section in the case when $\delta > 0$, in which case $a = m/(2m+\delta) < 1/2$. From the above discussion, we conclude that

$$\mathbb{P}\Big(\mathrm{dist}_{G_{m,\delta(t)}}(1,t) = k\Big) \leq c^k \sum_{\vec{s}} \prod_{j=0}^{k-1} \frac{1}{(s_j \wedge s_{j+1})^a (s_j \vee s_{j+1})^{1-a}}, \tag{2.17}$$

where $c = m^2 C$, and where the sum is over $\vec{s} = (s_0, \ldots, s_k)$ with $s_k = t$, $s_0 = 1$, $s_l \geq 1$ for all $l = 1, \ldots, k-1$ and $s_l \neq s_n$ for all $l \neq n$. Define

$$f_k(i,t) = \sum_{\vec{s}} \prod_{j=0}^{k-1} \frac{1}{(s_j \wedge s_{j+1})^a (s_j \vee s_{j+1})^{1-a}}, \tag{2.18}$$

where now the sum is over $\vec{s} = (s_0, \ldots, s_k)$ with $s_k = t, s_0 = i$, $s_l \geq 1$ for all $l = 1, \ldots, k-1$ and $s_l \neq s_n$ for all $l \neq n$, so that

$$\mathbb{P}\Big(\mathrm{dist}_{G_{m,\delta(t)}}(i,t) = k\Big) \leq c^k f_k(i,t). \tag{2.19}$$

We study the function $f_k(i,t)$ in the following lemma:

Lemma 2.4 (A bound on f_k)**.** *Fix $a < 1/2$. Then, for every $b > a$ such that $a + b < 1$, there exists a $C_{a,b} > 0$ such that, for every $1 \leq i < t$ and all $k \geq 1$,*

$$f_k(i,t) \leq \frac{C_{a,b}^k}{i^b t^{1-b}}. \tag{2.20}$$

Proof. We prove the lemma using induction on $k \geq 1$. To initialize the induction hypothesis, we note that, for $1 \leq i < t$ and every $b \geq a$,

$$f_1(i,t) = \frac{1}{(i \wedge t)^a (i \vee t)^{1-a}} = \frac{1}{i^a t^{1-a}} = \frac{1}{t}\Big(\frac{t}{i}\Big)^a \leq \frac{1}{t}\Big(\frac{t}{i}\Big)^b = \frac{1}{i^b t^{1-b}}. \tag{2.21}$$

This initializes the induction hypothesis as long as $C_{a,b} \geq 1$.

To advance the induction hypothesis, note that we have the recursion relation

$$f_k(i,t) = \sum_{s=1}^{i-1} \frac{1}{s^a i^{1-a}} f_{k-1}(s,t) + \sum_{s=i+1}^{\infty} \frac{1}{i^a s^{1-a}} f_{k-1}(s,t). \tag{2.22}$$

We now bound each of these two contributions, making use of the induction hypothesis. For the first sum, we bound

$$\sum_{s=1}^{i-1} \frac{1}{s^a i^{1-a}} f_{k-1}(s,t) \leq C_{a,b}^{k-1} \sum_{s=1}^{i-1} \frac{1}{s^a i^{1-a}} \frac{1}{s^b t^{1-b}} = \frac{C_{a,b}^{k-1}}{i^{1-a} t^{1-b}} \sum_{s=1}^{i-1} \frac{1}{s^{a+b}} \leq \frac{1}{1-a-b} \frac{C_{a,b}^{k-1}}{i^b t^{1-b}}, \tag{2.23}$$

since $a+b<1$. For the second sum, we bound

$$\sum_{s=i+1}^{\infty} \frac{1}{i^a s^{1-a}} f_{k-1}(s,t) \leq C_{a,b}^{k-1} \sum_{s=i+1}^{\infty} \frac{1}{i^a s^{1-a}} \frac{1}{s^b t^{1-b}} = \frac{C_{a,b}^{k-1}}{i^a t^{1-b}} \sum_{s=i+1}^{\infty} \frac{1}{s^{1+b-a}} \leq \frac{1}{b-a} \frac{C_{a,b}^{k-1}}{i^b t^{1-b}}, \tag{2.24}$$

since $1+b-a>1$. We conclude that

$$f_k(i,t) \leq \frac{C_{a,b}^{k-1}}{i^b t^{1-b}} \Big(\frac{1}{b-a} + \frac{1}{1-a-b} \Big) \leq \frac{C_{a,b}^{k}}{i^b t^{1-b}}, \tag{2.25}$$

when

$$C_{a,b} = \frac{1}{b-a} + \frac{1}{1-a-b} \geq 1. \tag{2.26}$$

This advances the induction hypothesis, and completes the proof. □

Using Lemma 2.4 and (2.19), we obtain that

$$\mathbb{P}\Big(\mathrm{dist}_{G_{m,\delta(t)}}(1,t) = k \Big) \leq \frac{(cC_{a,b})^k}{t^{1-b}}. \tag{2.27}$$

As a result, we obtain that

$$\mathbb{P}\Big(\mathrm{diam}(G_{m,\delta}(t)) \leq k \Big) \leq \mathbb{P}\Big(\mathrm{dist}_{G_{m,\delta(t)}}(1,t) \leq k \Big) \leq \frac{(cC_{a,b})^{k+1}}{t^{1-b}(cC_{a,b}-1)} = o(1), \tag{2.28}$$

whenever $k \leq \frac{1-b}{\log(cC_{a,b})} \log t$. We conclude that there exists $c_2 = c_2(m,\delta)$ such that, with high probability $\mathrm{diam}\Big(G_{m,\delta}(t)\Big) \geq c_2 \log t$. □

We next extend the above discussion to *average* distances.

Lemma 2.5 (Average distances for $\delta > 0$)**.** *Fix $m \geq 1$ and $\delta > 0$. Let $H_t = \mathrm{dist}_t(A_1, A_2)$ be the distance between two uniformly chosen vertices. Then, for $c_2 = c_2(m,\delta) > 0$ sufficiently small,* **whp**, *$H_t \geq c_2 \log t$.*

Proof. For $c_2 = c_2(m,\delta) > 0$, define

$$B_t \equiv \#\{i,j \in [t] : i < j : \mathrm{dist}_{G_{m,\delta(t)}}(i,j) \leq c_2 \log t\}, \tag{2.29}$$

where $\#\{A\}$ denotes the cardinality of A.

By Lemma 2.4, with $K = \log(cC_{a,b} \vee 2)$ and $a < b < 1-a$, and for all $1 \leq i < j \leq t$,

$$\mathbb{P}\Big(\mathrm{dist}_{G_{m,\delta(t)}}(i,j) = k \Big) \leq c^k f_k(i,j) \leq \frac{e^{Kk}}{i^{1-b} j^b}. \tag{2.30}$$

As a result, we obtain that

$$\mathbb{P}\Big(\mathrm{dist}_{G_{m,\delta(t)}}(i,j) \leq c_2 \log t \Big) \leq \frac{t^{Kc_2}}{i^{1-b} j^b} \frac{e^K}{e^K - 1}, \tag{2.31}$$

and thus, using also $\sum_{i=1}^{j-1} i^{b-1} \leq j^b/b$,

$$\mathbb{E}[B_t] \leq O(1) \sum_{1\leq i<j\leq t} \frac{t^{Kc_2}}{i^{1-b}j^b} = O(t^{Kc_2+1}). \tag{2.32}$$

It now suffices to note that

$$\mathbb{P}(H_t \leq c_2 \log t) = \mathbb{E}\big[I[\mathrm{dist}_{G_{m,\delta(t)}}(A_1, A_2) \leq c_2 \log t]\big] \leq \frac{2\mathbb{E}[B_t] + t}{t^2} = o(1), \tag{2.33}$$

by (2.32), for every $c_2 > 0$ such that $Kc_2 + 1 < 2$. □

Note that (2.16) is also a lower bound on average distances in case $\delta = 0$, which can be proven as above.

3 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 3.1, we bound 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 [18] or [43] used to prove bounds on the average distance for the inhomogeneous random graph and the configuration model, respectively, but substantial adaptations are necessary to deal with preferential attachment. After this, in Theorem 3.6, 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 PA-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. We make use of some technical results, stated in the appendix.

3.1 The diameter of the core

We recall that $\tau = 3 + \delta/m$, so that $-m < \delta < 0$ corresponds to $\tau \in (2, 3)$. We take $\sigma > 1/(3-\tau) = -m/\delta > 1$ and define the *core* Core_t to be

$$\mathrm{Core}_t = \big\{i \in [t] : D_i(t) \geq (\log t)^\sigma\big\}, \tag{3.1}$$

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

For $A \subseteq [t]$, we write

$$\mathrm{diam}_t(A) = \max_{i,j\in A} \mathrm{dist}_{G_{m,\delta(t)}}(i, j). \tag{3.2}$$

Then, $\mathrm{diam}_{2t}(\mathrm{Core}_t)$ is bounded in the following theorem:

Theorem 3.1 (The diameter of the core)**.** *Fix $m \geq 2$ and $\delta \in (-m, 0)$. For every $\sigma > 1/(3-\tau)$,* **whp***,*

$$\mathrm{diam}_{2t}(\mathrm{Core}_t) \leq (1 + o(1))\frac{4\log\log t}{|\log(\tau - 2)|}. \tag{3.3}$$

The proof of Theorem 3.1 is divided into several smaller steps. We start by proving that the diameter $\mathrm{diam}_{2t}(\mathrm{Inner}_t)$, where

$$\mathrm{Inner}_t = \big\{i \in [t] : D_i(t) \geq u_1\big\}, \quad \text{and where} \quad u_1 = t^{\frac{1}{2(\tau-1)}}(\log t)^{-\frac{1}{2}}, \tag{3.4}$$

is, **whp**, bounded. The choice of u_1 is a technical one: u_1 is the largest value l so that, **whp**, the total degree of vertices with degree exceeding l can be bounded from below by $tl^{2-\tau}$, see Lemma A.1. In

Proposition 3.2, we will show that the diameter of Inner_t is bounded. After this, we will show that the distance from any vertex in the core Core_t to the inner core Inner_t can be bounded by a fixed constant times $\log\log t$. This also shows that $\text{diam}_{2t}(\text{Core}_t)$ is bounded by a different constant times $\log\log t$. We now give the details.

Proposition 3.2 (The diameter of the inner core)**.** *Fix $m \geq 2$ and $\delta \in (-m, 0)$. Then* **whp**,

$$\text{diam}_{2t}(\text{Inner}_t) \leq \frac{2(\tau-1)}{3-\tau} + 6. \tag{3.5}$$

Proof. We first introduce the important notion of a t-connector between a vertex $i \in [t]$ and a set of vertices $A \subseteq [t]$. This notion will play a crucial role throughout the proof. We say that the vertex $j \in [2t] \setminus [t]$ is a *t-connector* between i and A if one of the *first two* edges incident to j connects to i and the other of the *first two* edges 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 $G_{m,\delta}(2t)$ 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 $G_{m,\delta}(t)$, the probability that $j \in [2t] \setminus [t]$ is a t-connector for i and A is at least

$$\frac{(D_A(t) + \delta|A|)(D_i(t) + \delta)}{[2t(2m+\delta)]^2} \geq \frac{\eta D_A(t) D_i(t)}{t^2}, \tag{3.6}$$

where in the inequality, we use that, almost surely, $D_i(t) \geq m$, and we let $\eta = (m+\delta)^2/(2m(2m+\delta))^2 > 0$, while, for any $A \subseteq [t]$, we write

$$D_A(t) = \sum_{i \in A} D_i(t). \tag{3.7}$$

Note that for fixed $j \in [2t] \setminus [t]$ the lower bound (3.6) holds independently of the fact whether the other vertices are t-connectors or not.

We now give a coupling proof which shows that a subset of size $n_t = \lfloor \sqrt{t} \rfloor$ of the set Inner_t has, **whp**, a bounded diameter. The last line of the proof of Lemma A.1 in the appendix shows that, **whp**, Inner_t contains at least $\sqrt{t}$ vertices. Denote the first $\lfloor \sqrt{t} \rfloor$ vertices of Inner_t by I. For each pair $i_1, i_2 \in I$ and each $j \in [2t] \setminus [t]$, the probability that j is a t-connector for i_1, i_2 is, by (3.6), at least

$$\frac{\eta u_1^2}{t^2} = \frac{\eta t^{\frac{1}{\tau-1}}}{t^2 \log t} \geq \frac{t^{\frac{1}{\tau-1}-2}}{\log^2 t} = q_t, \tag{3.8}$$

independently of the fact whether the other vertices are t-connectors or not. In the coupling we intend to compare the set I and all pairs of vertices of the set I, which are t-connected by some $j \in [2t] \setminus [t]$ with a so-called *multinomial* random graph H_{n_t}. The graph H_{n_t} has n_t vertices and we identify the $e_t = n_t(n_t-1)/2 \sim t/2$ pairs of vertices, which we number from 1 to e_t in an arbitrary order, with e_t cells of a multinomial experiment with t trials and probabilities given by

$$p_k = q_t, \quad 1 \leq k \leq e_t, \qquad p_0 = 1 - e_t q_t. \tag{3.9}$$

We can represent the t trials by independent random vectors $N_1, N_2, \ldots, N_t$, where

$$N_j = (N_{j,1}, N_{j,2}, \ldots, N_{j,e_t}), \quad 1 \leq j \leq t, \tag{3.10}$$

with distribution

$$\mathbb{P}(N_j = 1_i) = q_t, \quad \mathbb{P}(N_j = 0) = 1 - e_t q_t, \tag{3.11}$$

where 1_i is the i^{th} unit vector of length e_t, and 0 the null vector. If cell k of the multinomial experiment is not empty, i.e., if $\sum_{j=1}^{t} N_{j,k} > 0$, then we draw the edge with number k in the graph H_{n_t}, if the cell is empty then this edge is left out. Note that cell 0 is just an overflow cell, which counts the number of trials that not resulted in one of the cells $1, 2, \ldots, e_t$.

By the statement in (3.8) the distance in $G_{m,\delta}(2t)$ between any two vertices in I is at most two times the distance between the corresponding vertices in H_{n_t}. In Lemma A.2 of the appendix we will show that the diameter of H_{n_t} is at most the diameter of a *uniform* Erdős-Rényi graph $G(n_t, m_t)$, with n_t vertices and m_t edges, where

$$m_t = \frac{1}{2} e_t \Big(1 - (1 - q_t)^t\Big). \tag{3.12}$$

From [35, Section 1.4] we conclude that the above mentioned uniform Erdős-Rényi graph $G(n_t, m_t)$ is asymptotically equivalent with the classical *binomial* Erdős-Rényi graph $G(n_t, \lambda_t)$, where the edge probability λ_t is defined by

$$\lambda_t = \frac{1}{2}\Big(1 - (1 - q_t)^t\Big) \sim \frac{t^{\frac{1}{\tau-1}-1}}{2\log^2 t}. \tag{3.13}$$

Next, we show that $\operatorname{diam}(G(n_t, \lambda_t))$ is, **whp**, bounded by $\frac{\tau-1}{3-\tau} + 1$. For this we use the result in [9, 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 $\operatorname{diam}(G(n,p)) = d$, **whp**. In our case, $n = n_t = \lfloor t^{1/2} \rfloor$ and $p = \lambda_t$, which implies that, **whp**, $\frac{\tau-1}{3-\tau} < d \le \frac{\tau-1}{3-\tau} + 1$. We therefore obtain that the diameter of I in $G_{m,\delta}(2t)$ is, **whp**, bounded by

$$\operatorname{diam}_{2t}(I) \le \frac{2(\tau-1)}{3-\tau} + 2. \tag{3.14}$$

We finally show that for any $i \in \operatorname{Inner}_t \setminus I$, the probability that there does not exist a t-connector connecting i and I is small. Indeed, since $D_I(t) \ge \sqrt{t} u_1$ and $D_i(t) \ge u_1$, the mentioned probability is bounded above by

$$\Big(1 - \frac{\eta D_I(t) D_i(t)}{t^2}\Big)^t \le \exp\Big\{-\frac{\eta D_I(t) D_i(t)}{t}\Big\} \le \exp\Big\{-\frac{\eta u_1^2}{\sqrt{t}}\Big\} \le \exp\Big\{-\frac{\eta t^{\frac{1}{\tau-1}-\frac{1}{2}}}{\log t}\Big\} = o(t^{-1}), \tag{3.15}$$

for $\tau < 3$. Thus, **whp**, such a vertex i does not exist. This proves that **whp** the distance between any vertex $i \in \operatorname{Inner}_t \setminus I$ and I is bounded by 2, and, together with the above bound on $\operatorname{diam}_{2t}(I)$ we thus obtain (3.5). □

Proposition 3.3 (Distance from the core to the inner core)**.** *Fix $m \ge 2$ and $\delta \in (-m, 0)$. With high probability, the inner core* Inner_t *can be reached from any vertex in the core* Core_t *using no more than* $\frac{2\log\log t}{|\log(\tau-2)|}$ *edges in $G_{m,\delta}(2t)$. More precisely,* **whp**,

$$\max_{i \in \operatorname{Core}_t} \min_{j \in \operatorname{Inner}_t} \operatorname{dist}_{G_{m,\delta}(2t)}(i,j) \le \frac{2\log\log t}{|\log(\tau-2)|}. \tag{3.16}$$

Proof. For $k \ge 1$, we define

$$\mathcal{N}^{(k)} = \{i \in [t] : D_i(t) \ge u_k\}, \tag{3.17}$$

with u_1 defined in (3.4), and where we define u_k, for $k \ge 2$, recursively, so that for any vertex $i \in [t]$ with degree at least u_k, the probability that there is no t-connector for the vertex i and the set $\mathcal{N}^{(k-1)}$, conditionally on $G_{m,\delta}(t)$, is tiny. According to (3.6) and (A.1) in the appendix, this probability is at most

$$\Big(1 - \frac{\eta d_{\mathcal{N}^{(k-1)}} D_i(t)}{t^2}\Big)^t \le \exp\Big\{-\frac{\eta B t \big(u_{k-1}\big)^{2-\tau} u_k}{t}\Big\} = o(t^{-2}), \tag{3.18}$$

for some $B > 0$, when we define

$$u_k = D \log t (u_{k-1})^{\tau-2}, \tag{3.19}$$

with D exceeding $2(\eta B)^{-1}$. The following lemma identifies u_k:

Lemma 3.4 (Identification of u_k)**.** *For each $k \in \mathbb{N}$,*

$$u_k = D^{a_k} (\log t)^{b_k} t^{c_k}, \tag{3.20}$$

where

$$a_k = \frac{1-(\tau-2)^{k-1}}{3-\tau}, \qquad b_k = \frac{1-(\tau-2)^{k-1}}{3-\tau} - \frac{1}{2}(\tau-2)^{k-1}, \qquad c_k = \frac{(\tau-2)^{k-1}}{2(\tau-1)}. \tag{3.21}$$

Proof. We leave the straightforward induction proof to the reader. □

Then, the key step in the proof of Proposition 3.3 is the following lemma:

Lemma 3.5 (Connectivity between $\mathcal{N}^{(k-1)}$ and $\mathcal{N}^{(k)}$)**.** *Fix $m \geq 2$ and $\delta \in (-m, 0)$. Then, uniformly in k, the probability that there exists an $i \in \mathcal{N}^{(k)}$ that is not at distance at most two from $\mathcal{N}^{(k-1)}$ in $G_{m,\delta}(2t)$ is $o(t^{-1})$.*

Proof. It follows from (3.19) that the probability in the statement is by Boole's inequality bounded by

$$t \exp\Big(- \frac{\eta B t [u_{k-1}]^{2-\tau} u_k}{t} \Big) = t \cdot o(t^{-2}) = o(t^{-1}), \tag{3.22}$$

where we have used Lemma A.1 to get a lower bound to the degree of $\mathcal{N}^{(k-1)}$. □

We now complete the proof of Proposition 3.3. Fix

$$k^* = \Big\lfloor \frac{\log \log t}{|\log(\tau-2)|} \Big\rfloor. \tag{3.23}$$

As a result of Lemma 3.5, we have that the distance between $\mathcal{N}^{(k^*)}$ and $\mathrm{Inner}_t = \mathcal{N}^{(1)}$ is at most $2k^*$. Therefore, Proposition 3.3 follows when we can show that

$$\mathrm{Core}_t = \{i : D_i(t) \geq (\log t)^\sigma\} \subseteq \mathcal{N}^{(k^*)} = \{i : D_i(t) \geq u_{k^*}\}, \tag{3.24}$$

so that it suffices to prove that $(\log t)^\sigma \geq u_{k^*}$, for any $\sigma > 1/(3-\tau)$. This follows trivially for t large from the explicit representation of u_{k^*} given by Lemma 3.4. □

Proof of Theorem 3.1. We note that **whp**

$$\mathrm{diam}_{2t}(\mathrm{Core}_t) \leq \frac{2(\tau-1)}{3-\tau} + 6 + 2k^*, \tag{3.25}$$

where k^* is given in (3.23), and where we have made use of Proposition 3.2. This proves Theorem 3.1. □

3.2 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 is the following theorem:

Theorem 3.6 (Connecting the periphery to the core)**.** *Fix $m \geq 2$ and $\delta \in (-m, 0)$. For every $\sigma > 1/(3-\tau)$,* **whp***, the maximal distance between any vertex and Core_t in $G_{m,\delta}(2t)$ is bounded from above by $2\sigma \log\log t/\log m$.*

Together with Theorem 3.1, Theorem 3.6 proves the main result in Theorem 1.6.

The proof of Theorem 3.6 consists of two key steps. The first key step in Proposition 3.7 states that the distance between any vertex in $[t]$ and the core Core_t is bounded by a constant times $\log\log t$. The second key step in Proposition 3.10 shows that the distance between any vertex in $[2t] \setminus [t]$ and $[t]$ is bounded by another constant times $\log\log t$.

Proposition 3.7 (Connecting half of the periphery to the core)**.** *Fix $m \geq 2$ and $\delta \in (-m, 0)$. For every $\sigma > 1/(3-\tau)$,* **whp***, the distance between any vertex in $[t]$ and the core Core_t in $G_{m,\delta}(2t)$ is bounded from above by $\sigma \log\log t/\log m$.*

Proof. We start from a vertex $i \in [t]$ and will show that the probability that the distance between i and Core_t is at least $\sigma \log\log t/\log m$ is $o(t^{-1})$. 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* of vertex i.

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 its k-exploration tree before it hits the core is small, as we prove now:

Lemma 3.8 (A bound on the probability of multiple collisions)**.** *Fix $m \geq 2$ and $\delta \in (-m, 0)$. Fix $l \geq 1$, $b \in (0, 1]$ and take $k \leq \sigma \log\log t/\log m$. Then, for every vertex $i \in [t]$, the probability that its k-exploration tree has at least l collisions before it hits $\mathrm{Core}_t \cup [t^b]$ is bounded above by*

$$m^l (\log t)^{2\sigma l} \Big/ t^{bl}. \tag{3.26}$$

Proof. Take $i \in [t] \setminus [t^b]$ 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

$$\sum_{v \in \mathcal{T}_i^{(k)}} \frac{D_v(t) + \delta}{t^b} \leq \sum_{v \in \mathcal{T}_i^{(k)}} \frac{(\log t)^\sigma}{t^b} \leq \frac{m^{k+1} (\log t)^\sigma}{t^b}, \tag{3.27}$$

where the bound follows from $\delta < 0$ and $\#\{v \in \mathcal{T}_i^{(k)}\} \leq m^{k+1}$. For general l this upper bound becomes:

$$\left(\frac{m^{k+1} (\log t)^\sigma}{t^b} \right)^l. \tag{3.28}$$

When $k = \sigma \log\log t/\log m$, we have that $m^{kl} = (\log t)^{\sigma l}$. Therefore, the claim in Lemma 3.8 holds. □

We next prove that there exists a $b > 0$ such that, **whp**, $[t^b]$ is a subset of the core. Note that in this lemma the conditions $m \geq 2$ or $\delta \in (-m, 0)$ are not necessary.

Lemma 3.9 (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 every $\sigma > 1/(3-\tau)$. As a result,* **whp**, $[t^b] \subseteq \mathrm{Core}_t$.

We defer the proof of Lemma 3.9 to Section A.3 of the appendix. Now we are ready to complete the proof of Proposition 3.7:

Proof of Proposition 3.7. By combining Lemmas 3.8 and 3.9, the probability that there exists an $i \in [t]$ 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

$$m^l \sum_{i=1}^{t} (\log t)^{2\sigma l} \Big/ t^{bl} = m^l (\log t)^{2\sigma l} t^{-bl+1} = o(1), \tag{3.29}$$

when $l > \frac{1}{b}$. When the k-exploration tree hits the core, then we are done. 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 discarding at most l vertices and their complete offspring, and the size of the pruned tree has size at least m^{k-l}.

When $k = \sigma \log\log t / \log m - 2$, $m^{k-l} \geq (\log t)^{\sigma + o(1)}$. The total weight of the core is, by (A.1) in the appendix, at least

$$\sum_{i \in \mathrm{Core}_t} (D_i(t) + \delta) \geq Bt(\log t)^{-(\tau-2)\sigma}, \tag{3.30}$$

for some $B > 0$. The probability that there does not exist a t-connector between the k-exploration tree and the core is, by (3.6) and (A.1) in the appendix, bounded above by

$$\exp\left\{ -\frac{\eta B t (\log t)^{-(\tau-2)\sigma} (\log t)^{\sigma + o(1)}}{t} \right\} = o(t^{-1}), \tag{3.31}$$

since $\sigma > 1/(3-\tau)$. This completes the proof. □

Proposition 3.10 (Connecting the remaining periphery)**.** *Fix $m \geq 2$ and $\delta \in (-m, 0)$. For every $\sigma > 1/(3-\tau)$,* **whp**, *the maximal distance between any vertex and $[t]$ in $G_{m,\delta}(2t)$ is bounded from above by $\sigma \log\log t / \log m$.*

Proof. Take $k = \sigma \log\log t / \log m - 1$, and $j \in [2t] \setminus [t]$ with distance larger than k to the set of vertices $[t]$. We now apply Lemma 3.8 with t replaced by $2t$ and letting $l = 2$ and b such that $(2t)^b = t$, to conclude that with probability exceeding $1 - o(1)$, the k-exploration tree of j has at most 1 collision before it hits $\mathrm{Core}_{2t} \cup [t]$. We can hence conclude that with probability exceeding $1 - o(1)$, there are at least $m_k = (m-1)m^{k-1}$ vertices in $[2t] \setminus [t]$ at distance precisely equal to k from our starting vertex j. Denote these vertices by $i_1, \ldots, i_{m_k}$. We consider case (c), the proof for (a) and (b) is similar. Note that, uniformly in $s \in [2t] \setminus [t]$,

$$\frac{\sum_{i=1}^{t} (D_i(s) + \delta)}{(2m + \delta)s} \geq \frac{1}{2}. \tag{3.32}$$

Hence,

$$\mathbb{P}\Big(\nexists l \in [m_k] \text{ such that } \mathrm{dist}_{G_{m,\delta}(2t)}(i_l, \mathrm{Core}_{2t} \cup [t]) > 1 \Big) \leq 2^{-m_k} = o(t^{-1}), \tag{3.33}$$

since $m_k = \frac{m-1}{m^2}(\log t)^\sigma$, with $\sigma > 1/(3-\tau) > 1$. Therefore, any vertex $j \in [2t] \setminus [t]$ is, **whp**, within distance $k+1$ from $\mathrm{Core}_{2t} \cup [t]$. Proposition A.3 shows that, **whp** the set $\mathrm{Core}_{2t} \subseteq [t]$, so that, **whp**, $\mathrm{Core}_{2t} \cup [t] = [t]$ and the proposition follows. □

Proof of Theorem 3.6. Proposition 3.10 states that **whp** every vertex in $G_{m,\delta}(2t)$ is within distance $\sigma \log\log t/\log m$ of $[t]$ and Proposition 3.7 states that **whp** every vertex in $[t]$ is at most distance $\sigma \log\log t/\log m$ from the core Core_t. This shows that every vertex in $G_{m,\delta}(2t)$ is **whp** within distance $2\sigma \log\log t/\log m$ from the core. □

Proof of Theorem 1.6. Theorem 3.6 states that every vertex in $G_{m,\delta}(2t)$ is within distance $\frac{2\sigma \log\log t}{\log m}$ of the core Core_t. Theorem 3.1 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 $G_{m,\delta}(2t)$ is at most $C_G \log\log t$, where C_G is given in (1.7), because we can choose any $\sigma > 1/(3-\tau)$. This completes the proof of Theorem 1.6. □

4 A $\log\log t$ lower bound on the diameter: Proof of Theorem 1.7

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 $\mathcal{T}_i^{(k)}$ of vertices $i \in [2t]\backslash[t]$ as defined above. We shall call the tree $\mathcal{T}_i^{(k)}$ *proper* if the following conditions hold:

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

When such a tree exists in $G_{m,\delta}(2t)$ 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 $G_{m,\delta}(2t)$, 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 $G_{m,\delta}(2t)$ 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 $G_{m,\delta}(2t)$, i.e.,

$$Z_{m,\delta}^{(k)}(2t) = \sum_{\mathcal{T}\in\mathfrak{T}_m^k(2t)} I[\mathcal{T} \subseteq G_{m,\delta}(2t) \text{ and } \mathcal{T} \text{ is proper}]. \tag{4.1}$$

Here the event that all edges of $\mathcal{T}$ have been formed in $G_{m,\delta}(2t)$ is denoted by $\mathcal{T} \subseteq G_{m,\delta}(2t)$.

In Section 4.1 we will investigate the first moment of $Z_{m,\delta}^{(k)}(2t)$ and prove the following:

Proposition 4.1 (Expected number of proper trees tends to infinity)**.** *Fix $m \geq 2$ and $\delta > -m$. Let $k = \frac{\varepsilon}{\log m}\log\log t$, with $0 < \varepsilon < 1$. Then*

$$\lim_{t\to\infty} \mathbb{E}\left[Z_{m,\delta}^{(k)}(2t)\right] = \infty. \tag{4.2}$$

The variance of $Z_{m,\delta}^{(k)}(2t)$ will be the subject of Section 4.2, where we will prove the following:

Proposition 4.2 (Concentration of the number of proper trees)**.** *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,*

$$\mathrm{Var}\left(Z^{(k)}_{m,\delta}(2t)\right) \leq c_{m,\delta}\frac{(\log t)^2}{t}\mathbb{E}\left[Z^{(k)}_{m,\delta}(2t)\right]^2 + \mathbb{E}\left[Z^{(k)}_{m,\delta}(2t)\right]. \tag{4.3}$$

We use these two propositions to prove Theorem 1.7:

Proof of Theorem 1.7. We first use the Chebychev inequality to obtain that

$$\mathbb{P}\left(\mathrm{diam}(G_{m,\delta}(2t)) < k\right) \leq \mathbb{P}\left(Z^{(k)}_{m,\delta}(2t) = 0\right) \leq \frac{\mathrm{Var}\left(Z^{(k)}_{m,\delta}(2t)\right)}{\mathbb{E}\left[Z^{(k)}_{m,\delta}(2t)\right]^2}. \tag{4.4}$$

By Proposition 4.2, the right-hand side of (4.4) is, for some constant $c_{m,\delta} > 0$, at most

$$c_{m,\delta}\frac{(\log t)^2}{t} + \frac{1}{\mathbb{E}\left[Z^{(k)}_{m,\delta}(2t)\right]} = o(1), \tag{4.5}$$

by Proposition 4.1. □

4.1 The first moment of the number of proper trees

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 $G_{m,\delta}(2t)$ equals

$$\begin{aligned}
\mathbb{E}\left[Z^{(k)}_{m,\delta}(2t)\right] &= \sum_{\mathcal{T}\in\mathfrak{T}^k_m(2t)} \mathbb{P}\Big(\mathcal{T} \subseteq G_{m,\delta}(2t) \text{ and } \mathcal{T} \text{ is proper}\Big) \\
&= \sum_{\mathcal{T}\in\mathfrak{T}^k_m(2t)} \mathbb{P}\Big(\mathcal{T} \text{ is proper} | \mathcal{T} \subseteq G_{m,\delta}(2t)\Big)\mathbb{P}\Big(\mathcal{T} \subseteq G_{m,\delta}(2t)\Big) \\
&= \sum_{\mathcal{T}\in\mathfrak{T}^k_m(2t)} \mathbb{P}\Big(\mathcal{B}_{\mathcal{T}} | \mathcal{T} \subseteq G_{m,\delta}(2t)\Big) \cdot \mathbb{P}\Big(\mathcal{T} \subseteq G_{m,\delta}(2t)\Big).
\end{aligned} \tag{4.6}$$

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 4.3 (Lower bound on existence probability)**.** *Fix $m \geq 2$, $\delta > -m$ and $k \geq 0$. Given a proper k-exploration tree $\mathcal{T} \in \mathfrak{T}^k_m(2t)$, then, for t sufficiently large,*

$$\mathbb{P}\Big(\mathcal{T} \subseteq G_{m,\delta}(2t)\Big) \geq \left(\frac{a_{m,\delta}}{t}\right)^{m^{(k)}-1}, \tag{4.7}$$

where $m^{(k)} = \frac{m^{k+1}-1}{m-1}$.

Proof. Since every vertex is added before time $2t$, the denominator in (1.1)–(1.3) 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

$$\frac{m+\delta}{3t(2m+\delta)} = \frac{a_{m,\delta}}{t}. \tag{4.8}$$

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

$$\mathbb{P}\Big(\mathcal{T} \subseteq G_{m,\delta}(2t)\Big) \geq \left(\frac{a_{m,\delta}}{t}\right)^{m^{(k)}-1}. \tag{4.9}$$

□

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

Lemma 4.4 (No other vertex connects to $\mathcal{T}$)**.** *Fix $m \geq 2$, $\delta > -m$ and $0 \leq k \leq \frac{\log\log t}{\log m}$. Given a proper k-exploration tree $\mathcal{T} \in \mathfrak{T}_m^k(2t)$, then, for t sufficiently large,*

$$\mathbb{P}\Big(\mathcal{B}_{\mathcal{T}} | \mathcal{T} \subseteq G_{m,\delta}(2t)\Big) \geq \left(1 - \frac{m_\delta m^{k+1}}{t}\right)^{mt}. \tag{4.10}$$

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 in $[t]$ cannot connect to a vertex in $\mathcal{T}$, since $\mathcal{T} \subseteq [2t]\backslash[t]$. 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 $mt - (m^{(k)} - 1)$ added after time t. For A a set of vertices, 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 $\overline{\mathcal{E}}_n(A)$ 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:

$$\mathbb{P}\left(\bigcap_{i=1}^{n} \overline{\mathcal{E}}_i(\mathcal{T}) \Big| \mathcal{T} \subseteq G_{m,\delta}(2t)\right) \geq \left(1 - \frac{m_\delta m^{k+1}}{t}\right)^{n}, \tag{4.11}$$

by induction on $n = 0, \ldots, mt - (m^{(k)} - 1)$. For $n = 0$ the above holds, because both sides equal 1. Now assume that the above holds for $0 \leq n < mt - (m^{(k)} - 1)$, then

$$\begin{aligned}
&\mathbb{P}\left(\bigcap_{i=1}^{n+1} \overline{\mathcal{E}}_i(\mathcal{T}) \Big| \mathcal{T} \subseteq G_{m,\delta}(2t)\right) \\
&\quad = \mathbb{P}\left(\overline{\mathcal{E}}_{n+1}(\mathcal{T}) \Big| \bigcap_{i=1}^{n} \overline{\mathcal{E}}_i(\mathcal{T}) \cap \{\mathcal{T} \subseteq G_{m,\delta}(2t)\}\right) \mathbb{P}\left(\bigcap_{i=1}^{n} \overline{\mathcal{E}}_i(\mathcal{T}) \Big| \mathcal{T} \subseteq G_{m,\delta}(2t)\right) \\
&\quad \geq \left(1 - \mathbb{P}\left(\mathcal{E}_{n+1}(\mathcal{T}) \Big| \bigcap_{i=1}^{n} \overline{\mathcal{E}}_i(\mathcal{T}) \cap \{\mathcal{T} \subseteq G_{m,\delta}(2t)\}\right)\right) \cdot \left(1 - \frac{m_\delta m^{k+1}}{t}\right)^{n}.
\end{aligned} \tag{4.12}$$

Since it is known that at the time that the $(n+1)$-st 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 (1.1)–(1.3) will be at least t. Thus, the right-hand side of (4.12) is at least

$$\begin{aligned}
\left(1 - \sum_{i \in \mathcal{T}} \frac{m + 1 + \delta}{t}\right) \cdot \left(1 - \frac{m_\delta m^{k+1}}{t}\right)^{n} &\geq \left(1 - \frac{m_\delta m^{k+1}}{t}\right) \cdot \left(1 - \frac{m_\delta m^{k+1}}{t}\right)^{n} \\
&= \left(1 - \frac{m_\delta m^{k+1}}{t}\right)^{n+1},
\end{aligned} \tag{4.13}$$

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

$$\mathbb{P}\Big(\mathcal{B}_{\mathcal{T}} | \mathcal{T} \subseteq G_{m,\delta}(2t)\Big) \geq \left(1 - \frac{m_\delta m^{k+1}}{t}\right)^{mt - (m^{(k)} - 1)} \geq \left(1 - \frac{m_\delta m^{k+1}}{t}\right)^{mt}. \tag{4.14}$$

□

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 4.5 (Number of proper trees)**.** *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 $\left(t/m^{k+1}\right)^{m^{(k)}}$, where we recall that $m^{(k)} = \frac{m^{k+1}-1}{m-1}$.*

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

$$\#\{v \in \mathcal{T}_i^{(k)}\} = m^{(k)}. \tag{4.15}$$

For any subset $X \subseteq [2t]\backslash[t]$ with $\#\{v \in X\} = m^{(k)}$ 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 $[2t]\backslash[t]$ of size $m^{(k)}$ is $\binom{t}{m^{(k)}}$ which is at least

$$\left(\frac{t}{m^{(k)}}\right)^{m^{(k)}} \geq \left(\frac{t}{m^{k+1}}\right)^{m^{(k)}}, \tag{4.16}$$

where 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

$$\binom{a}{b} = \prod_{i=0}^{b-1} \frac{a-i}{b-i} \geq \left(\frac{a}{b}\right)^b. \tag{4.17}$$

□

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

Corollary 4.6 (Lower bound on expected number of proper trees)**.** *Fix $m \geq 2$, $\delta > -m$ and $0 \leq k \leq \frac{\log\log t}{\log m}$. Then, for t sufficiently large,*

$$\mathbb{E}\left[Z_{m,\delta}^{(k)}(2t)\right] \geq \frac{t}{a_{m,\delta}} \left(\frac{a_{m,\delta}}{m^{k+1}}\right)^{m^{k+1}} \left(1 - \frac{m_\delta m^{k+1}}{t}\right)^{mt}. \tag{4.18}$$

Proof. Using the bounds from Lemmas 4.3, 4.4 and 4.5 we get that

$$\begin{aligned}
\mathbb{E}\left[Z^{(k)}_{m,\delta}(2t)\right] &= \sum_{\mathcal{T}\in\mathfrak{T}^k_m(2t)} \mathbb{P}\Big(\mathcal{B}_{\mathcal{T}} | \mathcal{T}\subseteq G_{m,\delta}(2t)\Big)\cdot\mathbb{P}\Big(\mathcal{T}\subseteq G_{m,\delta}(2t)\Big) \\
&\geq \#\{\mathcal{T}\in\mathfrak{T}^k_m(2t)\}\left(1-\frac{m_\delta m^{k+1}}{t}\right)^{mt}\left(\frac{a_{m,\delta}}{t}\right)^{m^{(k)}-1} \\
&\geq \left(\frac{t}{m^{k+1}}\right)^{m^{(k)}}\left(1-\frac{m_\delta m^{k+1}}{t}\right)^{mt}\left(\frac{a_{m,\delta}}{t}\right)^{m^{(k)}-1} \\
&\geq \frac{t}{a_{m,\delta}}\left(\frac{a_{m,\delta}}{m^{k+1}}\right)^{m^{k+1}}\left(1-\frac{m_\delta m^{k+1}}{t}\right)^{mt}. \qquad (4.19)
\end{aligned}$$

□

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 4.1. 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 4.6 to get that

$$\lim_{t\to\infty}\mathbb{E}\left[Z^{(k)}_{m,\delta}(2t)\right] \geq \lim_{t\to\infty}\frac{t}{a_{m,\delta}}\left(\frac{a_{m,\delta}}{m^{k+1}}\right)^{m^{k+1}}\left(1-\frac{mm_\delta m^{k+1}}{mt}\right)^{mt}=\infty, \qquad (4.20)$$

since

$$\left(\frac{a_{m,\delta}}{m^{k+1}}\right)^{m^{k+1}}=\left(\frac{a_{m,\delta}}{m(\log t)^\varepsilon}\right)^{m(\log t)^\varepsilon}, \quad \text{and} \quad \left(1-\frac{mm_\delta m^{k+1}}{mt}\right)^{mt}\sim e^{-m^2 m_\delta(\log t)^\varepsilon}. \qquad (4.21)$$

□

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

4.2 The second moment of the number of proper trees

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

$$\begin{aligned}
\mathrm{Var}\left(Z^{(k)}_{m,\delta}(2t)\right) &= \mathrm{Var}\left(\sum_{\mathcal{T}\in\mathfrak{T}^k_m(2t)} I[\mathcal{T}\subseteq G_{m,\delta}(2t) \text{ and } \mathcal{T} \text{ is proper}]\right) \\
&= \mathrm{Var}\left(\sum_{\mathcal{T}\in\mathfrak{T}^k_m(2t)} I[F_{\mathcal{T}}]\right) = \sum_{\mathcal{T},\mathcal{T}'\in\mathfrak{T}^k_m(2t)}\mathrm{Cov}\left(I[F_{\mathcal{T}}],I[F_{\mathcal{T}'}]\right) \\
&= \sum_{\substack{\mathcal{T},\mathcal{T}'\in\mathfrak{T}^k_m(2t)\\ \mathcal{T}\neq\mathcal{T}'}}\left(\mathbb{P}\left(F_{\mathcal{T}}\cap F_{\mathcal{T}'}\right)-\mathbb{P}\left(F_{\mathcal{T}}\right)\mathbb{P}\left(F_{\mathcal{T}'}\right)\right)+\sum_{\mathcal{T}\in\mathfrak{T}^k_m(2t)}\mathbb{P}\left(F_{\mathcal{T}}\right)\left(1-\mathbb{P}\left(F_{\mathcal{T}}\right)\right). \qquad (4.22)
\end{aligned}$$

We start by studying the terms of the first sum in the following lemma.

Lemma 4.7 (Weak dependence of tree occurrences)**.** *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,*

$$\mathbb{P}\left(F_{\mathcal{T}} \cap F_{\mathcal{T}'}\right) - \mathbb{P}\left(F_{\mathcal{T}}\right)\mathbb{P}\left(F_{\mathcal{T}'}\right) \leq \left(\left(1 + \frac{2m_\delta m \log t}{t}\right)^{2m\log t} - 1\right)\mathbb{P}\left(F_{\mathcal{T}}\right)]\mathbb{P}\left(F_{\mathcal{T}'}\right). \tag{4.23}$$

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 (4.23) 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 $\overline{\mathcal{E}}_{j,s}(A)$ be the complement of $\mathcal{E}_{j,s}(A)$. We have that

$$\mathbb{P}(\mathcal{E}_{j,s}(A)) = \sum_{i\in A}\mathbb{P}(\mathcal{E}_{j,s}(i)), \tag{4.24}$$

because the events on the right-hand side are disjunct. These probabilities are given by the growth rules (1.1)–(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 $\mathbb{P}\left(F_{\mathcal{T}} \cap F_{\mathcal{T}'}\right)$, there will be a factor

$$\mathbb{P}\left(\overline{\mathcal{E}}_{j,t_0}(\mathcal{T}\cup\mathcal{T}')\right) = 1 - \mathbb{P}\left(\mathcal{E}_{j,t_0}(\mathcal{T}\cup\mathcal{T}')\right) = 1 - \sum_{i\in\mathcal{T}\cup\mathcal{T}'}\mathbb{P}\left(\mathcal{E}_{j,t_0}(i)\right). \tag{4.25}$$

In $\mathbb{P}\left(F_{\mathcal{T}}\right)\mathbb{P}\left(F_{\mathcal{T}'}\right)$, there will be a factor

$$\left(1 - \sum_{i\in\mathcal{T}}\mathbb{P}\left(\mathcal{E}_{j,t_0}(i)\right)\right)\left(1 - \sum_{i\in\mathcal{T}'}\mathbb{P}\left(\mathcal{E}_{j,t_0}(i)\right)\right). \tag{4.26}$$

It is easy to see that $1 - x - y \leq (1-x)(1-y)$ for $x, y \geq 0$, so (4.26) is at least as big as (4.25).

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 $\mathbb{P}\left(F_{\mathcal{T}} \cap F_{\mathcal{T}'}\right)$ there will only be a factor

$$\mathbb{P}\left(\mathcal{E}_{j,t_0}(h)\right), \tag{4.27}$$

since it will then automatically not connect to a vertex in $\mathcal{T}'$. In $\mathbb{P}\left(F_{\mathcal{T}}\right)\mathbb{P}\left(F_{\mathcal{T}'}\right)$, however, there will be a factor

$$\mathbb{P}\left(\mathcal{E}_{j,t_0}(h)\right)\left(1 - \sum_{i\in\mathcal{T}'}\mathbb{P}\left(\mathcal{E}_{j,t_0}(i)\right)\right). \tag{4.28}$$

When we multiply (4.28) by $\left(1 - \sum_{i\in\mathcal{T}'}\mathbb{P}\left(\mathcal{E}_{j,t_0}(i)\right)\right)^{-1}$ we obtain precisely (4.27). 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

$$\left(1 - \sum_{i\in\mathcal{T}'}\mathbb{P}\left(\mathcal{E}_{j,t_0}(i)\right)\right)^{-1} \leq \left(1 - \frac{m_\delta m^{k+1}}{t}\right)^{-1}. \tag{4.29}$$

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

$$\begin{aligned}\frac{\mathbb{P}\left(F_{\mathcal{T}} \cap F_{\mathcal{T}'}\right)}{\mathbb{P}\left(F_{\mathcal{T}}\right)\mathbb{P}\left(F_{\mathcal{T}'}\right)} &\leq \prod_{h\in\mathcal{T}}\left(1-\sum_{h\in\mathcal{T}'}\mathbb{P}\left(\mathcal{E}_{j,t_0}(h)\right)\right)^{-1}\prod_{k\in\mathcal{T}'}\left(1-\sum_{k\in\mathcal{T}'}\mathbb{P}\left(\mathcal{E}_{j,t_0}(k)\right)\right)^{-1}\\ &\leq \left(1-\frac{m_\delta m^{k+1}}{t}\right)^{-2m^{k+1}} = \left(1+\frac{m_\delta m^{k+1}}{t-m_\delta m^{k+1}}\right)^{2m^{k+1}}\\ &\leq \left(1+\frac{m_\delta m\log t}{t-m_\delta m\log t}\right)^{2m\log t} \leq \left(1+\frac{2m_\delta m\log t}{t}\right)^{2m\log t}. \end{aligned} \tag{4.30}$$

□

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

Proof of Proposition 4.2. Let $c_{m,\delta} = 8m_\delta m^2$. Then, using Lemma 4.7, we have that

$$\begin{aligned}\mathrm{Var}\left(Z^{(k)}_{m,\delta}(2t)\right) &= \sum_{\substack{\mathcal{T},\mathcal{T}'\in\mathfrak{T}^k_m(2t)\\ \mathcal{T}\neq\mathcal{T}'}}\left(\mathbb{P}\left(F_{\mathcal{T}}\cap F_{\mathcal{T}'}\right)-\mathbb{P}\left(F_{\mathcal{T}}\right)\mathbb{P}\left(F_{\mathcal{T}'}\right)\right)+\sum_{\mathcal{T}\in\mathfrak{T}^k_m(2t)}\mathbb{P}\left(F_{\mathcal{T}}\right)\left(1-\mathbb{P}\left(F_{\mathcal{T}}\right)\right)\\ &\leq \sum_{\substack{\mathcal{T},\mathcal{T}'\in\mathfrak{T}^k_m(2t)\\ \mathcal{T}\neq\mathcal{T}'}}\left(\left(1+\frac{2m_\delta m\log t}{t}\right)^{2m\log t}-1\right)\mathbb{P}\left(F_{\mathcal{T}}\right)\mathbb{P}\left(F_{\mathcal{T}'}\right)+\sum_{\mathcal{T}\in\mathfrak{T}^k_m(2t)}\mathbb{P}\left(F_{\mathcal{T}}\right). \end{aligned} \tag{4.31}$$

Since

$$\left(1+\frac{2m_\delta m\log t}{t}\right)^{2m\log t}-1 \leq e^{\frac{c_{m,\delta}}{2}\frac{(\log t)^2}{t}}-1 \leq c_{m,\delta}\frac{(\log t)^2}{t}, \tag{4.32}$$

we have that (4.31) is at most

$$\begin{aligned} &c_{m,\delta}\frac{(\log t)^2}{t}\sum_{\substack{\mathcal{T},\mathcal{T}'\in\mathfrak{T}^k_m(2t)\\ \mathcal{T}\neq\mathcal{T}'}}\mathbb{P}\left(F_{\mathcal{T}}\right)\mathbb{P}\left(F_{\mathcal{T}'}\right)+\mathbb{E}\left[Z^{(k)}_{m,\delta}(2t)\right]\\ &\qquad\leq c_{m,\delta}\frac{(\log t)^2}{t}\sum_{\mathcal{T},\mathcal{T}'\in\mathfrak{T}^k_m(2t)}\mathbb{P}\left(F_{\mathcal{T}}\right)\mathbb{P}\left(F_{\mathcal{T}'}\right)+\mathbb{E}\left[Z^{(k)}_{m,\delta}(2t)\right]\\ &\qquad= c_{m,\delta}\frac{(\log t)^2}{t}\mathbb{E}\left[Z^{(k)}_{m,\delta}(2t)\right]^2+\mathbb{E}\left[Z^{(k)}_{m,\delta}(2t)\right]. \end{aligned} \tag{4.33}$$

□

A Appendix

A.1 The tails of the degree of sequence

Lemma A.1 (The total degree of high degree vertices)**.** *Fix $m \geq 1$ and $\delta > -m$. Assume that $l_t \to \infty$, as $t \to \infty$ and that $l_t \leq u_1 = t^{\frac{1}{2(\tau-1)}}(\log t)^{-\frac{1}{2}}$. Then there exists a constant $B > 0$ such that with probability exceeding $1 - o(t^{-1})$,*

$$\sum_{i:D_i(t)\geq l_t} D_i(t) \geq Btl_t^{2-\tau}. \tag{A.1}$$

Proof. We note that

$$\sum_{i:D_i(t)\geq l_t} D_i(t) \geq l_t N_{\geq l_t}(t), \tag{A.2}$$

where $N_{\geq l_t}(t) = \#\{i \leq t : D_i(t) \geq l_t\}$ is the number of vertices with degree at least l_t.

In [23], detailed asymptotics for $N_{\geq l_t}(t)$ were proved for model (c) that we will survey now. These asymptotics play a key role throughout the proof.

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

$$\mathbb{P}\Big(|N_{\geq l_t}(t) - \mathbb{E}[N_{\geq l_t}(t)]| \geq B_1\sqrt{t\log t}\Big) = o(t^{-1}). \tag{A.3}$$

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 [17], and holds for any of the models (a)–(c).

Secondly, with

$$N_{l_t}(t) = \#\{i \leq t : D_i(t) = l_t\}, \tag{A.4}$$

the total number of vertices of degree equal to l_t, and with p_{l_t} defined by

$$p_{l_t} = \frac{\theta\Gamma(l_t+\delta)\Gamma(m+\delta+\theta)}{\Gamma(m+\delta)\Gamma(l_t+1+\delta+\theta)}, \quad l_t \geq m, \tag{A.5}$$

so that $p_k \sim k^{-\tau}$ with $\tau = 3+\delta/m$, there exists a constant B_2 such that

$$\sup_{l\geq 1} |\mathbb{E}[N_{l_t}(t)] - tp_{l_t}| \leq B_2. \tag{A.6}$$

For model (c), this is shown in [23], for model (a) this is shown in [31, 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 (A.6)) are proved in [30].

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

$$\begin{aligned} N_{\geq l_t}(t) &\geq \mathbb{E}[N_{\geq l_t}(t)] - B_1\sqrt{t\log t} \geq \mathbb{E}[N_{\geq l_t}(t)] - \mathbb{E}[N_{\geq 2l_t}(t)] - B_1\sqrt{t\log t} \\ &\geq \sum_{l=l_t}^{2l_t-1} [tp_{l_t} - B_2] - B_1\sqrt{t\log t} \geq B_3 t l_t^{1-\tau} - B_2 l_t - B_1\sqrt{t\log t}, \end{aligned} \tag{A.7}$$

for some $B_3 > 0$. We now wish to pick l_t such that $tl_t^{1-\tau}$ is the dominating term in the right-hand side of (A.7), i.e., $l_t/t^{1/\tau} \to 0$ and $\sqrt{t\log t}/tl_t^{1-\tau} \to 0$, as $t\to\infty$. Note that $\frac{1}{\tau} \geq \frac{1}{2(\tau-1)}$ for all $\tau > 2$, so we need to take u_1 as in the statement of the lemma and $l_t \leq u_1$. For this choice we find that (A.1) holds with probability exceeding $1-o(t^{-1})$ and that, **whp**, $N_{\geq l_t}(t)/\sqrt{t} \to \infty$, as $t\to\infty$. □

A.2 The diameter of the multinomial graph

Lemma A.2 (Diameter multinomial graph)**.** *Let H_{n_t} be the multinomial graph with parameters defined in* (3.9)*. Then,* **whp***, the diameter of H_{n_t} is bounded from above by the diameter of the uniform Erdős-Rényi graph $G(n_t, m_t)$, where the number m_t of edges is equal to*

$$m_t = \frac{1}{2}e_t\Big(1-(1-q_t)^t\Big). \tag{A.8}$$

Proof. Observe that by definition of the multinomial graph, and with $e_t = n_t(n_t-1)/2$,

$$M_{n_t} = \sum_{i=1}^{e_t} I[\{\sum_{j=1}^{t} N_{j,i} > 0\}]. \tag{A.9}$$

We only have to show that, **whp**, the random number of edges M_{n_t} dominates the deterministic number m_t. This can be deduced from Chebychev's inequality as follows.

From an easy calculation,

$$\mathbb{E}[M_{n_t}] = e_t(1-(1-q_t)^t) = 2m_t, \tag{A.10}$$

and

$$\mathrm{Var}(M_{t_n}) \;=\; e_t^2\left((1-2q_t)^t-(1-q_t)^{2t}\right) - e_t\left((1-q_t)^t-(1-q_t)^{2t}\right). \tag{A.11}$$

From (3.8), (3.9) and a binomial expansion we conclude

$$\mathrm{Var}(M_{n_t}) \sim tq_te_t \sim \mathbb{E}[M_{n_t}] = 2m_t. \tag{A.12}$$

so that the variance is of the same order as the first moment. Applying the Chebychev inequality yields

$$\mathbb{P}(M_{n_t} < m_t) \;\le\; \mathbb{P}(|M_{n_t}-\mathbb{E}[M_{n_t}]| > m_t) \le \frac{\mathrm{Var}(M_{n_t})}{m_t^2} \sim \frac{4\mathrm{Var}(M_{n_t})}{\mathbb{E}[M_{n_t}]^2} \to 0. \tag{A.13}$$

□

A.3 Proof of Lemma 3.9

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

We first note that, for models (a) and (b), the model for general $m \ge 1$ is obtained from the model for $m = 1$ by taking $\delta' = \delta/m$ and identifying groups of m vertices. For $m = 1$ and $\delta > -1$, we shall show by induction on j, that for model (a)

$$\mathbb{P}(D_i(t) = j) \le C_j \frac{\Gamma(t)\Gamma(i+\Delta)}{\Gamma(t+\Delta)\Gamma(i)}, \tag{A.14}$$

for all $t \ge i$ and $j \ge m$, with $\Delta = (1+\delta)/(2+\delta) \in (0,1)$ and where C_j will be determined in the course of the proof. Clearly, for every $t \ge i$, for model (a),

$$\mathbb{P}(D_i(t) = 1) = \prod_{s=i+1}^{t}\Big(1 - \frac{1+\delta}{(2+\delta)(s-1)+(1+\delta)}\Big) = \frac{\Gamma(t)\Gamma(i+\Delta)}{\Gamma(t+\Delta)\Gamma(i)}, \tag{A.15}$$

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

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

$$\mathbb{P}(D_i(t) = j) = \sum_{s=i+j-1}^{t} \mathbb{P}(D_i(s-1) = j-1)\frac{j-1+\delta}{(2+\delta)(s-1)+1+\delta}\mathbb{P}(D_i(t) = j | D_i(s) = j). \tag{A.16}$$

By the induction hypothesis, we have that

$$\mathbb{P}(D_i(s-1) = j-1) \le C_{j-1}\frac{\Gamma(s-1)\Gamma(i+\Delta)}{\Gamma(s-1+\Delta)\Gamma(i)}. \tag{A.17}$$

Moreover, analogously to (A.15), we have that

$$\mathbb{P}(D_i(t) = j | D_i(s) = j) = \prod_{q=s+1}^{t} \Big(1 - \frac{j+\delta}{(2+\delta)(q-1)+(1+\delta)}\Big) \tag{A.18}$$
$$= \frac{\Gamma(t - \frac{j-1}{2+\delta})\Gamma(s+\Delta)}{\Gamma(t+\Delta)\Gamma(s - \frac{j-1}{2+\delta})}.$$

Combining (A.16), (A.17) and (A.18), we arrive at

$$\mathbb{P}(D_i(t) = j) \le C_{j-1} \sum_{s=i+j-1}^{t} \frac{\Gamma(s-1)\Gamma(i+\Delta)}{\Gamma(s-1+\Delta)\Gamma(i)} \frac{j-1+\delta}{(2+\delta)(s-1)+(1+\delta)} \frac{\Gamma(t - \frac{j-1}{2+\delta})\Gamma(s+\Delta)}{\Gamma(t+\Delta)\Gamma(s - \frac{j-1}{2+\delta})}$$
$$= C_{j-1} \frac{j-1+\delta}{2+\delta} \frac{\Gamma(i+\Delta)}{\Gamma(i)} \frac{\Gamma(t - \frac{j-1}{2+\delta})}{\Gamma(t+\Delta)} \sum_{s=i+j-1}^{t} \frac{\Gamma(s-1)}{\Gamma(s - \frac{j-1}{2+\delta})}. \tag{A.19}$$

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

$$\sum_{s=l}^{t} \frac{\Gamma(s+a)}{\Gamma(s+b)} = \frac{1}{a-b+1} \Big[\frac{\Gamma(t+1+a)}{\Gamma(t+b)} - \frac{\Gamma(l+a)}{\Gamma(l-1+b)} \Big] \le \frac{1}{a-b+1} \frac{\Gamma(t+1+a)}{\Gamma(t+b)}. \tag{A.20}$$

Application of (A.20) 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

$$\mathbb{P}(D_i(t) = j) \le C_{j-1} \frac{j-1+\delta}{j-1} \frac{\Gamma(i+\Delta)}{\Gamma(i)} \frac{\Gamma(t)}{\Gamma(t+\Delta)}. \tag{A.21}$$

Equation (A.21) advances the induction when we define

$$C_j = \frac{j-1+\delta}{j-1} C_{j-1}. \tag{A.22}$$

For $m > 1$, inequality (A.14) for model (a) generalizes to

$$\mathbb{P}(D_i(t) = j) \le C_j \frac{\Gamma(t)\Gamma(i + \frac{1+\delta'}{2+\delta'})}{\Gamma(t + \frac{1+\delta'}{2+\delta'})\Gamma(i)} = C_j \frac{\Gamma(t)\Gamma(i + \frac{m+\delta}{2m+\delta})}{\Gamma(t + \frac{m+\delta}{2m+\delta})\Gamma(i)}. \tag{A.23}$$

This completes the investigation of $\mathbb{P}(D_i(t) = j)$ for model (a). In an identical fashion, for model (b), we obtain for $m = 1$

$$\mathbb{P}(D_i(t) = j) \le C_j \frac{\Gamma(t-\Delta)\Gamma(i)}{\Gamma(t)\Gamma(i-\Delta)}, \tag{A.24}$$

where again $C_1 = 1$ and C_j satisfies (A.22). This generalizes to

$$\mathbb{P}(D_i(t) = j) \le C_j \frac{\Gamma(t - \frac{m+\delta}{2m+\delta})\Gamma(i)}{\Gamma(t)\Gamma(i - \frac{m+\delta}{2m+\delta})}. \tag{A.25}$$

We omit further details for model (b).

For models (a) and (b) we can generalize the inequality for $m = 1$ to $m > 1$. Unfortunately this 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 $R_i(t) \geq 0$, we proceed at 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 we do this 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 $E_i(t+1) \leq D_i(t+1)$ if $E_i(t) \leq D_i(t)$, 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

$$\mathbb{P}(E_i(t+1) = j | E_i(t) = j-1) = \frac{j-1+\delta}{(2m+\delta)t}. \tag{A.26}$$

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

$$\mathbb{P}(E_i(t) = j) \leq C_j \frac{\Gamma(t - \frac{m+\delta}{2m+\delta})\Gamma(i)}{\Gamma(t)\Gamma(i - \frac{m+\delta}{2m+\delta})}, \tag{A.27}$$

where $C_m = 1$ and, for $j > m$,

$$C_j = \frac{j-1+\delta}{j-m} C_{j-1}. \tag{A.28}$$

The verification of (A.27) is a straightforward adaptation of the one of (A.14).

We summarize the bounds in models (a)–(c): for all $m \geq 1$, and $i \in [t], j \geq m$,

$$\mathbb{P}(D_i(t) = j) \leq C_j \frac{\Gamma(t-a_1)\Gamma(i+a_2)}{\Gamma(t+a_2)\Gamma(i-a_1)}, \tag{A.29}$$

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 $p \geq 1$.

Consequently, we obtain

$$\mathbb{P}(D_i(t) \leq j) \leq j^p \frac{\Gamma(t-a_1)\Gamma(i+a_2)}{\Gamma(t+a_2)\Gamma(i-a_1)}. \tag{A.30}$$

We finally use (A.30) to complete the proof of Lemma 3.9. Take $0 < b < \frac{a_1+a_2}{a_1+a_2+1} = \frac{m+\delta}{3m+2\delta}$, then, by Boole's inequality,

$$\begin{aligned}\mathbb{P}(\exists i \leq t^b : D_i(t) \leq (\log t)^\sigma) &\leq \sum_{i=1}^{t^b} \mathbb{P}(D_i(t) \leq (\log t)^\sigma) \leq (\log t)^{\sigma p} \frac{\Gamma(t-a_1)}{\Gamma(t+a_2)} \sum_{i=1}^{t^b} \frac{\Gamma(i+a_2)}{\Gamma(i-a_1)} \\ &\leq (\log t)^{\sigma p} (a_1+a_2+1)^{-1} \frac{\Gamma(t-a_1)}{\Gamma(t+a_2)} \frac{\Gamma(t^b+a_2+1)}{\Gamma(t^b-a_1)} = o(1). \end{aligned} \tag{A.31}$$

This completes the proof of Lemma 3.9. □

A.4 Late vertices have small degree

Recall the definition of the core Core_t in (3.1), where we take $\sigma > 1$. 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 1.7.

Proposition A.3 (Late vertices have small degree)**.** *Fix $m \geq 2, \delta > -m$ and $\sigma > 1$. Then,* $\mathrm{Core}_{2t} \subseteq [t]$ **whp**.

Proof. Note that

$$\mathbb{P}\Big(\mathrm{Core}_{2t} \subseteq [t]\Big) \geq 1 - \sum_{i=t+1}^{2t} \mathbb{P}\Big(D_i(2t) \geq (\log 2t)^\sigma\Big) \geq 1 - \sum_{i=t+1}^{2t} \mathbb{P}\Big(D_t(2t) \geq (\log 2t)^\sigma\Big)$$
$$= 1 - t\mathbb{P}\Big(D_t(2t) \geq (\log 2t)^\sigma\Big), \tag{A.32}$$

because vertex t is more likely to have a large degree than vertices added after time t. In Lemma A.4 we will show that $\mathbb{P}(D_t(2t) \geq (\log 2t)^\sigma) = o\left(\frac{1}{t}\right)$, so that $\mathbb{P}\Big(\mathrm{Core}_{2t} \subseteq [t]\Big) \geq 1 - o(1)$. □

Lemma A.4 (Tails of degree distribution)**.** *Fix $m \geq 2, \delta > -m$ and $\sigma > 1$. Then,*

$$\mathbb{P}\Big(D_t(2t) \geq (\log 2t)^\sigma\Big) = o\left(1/t\right). \tag{A.33}$$

Proof. As noted in Section 1, $G_{m,\delta}(2t)$ can be constructed from $G_{1,\delta'}(2mt)$, with $\delta' = \delta/m$. We will include the superscripts to avoid confusion. Thus identify, for $i \in [2t]$, vertices $((i-1)m+1)^{(1)}, \ldots, (im)^{(1)}$ in $G_{1,\delta'}(2mt)$ with vertex $i^{(m)}$ in $G_{m,\delta}(2t)$. So (A.33) is equivalent to

$$\mathbb{P}\Big(D_{((t-1)m+1)^{(1)}}(2mt) + \ldots + D_{(tm)^{(1)}}(2mt) \geq (\log 2t)^\sigma\Big) = o\left(1/t\right). \tag{A.34}$$

We will now color the vertices and edges in the following way. Color the vertices $1^{(1)}, \ldots, ((t-1)m)^{(1)}$ and all edges between these vertices blue and color the vertices $((t-1)m+1)^{(1)}, \ldots, (tm)^{(1)}$ 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 $((t-1)m+1)^{(1)}, \ldots, (tm)^{(1)}$ 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,

$$\mathbb{P}\Big(D_{((t-1)m+1)^{(1)}}(2mt) + \ldots + D_{(tm)^{(1)}}(2mt) \geq (\log 2t)^\sigma\Big) \leq \mathbb{P}\Big(\#\{\text{red edges}\} \ + m \geq (\log 2t)^\sigma\Big). \tag{A.35}$$

Since we will bound the right-hand 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 consider 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

$$\mathbb{P}\Big((j+1)^{(1)} \text{ connects to a red vertex} \big| G^{(c)}_{1,\delta'}(j)\Big) = \frac{\sum_{v^{(1)} \text{ red}} \left(D_{v^{(1)}}(m\lfloor j/m \rfloor) + \delta'\right)}{m\lfloor j/m \rfloor (2+\delta')}$$
$$\leq \frac{\sum_{v^{(1)} \text{ red}} \left(D_{v^{(1)}}(j) + \delta'\right)}{j(2+\delta') - m(2+\delta')}. \tag{A.36}$$

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 $2mt$. Consequently,

$$\mathbb{P}\Big(D_{t^{(m)}}(2t) \geq (\log 2t)^\sigma\Big) \leq \mathbb{P}\left(\frac{S_1(mt)}{2+\delta'} + m \geq (\log 2t)^\sigma\right). \tag{A.37}$$

To analyse the probability on the right-hand side, we make use of De Finetti's Theorem [28]. 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 $1 \leq k \leq n$,

$$\mathbb{P}\Big(X_1 = \ldots = X_k = 1, X_{k+1} = 0, \ldots, X_n = 0\Big) = \mathbb{E}\Big[U^k(1-U)^{n-k}\Big]. \tag{A.38}$$

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

$$\mathbb{P}\left(\sum_{i=1}^n X_i = k\right) = \mathbb{E}\Big[\mathbb{P}\Big(\mathrm{BIN}(n,U) = k\Big|U\Big)\Big]. \tag{A.39}$$

Let X_i denote the indicator that the i-th ball drawn in the Pólya urn scheme described above is red. As shown in [31, Section 11.1], $\{X_i\}_{i=1}^\infty$ is an infinite exchangeable sequence. Note that

$$S_1(s) = (2+\delta')m + (2+\delta')\sum_{i=1}^s X_i. \tag{A.40}$$

Hence,

$$\mathbb{P}\left(\frac{S_1(mt)}{2+\delta'} + m \geq (\log 2t)^\sigma\right) = \mathbb{E}\Big[\psi(U)\Big], \tag{A.41}$$

where $0 \leq \psi(u) = \mathbb{P}\Big(\mathrm{BIN}(mt,u) \geq (\log 2t)^\sigma - 2m\Big) \leq 1$.

Now observe from [35] that

$$\psi(u) \leq e^{-(\log 2t)^\sigma + 2m}, \tag{A.42}$$

whenever u is such that $7mtu \leq (\log 2t)^\sigma - 2m$. We define $g(t) = ((\log 2t)^\sigma - 2m)/(7(mt))$. Since,

$$\begin{aligned}\mathbb{E}\Big[\psi(U)\Big] &= \mathbb{E}\Big[\psi(U)\Big|U \leq g(t)\Big]\mathbb{P}\Big(U \leq g(t)\Big) + \mathbb{E}\Big[\psi(U)\Big|U > g(t)\Big]\mathbb{P}\Big(U > g(t)\Big) \\ &\leq \psi(g(t)) + \mathbb{P}\Big(U > g(t)\Big),\end{aligned} \tag{A.43}$$

we obtain, according to (A.42),

$$\mathbb{P}\left(\frac{S_1(mt)}{2+\delta'} + m \geq (\log 2t)^\sigma\right) \leq e^{-(\log 2t)^\sigma + 2m} + \mathbb{P}\Big(U > g(t)\Big) = o\left(\frac{1}{t}\right) + \mathbb{P}\Big(U > g(t)\Big). \tag{A.44}$$

It remains to show that also $\mathbb{P}(U > g(t)) = o\left(\frac{1}{t}\right)$. It turns out that U has a Beta-distribution with parameters $\alpha = m$ and $\beta = m(t-2)$ ([31]), so $\alpha, \beta > 1$. Thus we have that the probability density function

of U is unimodular, with its turning point at $t = \frac{\alpha-1}{\alpha+\beta-2}$ ([45]). It is easy to verify that $g(t) \geq \frac{\alpha-1}{\alpha+\beta-2}$, for t sufficiently large, so that

$$\mathbb{P}(U > g(t)) \leq (1-g(t)) \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} (g(t))^{\alpha-1} (1-g(t))^{\beta-1} \leq \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} (1-g(t))^{\beta}. \tag{A.45}$$

Using Stirling's formula (see e.g., [1]), one can show that there exists a constant $C > 0$, such that (A.45) is at most

$$C\frac{\beta^{\alpha}}{\Gamma(\alpha)} (1-g(t))^{\beta} \leq C(mt)^m \left(1 - \frac{(\log 2t)^{\sigma}}{8m(t-2)}\right)^{m(t-2)} \leq C(mt)^m e^{-(\log 2t)^{\sigma}/8} = o\left(1/t\right), \tag{A.46}$$

because $\sigma > 1$.

Note that we in fact proved that $\mathbb{P}(D_{t^{(m)}}(2t) \geq (\log 2t)^{\sigma}) = o\left(t^{-\gamma}\right)$, for any constant γ. □

Acknowledgements. The work of RvdH and SD is 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] 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] 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).

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

[4] N. Alon and J.H. Spencer. *The Probabilistic Method*. John Wiley & Sons, New York, second edition, (2000).

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

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

[7] 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).

[8] S. Bhamidi. Universal techniques to analyze preferential attachment trees: Global and local analysis. In preparation (2007). Available from `http://www.stat.berkeley.edu/~shanky/preferent.pdf`.

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

[10] 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).

[11] B. Bollobás, S. Janson, and O. Riordan. The phase transition in inhomogeneous random graphs. *Random Structures Algorithms*, **31**(1):3–122, (2007).

[12] 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).

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

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

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

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

[17] 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).

[18] 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).

[19] F. Chung and L. Lu. The average distance in a random graph with given expected degrees. *Internet Math.*, **1**(1):91–113, (2003).

[20] F. Chung and L. Lu. Coupling online and offline analyses for random power law graphs. *Internet Math.*, **1**(4):409–461, (2004).

[21] F. Chung and L. Lu. *Complex graphs and networks*, volume **107** of *CBMS Regional Conference Series in Mathematics*. Published for the Conference Board of the Mathematical Sciences, Washington, DC, (2006).

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

[23] M. Deijfen, H. van den Esker, R. van der Hofstad, and G. Hooghiemstra. A preferential attachment model with random initial degrees. *Arkiv för Matematik*, **47**(1):41-72, (2009).

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

[25] R. Durrett. *Random Graph Dynamics*. Cambridge University Press, Cambridge, (2007).

[26] H. van den Esker, R. van der Hofstad, and G. Hooghiemstra. Universality for the distance in finite variance random graphs. *J. Stat. Phys.*, **133**(1):169-202, (2008).

[27] 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).

[28] W. Feller. *An Introduction to Probability Theory and Its Applications, Volume II*. John Wiley & Sons, London, (1970).

[29] D. Fernholz and V. Ramachandran. The diameter of sparse random graphs. *Random Structures Algorithms*, **31**(4):482–516, (2007).

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

[31] R. van der Hofstad. *Random Graphs and Complex Networks*. In preparation, (2009). Available on `http://www.win.tue.nl/~rhofstad/NotesRGCN.pdf`.

[32] 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).

[33] R. van der Hofstad, G. Hooghiemstra, and D. Znamenski. Distances in random graphs with finite mean and infinite variance degrees. *Electron. J. Probab.*, **12**(25):703–766 (electronic), (2007).

[34] R. van der Hofstad, G. Hooghiemstra and D. Znamenski. A phase transition for the diameter of the configuration model. *Internet Math.*, **4**(1):113-128, (2008).

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

[36] J. Jordan. The degree sequences and spectra of scale-free random graphs. *Random Structures Algorithms*, **29**(2):226–242, (2006).

[37] T. F. Móri. On random trees. *Studia Sci. Math. Hungar.*, **39**(1-2):143–155, (2002).

[38] T. F. Móri. The maximum degree of the Barabási-Albert random tree. *Combin. Probab. Comput.*, **14**(3):339–348, (2005).

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

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

[41] R. Oliveira and J. Spencer. Connectivity transitions in networks with super-linear preferential attachment. *Internet Math.*, **2**(2):121–163, (2005).

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

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

[44] A. Rudas, B. Tóth, and B. Valkó. Random trees and general branching processes. *Random Structures Algorithms*, **31**(2):186–202, (2007).

[45] G.P. Wadsworth and J.G. Bryan. *Introduction to Probability and Random Variables*. McGraw-Hill Book Company, Inc., New York, (1960).