Ark. Mat., **47** (2009), 41–72
DOI: 10.1007/s11512-007-0067-4

A preferential attachment model with random initial degrees

Maria Deijfen, Henri van den Esker, Remco van der Hofstad and Gerard Hooghiemstra

Abstract. In this paper, a random graph process $\{G(t)\}_{t\geq 1}$ is studied and its degree sequence is analyzed. Let $\{W_t\}_{t\geq 1}$ be an i.i.d. sequence. The graph process is defined so that, at each integer time t, a new vertex with W_t edges attached to it, is added to the graph. The new edges added at time t are then preferentially connected to older vertices, i.e., conditionally on $G(t-1)$, the probability that a given edge of vertex t is connected to vertex i is proportional to $d_i(t-1)+\delta$, where $d_i(t-1)$ is the degree of vertex i at time $t-1$, independently of the other edges. The main result is that the asymptotical degree sequence for this process is a power law with exponent $\tau=\min\{\tau_{\mathrm{W}}, \tau_{\mathrm{P}}\}$, where τ_{W} is the power-law exponent of the initial degrees $\{W_t\}_{t\geq 1}$ and τ_{P} the exponent predicted by pure preferential attachment. This result extends previous work by Cooper and Frieze.

1. Introduction

Empirical studies on real life networks, such as the Internet, the World-Wide Web, social networks, and various types of technological and biological networks, show fascinating similarities. Many of the networks are *small worlds*, meaning that typical distances in the network are small, and many of them have *power-law degree sequences*, meaning that the number of vertices with degree k falls off as $k^{-\tau}$ for some exponent $\tau>1$. See [16] for an example of these phenomena in the Internet, and [25] and [26] for an example on the World-Wide Web. Also, [27, Table 3.1] gives an overview of a large number of networks and their properties.

Incited by these empirical findings, random graphs have been proposed to model and/or explain these phenomena – see [3] for an introduction to random graph models for complex networks. Two particular classes of models that have been studied from a mathematical viewpoint are (i) graphs where the edge probabilities depend on certain weights associated with the vertices, see e.g. [7], [11], [12],

[13] and [29], and (ii) so-called preferential attachment models, see e.g. [2], [6], [8], [9] and [14]. The first class can be viewed as generalizations of the classical Erdős-Rényi graph allowing for power-law degrees. Typically, the degree of a vertex is determined by its weight. Preferential attachment models are *dynamic* in the sense that a new vertex is added to the graph at each integer time. Each new vertex comes with a number of edges attached to it which are connected to the old vertices in such a way that vertices with high degree are more likely to be attached to. This has been shown to lead to graphs with power-law degree sequences, and these results are extended in the current paper.

In preferential attachment models, the degree of a vertex increases over time, implying that the oldest vertices tend to have the largest degrees. Indeed, vertices with large degrees are the most likely vertices to obtain even larger degrees. This is sometimes called the *rich-get-richer* effect. Models where the vertex degrees are determined by associated weights, on the other hand, give rise to something which could be referred to as *rich-by-birth* effect (a vertex is *born* with a weight which controls its degree). In reality, both these effects could play a role.

The aim of the current paper is to formulate and analyze a model that combines the rich-get-richer and rich-by-birth effects. The model is a preferential attachment model where the number of edges added upon the addition of a new vertex is a random variable associated to the vertex. For bounded initial degrees, the model is included in the very general class of preferential attachment models treated in [14], but the novelty of the model lies in that the initial degrees can have an arbitrary distribution. In particular, we can take the weight distribution to be a power law, which gives a model with two "competing" power laws: the power law caused by the preferential attachment mechanism and the power law of the initial degrees. In such a situation it is indeed not clear which of the power laws will dominate in the resulting degrees of the graph. Our main result implies that the most heavy-tailed power law wins, that is, the degrees in the resulting graph will follow a power law with the same exponent as the initial degrees in case this is smaller than the exponent induced by the preferential attachment, and with an exponent determined by the preferential attachment in case this is smaller.

The proof of our main result requires finite moment of order $1+\varepsilon$ for the initial degrees. However, we believe that the conclusion is true also in the infinite mean case. More specifically, we conjecture that, when the distribution of the initial degrees is a power law with infinite mean, the degree sequence in the graph will obey a power law with the same exponent as the one of the initial degrees. Indeed, the power law of the initial degrees will always be the "strongest" in this case, since preferential attachment mechanisms only seem to be able to produce power laws with finite mean. In reality, power laws with infinite mean are not uncommon, see e.g. [27, Table 3.1] for some examples, and hence it is desirable to find a model that can capture this. We have not been able to give a full proof for the infinite mean case, but we present partial results in Section 1.2.

1.1. Definition of the model

The model that we consider is described by a graph process $\{G(t)\}_{t\geq 1}$. To define it, let $\{W_i\}_{i\geq 1}$ be an independent identically distributed (i.i.d.) sequence of positive integer-valued random variables and let $G(1)$ be a graph consisting of two vertices v_0 and v_1 with W_1 edges joining them. For $t\geq 2$, the graph $G(t)$ is constructed from $G(t-1)$ in such a way that a vertex v_t, with associated weight W_t, is added to the graph $G(t-1)$, and the edge set is updated by adding W_t edges between the vertex v_t and the vertices v_0, v_1, ..., v_{t-1}. Thus, W_t is the *random initial degree* of vertex v_t. Write $d_0(s), ..., d_{t-1}(s)$ for the degrees of the vertices v_0, v_1, ..., v_{t-1} at time $s\geq t-1$. The endpoints of the W_t edges emanating from vertex v_t are chosen independently (with replacement) from $\{v_0, ..., v_{t-1}\}$, and the probability that v_i is chosen as the endpoint of a fixed edge is equal to

$$\frac{d_i(t-1)+\delta}{\sum_{j=0}^{t-1}(d_j(t-1)+\delta)} = \frac{d_i(t-1)+\delta}{2L_{t-1}+t\delta}, \quad 0\leq i\leq t-1, \tag{1.1}$$

where $L_t=\sum_{i=1}^{t} W_i$, and δ is a fixed parameter of the model. Write S_{W} for the support of the distribution of the initial degrees. To ensure that the above expression defines a probability, we require that

$$\delta+\min\{x : x\in S_{\mathrm{W}}\}>0. \tag{1.2}$$

This model will be referred to as the PARID-model (preferential attachment with random initial degrees). Note that, when $W_i\equiv 1$ and $\delta=0$, we retrieve the original preferential attachment model from Barabási–Albert [2].

Remark 1.1. We shall give special attention to the case where $\mathbb{P}(W_i=m)=1$ for some integer $m\geq 1$, since it turns out that sharper error bounds are possible in this case. These sharper bounds are needed in [22], where the diameter in preferential attachment models is studied.

1.2. Main result

Our main result concerns the degree sequence in the graph $G(t)$. To formulate it, let $N_k(t)$ be the number of vertices with degree k in $G(t)$ and define $p_k(t)=N_k(t)/(t+1)$ as the fraction of vertices with degree k. Furthermore, let $\{r_k\}_{k\geq 1}$ be the probabilities associated with the weight distribution, that is,

$$r_k = \mathbb{P}(W_1 = k), \quad k \geq 1. \tag{1.3}$$

Finally, assume that the weights have finite mean $\mu>0$ and define $\theta=2+\delta/\mu$. We are interested in the limiting distribution of $p_k(t)$, as $t\to\infty$. This distribution, denoted by $\{p_k\}_{k\geq 1}$, is obtained as the solution of the recurrence relation

$$p_k = \frac{k-1+\delta}{\theta} p_{k-1} - \frac{k+\delta}{\theta} p_k + r_k. \tag{1.4}$$

Roughly, this relation is derived by analyzing how the number of vertices with degree k is changed upon the addition of a new vertex; see e.g. [14] for some heuristic explanation. By iteration, it can be seen that the recursion is solved by

$$p_k = \frac{\theta}{k+\delta+\theta} \sum_{i=0}^{k-1} r_{k-i} \prod_{j=1}^{i} \frac{k-j+\delta}{k-j+\delta+\theta}, \quad k \geq 1, \tag{1.5}$$

where the empty product, arising when $i=0$, is defined to be equal to one. Since $\{p_k\}_{k\geq 1}$ satisfies (1.4) with $p_0=0$, we have that $\sum_{k=1}^{\infty} p_k = \sum_{k=1}^{\infty} r_k = 1$. Hence, $\{p_k\}_{k\geq 1}$ defines a probability distribution. Our main result states that the limiting degree distribution in the PARID-model is given by $\{p_k\}_{k\geq 1}$.

Theorem 1.2. *If the initial degrees $\{W_i\}_{i\geq 1}$ have finite moment of order $1+\varepsilon$ for some $\varepsilon>0$, then there exists a constant $\gamma\in\left(0, \frac{1}{2}\right)$ such that*

$$\lim_{t\to\infty} \mathbb{P}\Big(\max_{k\geq 1} |p_k(t) - p_k| \geq t^{-\gamma}\Big) = 0,$$

where $\{p_k\}_{k\geq 1}$ is defined in (1.5). When $r_m=1$ for some integer $m\geq 1$, then $t^{-\gamma}$ can be replaced by $C\sqrt{(\log t)/t}$ for some sufficiently large constant C.

To analyze the distribution $\{p_k\}_{k\geq 1}$, first consider the case when the initial degrees are almost surely constant, that is, when $r_m=1$ for some positive integer m. Then $r_j=0$ for all $j\neq m$, and (1.5) reduces to

$$p_k = \begin{cases} \dfrac{\theta\Gamma(k+\delta)\Gamma(m+\delta+\theta)}{\Gamma(m+\delta)\Gamma(k+1+\delta+\theta)} & \text{for } k\geq m; \\ 0 & \text{for } k<m, \end{cases}$$

where $\Gamma(\cdot)$ denotes the gamma function. By Stirling's formula, we have that $\Gamma(s+a)/\Gamma(s)\sim s^a$, as $s\to\infty$, and from this it follows that $p_k\sim ck^{-(1+\theta)}$ for some constant $c>0$. Hence, the degree sequence obeys a power law with exponent $1+\theta=3+\delta/m$. Note that, by choosing $\delta>-m$ appropriately, any value of the exponent larger than 2 can be obtained. For other choices of $\{r_k\}_{k\geq 1}$, the behavior of $\{p_k\}_{k\geq 1}$ is less transparent. The following proposition asserts that, if $\{r_k\}_{k\geq 1}$ is a power law, then $\{p_k\}_{k\geq 1}$ is a power law as well. It also gives the aforementioned characterization of the exponent as the minimum of the exponent of the r_k's and an exponent induced by the preferential attachment mechanism.

Proposition 1.3. *Assume that* $r_k=\mathbb{P}(W_1=k)=k^{-\tau_{\mathrm{W}}}L(k)$ *for some* $\tau_{\mathrm{W}}>2$ *and some function* $k\mapsto L(k)$ *which is slowly varying at infinity. Then* $p_k=k^{-\tau}\hat{L}(k)$ *for some slowly varying function* $k\mapsto\hat{L}(k)$ *and with power-law exponent* τ *given by*

$$\tau=\min\{\tau_{\mathrm{W}},\tau_{\mathrm{P}}\}, \tag{1.6}$$

where τ_{P} *is the power-law exponent of the pure preferential attachment model given by* $\tau_{\mathrm{P}}=3+\delta/\mu$. *When* r_k *decays faster than a power law, then* (1.6) *remains true with the convention that* $\tau_{\mathrm{W}}=\infty$.

Now assume that the mean of the initial degrees $\{W_i\}_{i\geq 1}$ is infinite. More specifically, suppose that $\{r_k\}_{k\geq 1}$ is a power law with exponent $\tau_{\mathrm{W}}\in[1,2]$. Then, we conjecture that the main result above remains true.

Conjecture 1.4. When $\{r_k\}_{k\geq 1}$ is a power-law distribution with exponent $\tau_{\mathrm{W}}\in[1,2]$, then the degree sequence in the PARID-model obeys a power law with the same exponent τ_{W}.

Unfortunately, we cannot quite prove Conjecture 1.4. However, we shall prove a slightly weaker version of it. To this end, write $N_{\geq k}(t)$ for the number of vertices with degree larger than or equal to k at time t, that is, $N_{\geq k}(t)=\sum_{i=0}^{t}\mathbf{1}_{\{d_i(t)\geq k\}}$, and let $p_{\geq k}(t)=N_{\geq k}(t)/(t+1)$. Since $d_i(t)\geq W_i$, obviously

$$\mathbb{E}[p_{\geq k}(t)]=\frac{\mathbb{E}[N_{\geq k}(t)]}{t+1}\geq\frac{\mathbb{E}[\sum_{i=1}^{t}\mathbf{1}_{\{W_i\geq k\}}]}{t+1}=\mathbb{P}(W_1\geq k)\frac{t}{t+1}=\mathbb{P}(W_1\geq k)(1+o(1)), \tag{1.7}$$

that is, the expected degree sequence in the PARID-model is always bounded from below by the weight distribution. In order to prove a related upper bound, we start by investigating the expectation of the degrees.

Theorem 1.5. *Suppose that*

$$\sum_{k>x}r_k=\mathbb{P}(W_1>x)=x^{1-\tau_{\mathrm{W}}}L(x),$$

where $\tau_{\mathrm{W}}\in(1,2)$ *and* $x\mapsto L(x)$ *is slowly varying at infinity. Then, for every* $s<\tau_{\mathrm{W}}-1$, *there exists a constant* $C>0$ *and a slowly varying function* $x\mapsto l(x)$ *such* *that, for* $i\in\{0,...,t\}$,

$$\mathbb{E}[d_i(t)^s]\le C\Big(\frac{t}{i\vee 1}\Big)^{s/(\tau_{\rm W}-1)}\Big(\frac{l(t)}{l(i)}\Big)^s,$$

where $x\vee y=\max\{x,y\}$.

As a consequence of Theorem 1.5, we obtain the following result.

Corollary 1.6. *If* $\sum_{k>x} r_k=\mathbb{P}(W_1>x)=x^{1-\tau_{\rm W}}L(x)$, *where* $\tau_{\rm W}\in(1,2)$ *and the function* $x\mapsto L(x)$ *is slowly varying at infinity, then, for every* $s<\tau_{\rm W}-1$, *there exists an* M (*independent of* t) *such that*

$$\mathbb{E}[p_{\ge k}(t)]\le Mk^{-s}.$$

Proof. For $s<\tau_{\rm W}-1$, it follows from Theorem 1.5 and Markov's inequality that

$$\begin{aligned}\mathbb{E}[p_{\ge k}(t)]&=\frac{1}{t+1}\sum_{i=0}^{t}\mathbb{P}(d_i(t)\ge k)=\frac{1}{t+1}\sum_{i=0}^{t}\mathbb{P}(d_i(t)^s\ge k^s)\\ &\le\frac{1}{t+1}\sum_{i=0}^{t}k^{-s}\mathbb{E}[d_i(t)^s]\le k^{-s}\frac{C}{t+1}\sum_{i=0}^{t}\Big(\frac{t}{i\vee 1}\Big)^{s/(\tau_{\rm W}-1)}\Big(\frac{l(t)}{l(i)}\Big)^s\le Mk^{-s},\end{aligned}\tag{1.8}$$

since, for $s<\tau_{\rm W}-1$ and using [17, Theorem 2, p. 283], there exists a constant $c>0$ such that

$$\sum_{i=0}^{t}(i\vee 1)^{-s/(\tau_{\rm W}-1)}l(i)^{-s}=ct^{1-(s/\tau_{\rm W-1})}l(t)^{-s}(1+o(1)).\qquad\square$$

Combining Corollary 1.6 with (1.7) yields that, when the weight distribution is a power law with exponent $\tau_{\rm W}\in(1,2)$, the only possible power law for the degrees has exponent equal to $\tau_{\rm W}$. This statement is obviously not as strong as Theorem 1.2, but it does offer convincing evidence for Conjecture 1.4. We prove Theorem 1.5 in Section 3.

1.3. Related work

Before proceeding with the proofs, we discuss how the proof of our main result is related to other proofs of similar results in the literature and describe some related work.

Virtually all proofs of asymptotic power laws in preferential attachment models consist of two steps: one step where it is proved that the degree sequence is concentrated around its mean, and one where the mean degree sequence is identified. In this paper, these two results are formulated in Propositions 2.1 and 2.2 below, respectively. For bounded support of W_i, the concentration result and its proof are identical in all proofs. To handle the case where W_i has unbounded support, we make use of an additional coupling argument. The main differences however arise in the statement and proof of the part where the expected degree sequence is characterized. In our Proposition 2.2, a stronger result is proved than the ones for $\delta=0$ appearing in [9] for the case of a fixed number of edges, and in [23] and [14] for the case of a random number of edges with bounded support and exponential moment, respectively. More precisely, Proposition 2.2 is valid for a wider range of k values and the error term is smaller. The model in [14] – which is much more general than the model discussed here – and the model in [23] indeed also allow for a random i.i.d. number of edges $\{W_i\}_{i\geq 1}$. However, as mentioned, there W_i is assumed to have bounded support and exponential moments, respectively, and hence, in those models, the competition of the exponents in (1.6) do not arise.

A related model which also tries to combine the rich-get-richer and the rich-by-birth effect is the so-called *fitness model*, formulated by Barabási and Bianconi [4] and [5], and later generalized by Ergün and Rodgers [15]. There the vertices are equipped with weights, referred to as *fitnesses*, which determine their ability to compete for edges. The number of edges emanating from each vertex however is fixed. Recently, the degree sequence in this model has been analyzed in [10]. Results similar to ours for various other random graph processes where a fixed number of edges emanates from each vertex can be found in [20]. Furthermore, in [6], a *directed* preferential attachment model is investigated, and it is proved that the degrees obey a power law similar to the one in [9]. In [1], the error bound in our concentration result (Proposition 2.1) is proved for $m=1$ for several models. For related references, see [20] and [30]. Finally, we mention [24], where a graph process is studied in which, conditionally on $G(t)$, edges to different vertices are added *independently* with probability proportional to the degree of the vertex. In this case, as in [9], the power-law exponent can only take the value $\tau=3$, but it can be expected that by incorporating an additive δ-term as in (1.1), the model can be generalized to $\tau\geq 3$. However, since $\delta<0$ is not allowed in this model (by the independence of the edges to different vertices, the degree of any vertex is zero with positive probability), we expect that $\tau<3$ is not possible.

2. Proofs of Theorem 1.2 and Proposition 1.3

In this section, we prove Theorem 1.2 and Proposition 1.3. We start by proving Proposition 1.3, since the proof of Theorem 1.2 makes use of it.

2.1. Proof of Proposition 1.3

Recall the definition (1.5) of p_k. Assume that $\{r_k\}_{k\geq 1}$ is a power-law distribution with exponent $\tau_{\rm W}>2$, that is, assume that $r_k=L(k)k^{-\tau_{\rm W}}$, for some slowly varying function $k\mapsto L(k)$. We want to show that then p_k is a power-law distribution as well, more precisely, we want to show that $p_k=\hat{L}(k)k^{-\tau}$, where $\tau=\min\{\tau_{\rm W},1+\theta\}$ and $k\mapsto \hat{L}(k)$ is again a slowly varying function. To this end, first note that the expression for p_k can be rewritten in terms of the gamma function as

$$p_k=\frac{\theta\Gamma(k+\delta)}{\Gamma(k+\delta+1+\theta)}\sum_{m=1}^{k}\frac{\Gamma(m+\delta+\theta)}{\Gamma(m+\delta)}r_m. \tag{2.1}$$

By Stirling's formula, we have that

$$\frac{\Gamma(k+\delta)}{\Gamma(k+\delta+1+\theta)}=k^{-(1+\theta)}(1+O(k^{-1})),\quad \text{as } k\to\infty, \tag{2.2}$$

and

$$\frac{\Gamma(m+\delta+\theta)}{\Gamma(m+\delta)}=m^{\theta}(1+O(m^{-1})),\quad \text{as } m\to\infty. \tag{2.3}$$

Furthermore, by assumption, $r_m=L(m)m^{-\tau_{\rm W}}$. It follows that

$$\sum_{m=1}^{k}\frac{\Gamma(m+\delta+\theta)}{\Gamma(m+\delta)}r_m \tag{2.4}$$

is convergent, as $k\to\infty$, if $\theta-\tau_{\rm W}<-1$, that is, if $\tau_{\rm W}>1+\theta$. For such values of $\tau_{\rm W}$, the distribution p_k is hence a power law with exponent $\tau_{\rm P}=1+\theta$. When $\theta-\tau_{\rm W}\geq -1$, that is, when $\tau_{\rm W}\leq\tau_{\rm P}$, the series in (2.4) diverges and, by [17, Lemma, p. 280], it can be seen that

$$k\longmapsto\sum_{m=1}^{k}\frac{\Gamma(m+\delta+\theta)}{\Gamma(m+\delta)}r_m$$

varies regularly with exponent $\theta-\tau_{\rm W}+1$. Combining this with (2.2) yields that p_k (compare (2.1)) varies regularly with exponent $\tau_{\rm W}$, as desired.

2.2. Proof of Theorem 1.2

As mentioned in Section 1.3, the proof of Theorem 1.2 consists of two parts: in the first part, we prove that the degree sequence is concentrated around its mean, and in the second part, the mean degree sequence is identified. These results are proved in two separate propositions – Propositions 2.1 and 2.2 – which are proved in Sections 2.3 and 2.4, respectively.

The result on the concentration of the degree sequence is as follows:

Proposition 2.1. *If the initial degrees* $\{W_i\}_{i\geq 1}$ *in the PARID-model have finite moments of order* $1+\varepsilon$*, for some* $\varepsilon>0$*, then there exists a constant* $\alpha\in\left(\frac{1}{2},1\right)$ *such that*

$$\lim_{t\to\infty}\mathbb{P}\Big(\max_{k\geq 1}|N_k(t)-\mathbb{E}[N_k(t)]|\geq t^{\alpha}\Big)=0.$$

When $r_m=1$ *for some* $m\geq 1$*, then* t^{α} *can be replaced by* $C\sqrt{t\log t}$ *for some sufficiently large* C*. Identical concentration estimates hold for* $N_{\geq k}(t)$*.*

As for the identification of the mean degree sequence, the following proposition says that the expected number of vertices with degree k is close to $(t+1)p_k$ for large t. More precisely, it asserts that the difference between $\mathbb{E}[N_k(t)]$ and $(t+1)p_k$ is bounded, uniformly in k, by a constant times t^{β}, for some $\beta\in[0,1)$.

Proposition 2.2. *Assume that the initial degrees* $\{W_i\}_{i\geq 1}$ *in the PARID-model have finite moment of order* $1+\varepsilon$ *for some* $\varepsilon>0$*, and let* $\{p_k\}_{k\geq 1}$ *be defined as in* (1.5)*. Then there exist constants* $c>0$ *and* $\beta\in[0,1)$ *such that*

$$\max_{k\geq 1}|\mathbb{E}[N_k(t)]-(t+1)p_k|\leq ct^{\beta}. \tag{2.5}$$

When $r_m=1$ *for some* $m\geq 1$*, then the above estimate holds with* $\beta=0$*.*

With Propositions 2.1 and 2.2 at hand it is not hard to prove Theorem 1.2.

Proof of Theorem 1.2. Combining (2.5) with the triangle inequality, it follows that

$$\mathbb{P}\Big(\max_{k\geq 1}|N_k(t)-(t+1)p_k|\geq ct^{\beta}+t^{\alpha}\Big)\leq\mathbb{P}\Big(\max_{k\geq 1}|N_k(t)-\mathbb{E}[N_k(t)]|\geq t^{\alpha}\Big).$$

By Proposition 2.1, the right-hand side tends to 0, as $t\to\infty$, and hence, since $p_k(t)=N_k(t)/(t+1)$, we have that

$$\lim_{t\to\infty}\mathbb{P}\left(\max_{k\geq 1}|p_k(t)-p_k|\geq\frac{ct^{\beta}+t^{\alpha}}{t+1}\right)=0.$$

The theorem follows from this by picking $0<\gamma<1-\max\{\alpha,\beta\}$. Note that, since $0\leq\beta<1$ and $\frac{1}{2}<\alpha<1$, we have $0<\gamma<\frac{1}{2}$. The proof for $r_m=1$ is analogous. □

2.3. Proof of Proposition 2.1

This proof is an adaption of a martingale argument, which first appeared in [9], and has been used for all proofs of power-law degree sequences since. The idea is to express the difference $N_k(t)-\mathbb{E}[N_k(t)]$ in terms of a Doob martingale. After bounding the martingale differences, which are bounded in terms of the random number of edges $\{W_i\}_{i\geq 1}$, the Azuma–Hoeffding inequality can be applied to conclude that the probability of observing large deviations is suitably small, at least when the initial number of edges has bounded support. When the initial degrees $\{W_i\}_{i\geq 1}$ are unbounded, an extra coupling step is required. The argument for $N_{\geq k}(t)$ is identical, so we focus on $N_k(t)$.

We start by giving an argument when $W_i \leq t^a$ for all $i\leq t$ and some $a\in\left(0,\frac{1}{2}\right)$. First note that

$$N_k(t) \leq \frac{1}{k}\sum_{l=k}^{\infty} l N_l(t) \leq \frac{1}{k}\sum_{l=1}^{\infty} l N_l(t) = \frac{L_t}{k}. \tag{2.6}$$

Thus, $\mathbb{E}[N_k(t)]\leq \mu t/k$. For $\alpha\in\left(\frac{1}{2},1\right)$, let $\eta>0$ be such that $\eta+\alpha>1$ (the choice of α will be specified in more detail below). Then, for any $k>t^\eta$, the event $|N_k(t)-\mathbb{E}[N_k(t)]|\geq t^\alpha$ implies that $N_k(t)\geq t^\alpha$, and hence that $L_t\geq kN_k(t)>t^{\eta+\alpha}$. It follows from Boole's inequality that

$$\mathbb{P}\Big(\max_{k\geq 1}|N_k(t)-\mathbb{E}[N_k(t)]|\geq t^\alpha\Big)\leq \sum_{k=1}^{t^\eta}\mathbb{P}(|N_k(t)-\mathbb{E}[N_k(t)]|\geq t^\alpha)+\mathbb{P}(L_t>t^{\eta+\alpha}).$$

Since $\eta+\alpha>1$ and $L_t/t\to\mu$ almost surely, the event $L_t>t^{\eta+\alpha}$ has small probability. To estimate the probability $\mathbb{P}(|N_k(t)-\mathbb{E}[N_k(t)]|\geq t^\alpha)$, introduce

$$M_n=\mathbb{E}[N_k(t)|G(n)],\quad n=0,...,t,$$

where $G(0)$ is defined as the empty graph. Since $\mathbb{E}[M_n]<\infty$, the process is a Doob martingale with respect to $\{G(n)\}_{n=0}^t$. Furthermore, we have that $M_t=N_k(t)$ and $M_0=\mathbb{E}[N_k(t)]$, so that

$$N_k(t)-\mathbb{E}[N_k(t)]=M_t-M_0.$$

Also, conditionally on the initial degrees $\{W_i\}_{i=1}^t$, the increments satisfy the inequality $|M_n-M_{n-1}|\leq 2W_n$. To see this, note that the additional information contained in $G(n)$ compared to $G(n-1)$ consists in how the W_n edges emanating from v_n are attached. This can affect the degrees of at most $2W_n$ vertices. By the assumption that $W_i \le t^a$ for all $i=1,...,t$, we obtain that $|M_n - M_{n-1}| \le 2t^a$. Combining all of this, it follows from the Azuma–Hoeffding inequality – see e.g. [18, Section 12.2] – that, conditionally on $W_i \le t^a$ for all $i=1,...,t$,

$$\mathbb{P}(|N_k(t) - \mathbb{E}[N_k(t)]| \ge t^\alpha) \le 2\exp\left(-\frac{t^{2\alpha}}{8\sum_{i=1}^t t^{2a}}\right) = 2\exp\left(-\frac{t^{2\alpha-1-2a}}{8}\right),$$

so that we end up with the estimate, again conditionally on $W_i \le t^a$ for all $i=1,...,t$,

$$\mathbb{P}\Big(\max_{k\ge1} |N_k(t) - \mathbb{E}[N_k(t)]| \ge t^\alpha\Big) \le 2t^\eta \exp\left(-\frac{t^{2\alpha-1-2a}}{8}\right) + \mathbb{P}(L_t > t^{\eta+\alpha}). \tag{2.7}$$

Since $a<\frac{1}{2}$, the above exponential tends to 0 for any $\alpha<1$ satisfying that $\alpha>a+\frac{1}{2}$. When the initial degrees are *bounded*, the above argument can be adapted to yield that the probability that $\max_{k\ge1} |N_k(t) - \mathbb{E}[N_k(t)]|$ exceeds $C\sqrt{t\log t}$ is $o(1)$ for some $C>0$ sufficiently large. We omit the details of this argument.

We conclude that Proposition 2.1 has been proved for all graphs $G(t)$ satisfying that $W_i \le t^a$ for arbitrary $a\in\left(0,\frac{1}{2}\right)$. Naturally, this assumption may not be true. When the initial degrees are bounded, the assumption is true, even with t^a replaced by m, but we are interested in graphs having initial degrees with finite $(1+\varepsilon)$-moments. We next extend the proof to this setting by a coupling argument.

Fix $a\in\left(0,\frac{1}{2}\right)$, arbitrarily, and define, for $i=1,2,...,t$ and $1\le s\le t$,

$$W_i' = W_i \wedge t^a \quad \text{and} \quad L_s' = \sum_{i=1}^s W_i', \tag{2.8}$$

where $x\wedge y = \min\{x,y\}$. Then, the above argument shows that the PARID-model with initial degrees $\{W_i'\}_{i=1}^t$ satisfies the claim in Proposition 2.1. Denote the graph process with initial degrees $\{W_i'\}_{i=1}^t$ by $\{G'(i)\}_{i=1}^t$ and its degrees by $d_i'(s)$, $i\le s\le t$. We now present a coupling between $\{G(i)\}_{i=1}^t$ and $\{G'(i)\}_{i=1}^t$.

Define the attachment probabilities in $\{G(i)\}_{i=1}^t$ and $\{G'(i)\}_{i=1}^t$ by

$$p_i(s) = \frac{d_i(s-1)+\delta}{2L_{s-1}+\delta s} \quad \text{and} \quad p_i'(s) = \frac{d_i'(s-1)+\delta}{2L_{s-1}'+\delta s}. \tag{2.9}$$

Observe that $p_i'(s)$ is properly defined since $d_i'(s-1)+\delta \ge W_i'+\delta = W_i\wedge t^a+\delta \ge 0$, for $t^a \ge \min\{x : x\in S_{\mathrm{W}}\}$, which is true for t not too small.

We number the edges by saying that the edge (s,l) is the lth edge of vertex s, where $1\le l\le W_s$. The aim is to couple all edges in such a way that most edges have the same starting and ending vertices in G and G'. For this, we shall split the set of edges into two classes, the *successfully coupled edges*, and the *miscoupled edges*. The successfully coupled edges will have identical starting and ending vertices in

G and in G', while the miscoupled edges will either only exist in G (when $l>W_s'$ for edge (s,l)) or will have the same starting vertex, but different ending vertices in G and in G' (when $l\le W_s'$ for edge (s,l)). We shall denote the set of miscoupled edges with number (s,l) with $s\le t$ by $U(t)$. We now explain when an edge is miscoupled. For any $W_s'<l\le W_s$, the edge with number (s,l) is miscoupled. In the graph $G(s)$, we attach the edge to a vertex i with probability $p_i(s)$, while in $G'(s)$ this edge is absent. For $1\le l\le W_s'$, the edge with number (s,l) is attached to i in *both* graphs with probability $m_i(s)=p_i(s)\wedge p_i'(s)$, where $i=0,1,...,s-1$. Observe that $\sum_{i=0}^{s-1} p_i(s)=\sum_{i=0}^{s-1} p_i'(s)=1$, but $\sum_{i=0}^{s-1} m_i(s)\le 1$. For each edge with number (s,l) with $1\le l\le W_s'$, we take *one* trial, independent of all randomness involved, with probability vector

$$(m_0(s), m_1(s), ..., m_{s-1}(s), \nu(s)), \tag{2.10}$$

where $\nu(s)=1-\sum_{i=0}^{s-1} m_i(s)$. If the trial ends in cell i, which happens with probability $m_i(s)$, then we attach the edge (s,l) to vertex i in both $G(s)$ and $G'(s)$, and the edge (s,l) is coupled successfully. If the trial ends in cell s, which happens with probability $\nu(s)$, then the edge (s,l) is miscoupled, so that $(s,l)\in U(s)$. Then, in the graphs $G(s)$ and $G'(s)$, respectively, we attach the edge (s,l) to vertex $0,1,...,s-1$ according to two further, independent trials with probability vectors

$$\begin{aligned} &\frac{1}{\nu(s)}(p_0(s)-m_0(s), ..., p_{s-1}(s)-m_{s-1}(s)),\\ &\frac{1}{\nu(s)}(p_0'(s)-m_0(s), ..., p_{s-1}'(s)-m_{s-1}(s)), \end{aligned} \tag{2.11}$$

respectively (note that since $m_i(s)=p_i(s)\wedge p_i'(s)$, these draws are indeed different a.s.). From this definition, we conclude that the probability of attaching any edge of vertex s to vertex i in the graph G has marginal probability

$$m_i(s)+\nu(s)\frac{p_i(s)-m_i(s)}{\nu(s)}=p_i(s), \tag{2.12}$$

as required. Similarly, this marginal probability equals $p_i'(s)$ in G', so that the graphs G and G' have the correct marginal distributions. We note that each miscoupled edge in $U(s)$ creates a difference in degrees of at most 2 in $G(s)$ and $G'(s)$, so that

$$\sum_{i=1}^{s} |d_i(s)-d_i'(s)|\le 2|U(s)|. \tag{2.13}$$

Indeed, when $l>W_s'$, the edge (s,l) is absent in $G'(s)$ and present in $G(s)$, so that the sum of absolute difference in degrees is increased by at most 2, while if $l\le W_s'$

and $(s,l)\in U(s)$, then only the ending vertices of the edge (s,l) are different in $G(s)$ and $G'(s)$, so that the sum of absolute difference in degrees is again increased by at most 2.

From the above construction we get

$$\mathbb{E}[|U(s)|]=\mathbb{E}[|U(s-1)|]+2\mathbb{E}[W_s-W_s']+\mathbb{E}[R_s], \tag{2.14}$$

where R_s is the total number of miscoupled edges during the attachment of the edges with numbers (s,l) and $l\le W_s'$. From (2.11), we obviously obtain

$$\mathbb{E}[R_s]=\mathbb{E}[\mathbb{E}[R_s \mid W_s]]=\mathbb{E}[W_s'\nu(s)]=\mathbb{E}[W_s']\mathbb{E}[\nu(s)], \tag{2.15}$$

because W_s' is independent of $m_i(s)$, $i=0,1,...,s-1$, and hence of $\nu(s)$.

In order to bound $\mathbb{E}[R_s]$, we observe that

$$\nu(s)=1-\sum_{i=0}^{s-1} m_i(s)=\sum_{i=0}^{s-1}[\,p_i(s)-(p_i(s)\wedge p_i'(s))]=\frac{1}{2}\sum_{i=0}^{s}|p_i(s)-p_i'(s)|.$$

We bound

$$\begin{aligned} |p_i(s)-p_i'(s)| &= \left|\frac{d_i(s-1)+\delta}{2L_{s-1}+\delta s}-\frac{d_i'(s-1)+\delta}{2L_{s-1}'+\delta s}\right| \\ &\le \frac{|d_i(s-1)-d_i'(s-1)|}{2L_{s-1}+\delta s}+\frac{2(L_{s-1}-L_{s-1}')(d_i'(s-1)+\delta)}{(2L_{s-1}+\delta s)(2L_{s-1}'+\delta s)}, \end{aligned} \tag{2.16}$$

because $L_{s-1}'\le L_{s-1}$. From (2.16) we obtain the following upper bound for $\nu(s)$:

$$\begin{aligned} \nu(s) &= \frac{1}{2}\sum_{i=0}^{s-1}|p_i(s)-p_i'(s)| \\ &\le \frac{1}{2}\sum_{i=0}^{s-1}\frac{|d_i(s-1)-d_i'(s-1)|}{2L_{s-1}+\delta s}+\frac{1}{2}\sum_{i=0}^{s-1}\frac{2(L_{s-1}-L_{s-1}')(d_i'(s-1)+\delta)}{(2L_{s-1}+\delta s)(2L_{s-1}'+\delta s)} \\ &\le \frac{|U(s-1)|}{2L_{s-1}+\delta s}+\frac{L_{s-1}-L_{s-1}'}{2L_{s-1}+\delta s}, \end{aligned} \tag{2.17}$$

by (2.13). The following lemma bounds the expected value of $|U(t)|$.

Lemma 2.3. *There exist constants $K>0$ and $b\in(0,1)$ such that for all $t\in\mathbb{N}$,*

$$\mathbb{E}[|U(t)|]\le Kt^b. \tag{2.18}$$

Proof. We prove Lemma 2.3 by induction. We start with some preparations for the induction step. Obviously, $\mathbb{E}[W_s'] \le \mathbb{E}[W_s] = \mu$ and, from the existence of the $(1+\varepsilon)$-moment of W_s, we obtain that

$$\mathbb{E}[W_s - W_s'] = \mathbb{E}[(W_s - t^a)\mathbf{1}_{\{W_s > t^a\}}] \le t^{-a\varepsilon}\mathbb{E}[W_s^{1+\varepsilon}] \le Ct^{-a\varepsilon}. \tag{2.19}$$

Secondly, from the strong law of large numbers $L_s/s \to \mu$ a.s. Using this in combination with (2.17), we find that, taking $\zeta > 0$ such that $2(1-\zeta)\mu + \delta = (1+\zeta)\mu > 1$, which is possible since $2\mu + \delta > \mu$,

$$\begin{aligned} \mathbb{E}[\nu(s)] &\le \frac{\mathbb{E}[|U(s-1)|]}{(s-1)(1+\zeta)\mu} + \frac{2\mathbb{E}[L_{s-1} - L_{s-1}']}{s-1} + \mathbb{P}(L_{s-1} \le (1-\zeta)\mu(s-1)) \\ &= \frac{\mathbb{E}[|U(s-1)|]}{(s-1)(1+\zeta)\mu} + 2\mathbb{E}[W_{s-1} - W_{s-1}'] + \mathbb{P}(L_{s-1} \le (1-\zeta)\mu(s-1)). \end{aligned} \tag{2.20}$$

We are now ready to prove (2.18). Obviously, for any finite set of natural numbers t, the inequality (2.18) holds by making K sufficiently large. This initializes the induction hypothesis, and we may assume in the induction step that t is large. So assume (2.18) for $s-1 < t$, with s large and we will show that (2.18) holds for s. From (2.14), (2.15), (2.19), (2.20) and the induction hypothesis, it follows that

$$\begin{aligned} \mathbb{E}[|U(s)|] &\le \mathbb{E}[|U(s-1)|] + 2\mathbb{E}[W_s - W_s'] + \mathbb{E}[R_s] \\ &\le K(s-1)^b + 2C(1+\mu)t^{-a\varepsilon} + \frac{K(s-1)^b}{(1+\zeta)(s-1)} + \mu\mathbb{P}(L_{s-1} \le (1-\zeta)\mu(s-1)) \\ &= Ks^b\left(\left(1 - \frac{1}{s}\right)^b + \frac{2C(1+\mu)}{Ks^{b+a\varepsilon}} + \frac{(1-1/s)^b}{(1+\zeta)(s-1)}\right) + \mu\mathbb{P}(L_{s-1} \le (1-\zeta)\mu(s-1)). \end{aligned}$$

Standard large deviation techniques and the fact that L_t is a sum of t i.i.d. *non-negative* random variables show that $s \mapsto \mathbb{P}(L_{s-1} \le (1-\zeta)\mu(s-1))$ converges to 0 exponentially fast for any $\zeta > 0$, so that we obtain the required bound Ks^b whenever s is sufficiently large and

$$\left(1 - \frac{1}{s}\right)^b + \frac{2C(1+\mu)}{Ks^{b+a\varepsilon}} + \frac{(1-1/s)^b}{(1+\zeta)(s-1)} < 1.$$

This can be established when $b + a\varepsilon \ge 1$, by taking s and K sufficiently large. □

We now complete the proof of Proposition 2.1. The Azuma–Hoeffding argument proves that $N_k'(t)$, the number of vertices with degree k in $G'(t)$, satisfies the bound in Proposition 2.1, i.e., that (recall (2.7))

$$\mathbb{P}\Big(\max_{k \ge 1} |N_k'(t) - \mathbb{E}[N_k'(t)]| \ge t^\alpha\Big) \le 2t^\eta \exp\left(-\frac{t^{2\alpha-1-2a}}{8}\right) + \mathbb{P}(L_t' > t^{\eta+\alpha}) \tag{2.21}$$

for $\alpha\in(\frac{1}{2},1)$ and $\eta>0$ such that $\alpha+\eta>1$ and $a\in(0,\frac{1}{2})$. Moreover, we have for every $k\geq 1$, that

$$|N_k(t)-N_k'(t)|\leq |U(t)|, \tag{2.22}$$

since every miscoupling can change the degree of at most one vertex. By (2.22) and (2.18), there is a $b\in(0,1)$ such that

$$|\mathbb{E}[N_k(t)]-\mathbb{E}[N_k'(t)]|\leq \mathbb{E}[|U(t)|]\leq Kt^b. \tag{2.23}$$

Also, by the Markov inequality, (2.22) and (2.18), for every $\alpha\in(b,1)$, we have that

$$\mathbb{P}\Big(\max_{k\geq 1}|N_k(t)-N_k'(t)|>t^\alpha\Big)\leq \mathbb{P}(|U(t)|>t^\alpha)\leq t^{-\alpha}\mathbb{E}[|U(t)|]=o(1). \tag{2.24}$$

Now fix $\alpha\in\big(b\vee\big(a+\frac{1}{2}\big),1\big)$, where $x\vee y=\max\{x,y\}$, and decompose

$$\begin{aligned}\max_{k\geq 1}|N_k(t)-\mathbb{E}[N_k(t)]|\leq \max_{k\geq 1}|N_k'(t)-\mathbb{E}[N_k'(t)]|+\max_{k\geq 1}|\mathbb{E}[N_k(t)]-\mathbb{E}[N_k'(t)]|\\ +\max_{k\geq 1}|N_k(t)-N_k'(t)|.\end{aligned} \tag{2.25}$$

The first term on the right-hand side is bounded by t^α with high probability by (2.21), the second term is, for t sufficiently large and with probability one, bounded by t^α by (2.23) while the third term is bounded by t^α with high probability by (2.24). This completes the proof.

2.4. Proof of Proposition 2.2

For $k\geq 1$, let

$$\overline{N}_k(t)=\mathbb{E}[N_k(t)|\{W_i\}_{i=1}^t] \tag{2.26}$$

denote the expected number of vertices with degree k at time t given the initial degrees $W_1,...,W_t$, and define

$$\varepsilon_k(t)=\overline{N}_k(t)-(t+1)p_k,\quad k\geq 1. \tag{2.27}$$

Also, for a sequence of real numbers $Q=\{Q_k\}_{k\geq 1}$, define the supremum norm of Q as $\|Q\|=\sup_{k\geq 1}|Q_k|$. Using this notation, since $\mathbb{E}[\overline{N}_k(t)]=\mathbb{E}[N_k(t)]$, we have to show that there are constants $c>0$ and $\beta\in[0,1)$ such that

$$\|\mathbb{E}[\varepsilon(t)]\|=\sup_{k\geq 1}|\mathbb{E}[\overline{N}_k(t)]-(t+1)p_k|\leq ct^\beta\quad \text{for } t=0,1,..., \tag{2.28}$$

where $\varepsilon(t)=\{\varepsilon_k(t)\}_{k\geq 1}$. The plan to do this is to formulate a recursion for $\varepsilon(t)$, and then to use induction in t to establish (2.28). The recursion for $\varepsilon(t)$ is obtained by combining a recursion for $\overline{N}(t)=\{\overline{N}_k(t)\}_{k\geq 1}$, that will be derived below, and the recursion for p_k in (1.4). The hard work then is to bound the error terms in this recursion; see Lemma 2.4 below.

Let us start by deriving a recursion for $\overline{N}(t)$. To this end, for a real-valued sequence $Q=\{Q_k\}_{k\geq 0}$, with $Q_0=0$, introduce the operator T_t, defined as

$$(T_tQ)_k=\left(1-\frac{k+\delta}{2L_{t-1}+t\delta}\right)Q_k+\frac{k-1+\delta}{2L_{t-1}+t\delta}Q_{k-1},\quad k\geq 1. \tag{2.29}$$

When applied to $\overline{N}(t-1)$, the operator T_t describes the effect of the addition of a single edge emanating from the vertex v_t, the vertex v_t itself being excluded from the degree sequence. Indeed, there are on the average $\overline{N}_{k-1}(t-1)$ vertices with degree $k-1$ at time $t-1$ and a new edge is connected to such a vertex with probability $(k-1+\delta)/(2L_{t-1}+t\delta)$. After this connection is made, the vertex will have degree k. Similarly, there are on the average $\overline{N}_k(t-1)$ vertices with degree k at time $t-1$. Such a vertex is hit by a new edge with probability $(k+\delta)/(2L_{t-1}+t\delta)$, and will then have degree $k+1$. The expected number of vertices with degree k after the addition of one edge is hence given by the operator in (2.29) applied to $\overline{N}(t)$.

Write T_t^n for the n-fold application of T_t, and define $T_t'=T_t^{W_t}$. Then T_t' describes the change in the expected degree sequence $\overline{N}(t)$ when all the W_t edges emanating from vertex v_t have been connected, ignoring vertex v_t itself. Hence, $\overline{N}(t)$ satisfies

$$\overline{N}_k(t)=(T_t'\overline{N}(t-1))_k+\mathbf{1}_{\{W_t=k\}},\quad k\geq 1. \tag{2.30}$$

Introduce a second operator S on sequences of real numbers $Q=\{Q_k\}_{k\geq 0}$, with $Q_0=0$, by (compare to (1.4))

$$(SQ)_k=\frac{k-1+\delta}{\theta}Q_{k-1}-\frac{k+\delta}{\theta}Q_k,\quad k\geq 1, \tag{2.31}$$

where $\theta=2+\delta/\mu$ and μ is the expectation of W_1.

The recursion (1.4) is given by $p_k=(Sp)_k+r_k$, with initial condition $p_0=0$. It is solved by $p=\{p_k\}_{k\geq 1}$, as defined in (1.5). Observe that

$$(t+1)p_k=tp_k+(Sp)_k+r_k=t(T_t'p)_k+r_k-\varkappa_k(t),\quad k\geq 1, \tag{2.32}$$

where

$$\varkappa_k(t)=t(T_t'p)_k-(Sp)_k-tp_k. \tag{2.33}$$

Combining (2.27), (2.30) and (2.32), and using the linearity of T'_t, it follows that $\varepsilon(t)=\{\varepsilon_k(t)\}_{k\geq 1}$ satisfies the recursion

$$\varepsilon_k(t)=(T'_t\varepsilon(t-1))_k+\mathbf{1}_{\{W_t=k\}}-r_k+\varkappa_k(t), \tag{2.34}$$

indeed,

$$\begin{aligned}\varepsilon_k(t)&=\overline{N}_k(t)-(t+1)p_k=(T'_t\overline{N}(t-1))_k+\mathbf{1}_{\{W_t=k\}}-t(T'_tp)_k-r_k+\varkappa_k(t)\\&=(T'_t\varepsilon(t-1))_k+\mathbf{1}_{\{W_t=k\}}-r_k+\varkappa_k(t).\end{aligned}$$

Now we define $k_t=\eta t$, where $\eta\in(\mu,2\mu+\delta)$. As, by (1.2), $\delta>-\min\{x:x\in S_{\mathrm{W}}\}\geq-\mu$, the interval $(\mu,2\mu+\delta)\neq\varnothing$. Also, by the law of large numbers, $L_t\leq k_t$, as $t\to\infty$, with high probability. Further, we define $\tilde{\varepsilon}_k(t)=\varepsilon_k(t)\mathbf{1}_{\{k\leq k_t\}}$ and note that, for $k\leq k_t$, the sequence $\{\tilde{\varepsilon}_k(t)\}_{k\geq1}$ satisfies

$$\tilde{\varepsilon}_k(t)=\mathbf{1}_{\{k\leq k_t\}}(T'_t\varepsilon(t-1))_k+\mathbf{1}_{\{W_t=k\}}-r_k+\widetilde{\varkappa}_k(t), \tag{2.35}$$

where $\widetilde{\varkappa}_k(t)=\varkappa_k(t)\mathbf{1}_{\{k\leq k_t\}}$. It follows from $\mathbb{E}[\mathbf{1}_{\{W_t=k\}}]=r_k$ and the triangle inequality that

$$\begin{aligned}\|\mathbb{E}[\varepsilon(t)]\|&\leq\|\mathbb{E}[\varepsilon(t)-\tilde{\varepsilon}(t)]\|+\|\mathbb{E}[\tilde{\varepsilon}(t)]\|\\&\leq\|\mathbb{E}[\varepsilon(t)-\tilde{\varepsilon}(t)]\|+\|\mathbb{E}[\mathbf{1}_{(-\infty,k_t]}(\,\cdot\,)T'_t\varepsilon(t-1)]\|+\|\mathbb{E}[\widetilde{\varkappa}(t)]\|,\end{aligned} \tag{2.36}$$

where $\mathbf{1}_{(-\infty,k_t]}(k)=\mathbf{1}_{\{k\leq k_t\}}$. Inequality (2.36) is the key ingredient in the proof of Proposition 2.2. We will derive the following bounds for the terms in (2.36).

Lemma 2.4. *There are constants $C_{\tilde{\varepsilon}}$, $C_\varepsilon^{(1)}$, $C_\varepsilon^{(2)}$ and $C_{\widetilde{\varkappa}}$, independent of t, such that for t sufficiently large and some $\beta\in[0,1)$,*

(a) $\|\mathbb{E}[\varepsilon(t)-\tilde{\varepsilon}(t)]\|\leq C_{\tilde{\varepsilon}}/t^{1-\beta}$;

(b) $\|\mathbb{E}[\mathbf{1}_{(-\infty,k_t]}(\,\cdot\,)T'_t\varepsilon(t-1)]\|\leq(1-C_\varepsilon^{(1)}/t)\|\mathbb{E}[\varepsilon(t-1)]\|+C_\varepsilon^{(2)}/t^{1-\beta}$;

(c) $\|\mathbb{E}[\widetilde{\varkappa}(t)]\|\leq C_{\widetilde{\varkappa}}/t^{1-\beta}$.

When $r_m=1$ for some integer $m\geq1$, then the above bounds hold with $\beta=0$.

Given these bounds, Proposition 2.2 is easily established.

Proof of Proposition 2.2. Recall that we want to establish (2.28). We shall prove this by induction on t. Fix $t_0\in\mathbb{N}$. We start by verifying the induction hypothesis for $t\leq t_0$, thus initializing the induction hypothesis. For any $t\leq t_0$, we have

$$\|\mathbb{E}[\varepsilon(t)]\|\leq\sup_{k\geq1}\mathbb{E}[\overline{N}_k(t)]+(t_0+1)\sup_{k\geq1}p_k\leq2(t_0+1), \tag{2.37}$$

since there are precisely t_0+1 vertices at time t_0 and $p_k\leq1$. This initializes the induction hypothesis, when c is so large that $2(t_0+1)\leq ct_0^\beta$. Next, we advance the induction hypothesis. Assume that (2.28) holds at time $t-1$ and apply Lemma 2.4 to (2.36) to get that

$$\begin{aligned}\|\mathbb{E}[\varepsilon(t)]\| &\leq \|\mathbb{E}[\varepsilon(t)-\tilde{\varepsilon}(t)]\|+\|\mathbb{E}[\mathbf{1}_{(-\infty,k_t]}(\cdot)T'_t\varepsilon(t-1)]\|+\|\mathbb{E}[\widetilde{\varkappa}(t)]\| \\ &\leq \frac{C_{\tilde{\varepsilon}}}{t^{1-\beta}}+\left(1-\frac{C_\varepsilon^{(1)}}{t}\right)c(t-1)^\beta+\frac{C_\varepsilon^{(2)}}{t^{1-\beta}}+\frac{C_{\widetilde{\varkappa}}}{t^{1-\beta}} \\ &\leq ct^\beta-\frac{cC_\varepsilon^{(1)}-(C_\varepsilon^{(2)}+C_{\tilde{\varepsilon}}+C_{\widetilde{\varkappa}})}{t^{1-\beta}},\end{aligned}$$

as long as $1-C_\varepsilon^{(1)}/t\geq 0$, which is equivalent to $t\geq C_\varepsilon^{(1)}$. If we then choose c large so that $cC_\varepsilon^{(1)}\geq C_\varepsilon^{(2)}+C_\varepsilon+C_{\widetilde{\varkappa}}$, $c\geq 2(t_0+1)t_0^{-\beta}$ (recall (2.37)) and $t_0\geq C_\varepsilon^{(1)}$, then we have that $\|\mathbb{E}[\varepsilon(t)]\|\leq ct^\beta$, and (2.28) follows by induction in t. □

It remains to prove Lemma 2.4. We shall prove Lemma 2.4(a)–(c) one by one, starting with (a).

Proof of Lemma 2.4(a). We have $\|\mathbb{E}[\varepsilon(t)-\tilde{\varepsilon}(t)]\|\leq\mathbb{E}[\|\varepsilon(t)-\tilde{\varepsilon}(t)\|]$, and, using the definition of $\tilde{\varepsilon}(t)$, we get that

$$\|\varepsilon(t)-\tilde{\varepsilon}(t)\| = \sup_{k>k_t}|\overline{N}_k(t)-(t+1)p_k| \leq \sup_{k>k_t}\overline{N}_k(t)+(t+1)\sup_{k>k_t}p_k.$$

The maximal possible degree of a vertex at time t is L_t, which implies that $\sup_{k>k_t}\overline{N}_k(t)=0$, when $L_t\leq k_t$. The latter is true almost surely when $r_m=1$ for some integer m, when t is sufficiently large, since for t large $L_t=mt\leq\eta t=k_t$, where $\eta\in(m,2m+\delta)$, by the fact that $\mu=m$ and $\delta>-m$. On the other hand, by (2.6), with $N_k(t)$ replaced by $\overline{N}_k(t)$ we find that $\overline{N}_k(t)\leq L_t/k_t$ for $k\geq k_t$, and we obtain that

$$\mathbb{E}\Big[\sup_{k>k_t}\overline{N}_k(t)\Big] \leq k_t^{-1}\mathbb{E}[L_t\mathbf{1}_{\{L_t>k_t\}}]. \tag{2.38}$$

With $k_t=\eta t$ for some $\eta\in(\mu,2\mu+\delta)$, we have that

$$\mathbb{E}[L_t\mathbf{1}_{\{L_t>k_t\}}]\leq k_t^{-\varepsilon}\mathbb{E}[L_t^{1+\varepsilon}\mathbf{1}_{\{L_t>k_t\}}]\leq k_t^{-\varepsilon}\mathbb{E}[|L_t-\mu t|^{1+\varepsilon}]+(\mu t)^{1+\varepsilon}k_t^{-\varepsilon}\mathbb{P}(L_t>k_t), \tag{2.39}$$

and, by the Markov inequality,

$$\mathbb{P}(L_t>k_t)\leq\mathbb{P}(|L_t-\mu t|^{1+\varepsilon}>(k_t-\mu t)^{1+\varepsilon})\leq(k_t-\mu t)^{-(1+\varepsilon)}\mathbb{E}[|L_t-\mu t|^{1+\varepsilon}].$$

Combining the two latter results, we obtain that

$$\mathbb{E}[L_t\mathbf{1}_{\{L_t>k_t\}}]\leq k_t^{-\varepsilon}\left(1+\left(\frac{\mu}{\eta-\mu}\right)^{1+\varepsilon}\right)\mathbb{E}[|L_t-\mu t|^{1+\varepsilon}]. \tag{2.40}$$

To bound the last expectation, we will use a consequence of the Marcinkiewicz–Zygmund inequality, see e.g. [19, Corollary 8.2 in §3], which runs as follows. Let $q\in[1,2]$, and suppose that $\{X_i\}_{i\geq 1}$ is an i.i.d. sequence with $\mathbb{E}[X_1]=0$ and $\mathbb{E}[|X_1|^q]<\infty$. Then there exists a constant c_q depending only on q, such that

$$\mathbb{E}\left[\left|\sum_{i=1}^{t} X_i\right|^q\right] \leq c_q t\mathbb{E}[|X_1{}^q|]. \tag{2.41}$$

Applying (2.41) with $q=1+\varepsilon$, we obtain that

$$\mathbb{E}\Big[\sup_{k>k_t} \overline{N}_k(t)\Big] \leq k_t^{-(1+\varepsilon)}\left(1+\left(\frac{\mu}{\eta-\mu}\right)^{1+\varepsilon}\right)\mathbb{E}[|L_t-\mu t|^{1+\varepsilon}] \leq c_{1+\varepsilon}t^{-\varepsilon}. \tag{2.42}$$

Furthermore, since by Proposition 1.3, we have $p_k\leq ck^{-\gamma}$ for some $\gamma>2$ (see also (1.6)), we have that $\sup_{k>k_t} p_k\leq ct^{-\gamma}$ for some constant c. It follows that

$$(t+1)\sup_{k>k_t} p_k \leq \frac{C_p}{t^{\gamma-1}},$$

and, since $\gamma>2$, part (a) is established with $C_{\tilde{\varepsilon}}=c_{1+\varepsilon}+C_p$, and $1-\beta=(\varepsilon\wedge\gamma)-1$. □

Proof of Lemma 2.4(b). We will start by showing that for t sufficiently large,

$$\|\mathbb{E}[\mathbf{1}_{(-\infty,k_t]}(\,\cdot\,)T_t\varepsilon(t-1)]\| \leq \left(1-\frac{C_\varepsilon^{(1)}}{t}\right)\|\mathbb{E}[\mathbf{1}_{(-\infty,k_t]}(\,\cdot\,)\varepsilon(t-1)]\|+\frac{C_\varepsilon^{(3)}}{t^{1-\beta}}, \tag{2.43}$$

which is (b) when we condition on $W_t=1$. We shall extend the proof to the case where $W_t\geq 1$ at a later stage. To prove (2.43), we shall prove a related bound, which also proves useful in the extension to $W_t\geq 1$. Indeed, we shall prove, that for any real-valued sequence $Q=\{Q_k\}_{k\geq 0}$ satisfying (i) $Q_0=0$ and (ii)

$$\sup_{k\geq 1} |k+\delta|\,|Q_k| \leq C_Q L_{t-1}, \tag{2.44}$$

there exists a $\beta\in(0,1)$ (independent of Q) and a constant $c>0$ such that for t sufficiently large,

$$\|\mathbb{E}[\mathbf{1}_{(-\infty,k_t]}(\,\cdot\,)T_tQ]\| \leq \left(1-\frac{C_\varepsilon^{(1)}}{t}\right)\|\mathbb{E}[\mathbf{1}_{(-\infty,k_t]}(\,\cdot\,)Q]\|+\frac{cC_Q}{t^{1-\beta}}. \tag{2.45}$$

Here we stress that Q can be *random*, for example, we shall apply (2.45) to $\varepsilon(t-1)$ in order to derive (2.43).

In order to prove (2.45), we recall that

$$\mathbb{E}[(T_tQ)_k] = \mathbb{E}\left[\left(1-\frac{k+\delta}{2L_{t-1}+t\delta}\right)Q_k+\frac{k-1+\delta}{2L_{t-1}+t\delta}Q_{k-1}\right], \quad k\geq 1. \tag{2.46}$$

In bounding this expectation we will encounter a problem in that Q_k, which is allowed to be random, and L_{t-1} are not independent (for example when $Q=\varepsilon(t-1)$). To get around this, we add and subtract the expression on the right-hand side but with the random quantities replaced by their expectations, that is, for $k\geq 1$, we write

$$\mathbb{E}[(T_tQ)_k] = \left(1-\frac{k+\delta}{2\mu(t-1)+t\delta}\right)\mathbb{E}[Q_k] + \frac{k-1+\delta}{2\mu(t-1)+t\delta}\mathbb{E}[Q_{k-1}] \tag{2.47}$$

$$+(k+\delta)\mathbb{E}\left[Q_k\frac{2L_{t-1}-2\mu(t-1)}{(2L_{t-1}+t\delta)(2\mu(t-1)+t\delta)}\right] \tag{2.48}$$

$$+(k+\delta-1)\mathbb{E}\left[Q_{k-1}\frac{2\mu(t-1)-2L_{t-1}}{(2L_{t-1}+t\delta)(2\mu(t-1)+t\delta)}\right]. \tag{2.49}$$

Note that, when $r_m=1$ for some integer $m\geq 1$, then $L_t=\mu t=mt$. Hence the terms in (2.48) and (2.49) are both equal to zero, and only (2.47) contributes. We first deal with (2.47). Observe that $k\leq k_t=\eta t$, with $\eta\in(\mu, 2\mu+\delta)$, implies that $k\leq(2\mu+\delta)(t-1)$ for t sufficiently large, and hence

$$1-\frac{k+\delta}{2\mu(t-1)+t\delta}\geq 0. \tag{2.50}$$

It follows that, for t sufficiently large,

$$\sup_{k\leq k_t}\left|\left(1-\frac{k+\delta}{2\mu(t-1)+t\delta}\right)\mathbb{E}[Q_k]+\frac{k-1+\delta}{2\mu(t-1)+t\delta}\mathbb{E}[Q_{k-1}]\right| \tag{2.51}$$
$$\leq\left(1-\frac{1}{2\mu(t-1)+t\delta}\right)\|\mathbb{E}[\mathbf{1}_{(-\infty,k_t]}(\,\cdot\,)Q]\|\leq\left(1-\frac{C_\varepsilon^{(1)}}{t}\right)\|\mathbb{E}[\mathbf{1}_{(-\infty,k_t]}(\,\cdot\,)Q]\|,$$

for some constant $C_\varepsilon^{(1)}$. This proves (2.45) – with $C_Q=0$ – when the number of edges is a.s. constant since (2.48) and (2.49) are zero. It remains to bound the terms (2.48) and (2.49) in the case where the number of edges is not a.s. constant. We will prove that the supremum over k of the absolute values of both these terms are bounded by constants divided by $t^{1-\beta}$ for some $\beta\in[0,1)$. Starting with (2.48), by using the assumption (ii) in (2.44), as well as $2L_{t-1}+\delta t\geq L_{t-1}$ for t sufficiently large, it follows that

$$\sup_{k\geq 1}\left|(k+\delta)\mathbb{E}\left[Q_k\frac{2L_{t-1}-2\mu(t-1)}{(2L_{t-1}+t\delta)(2\mu(t-1)+t\delta)}\right]\right|\leq\frac{cC_Q}{t}\mathbb{E}[|L_{t-1}-\mu(t-1)|].$$

To bound the latter expectation, we combine (2.41) for $q=1+\varepsilon$, with Hölder's inequality, to obtain that

$$\mathbb{E}[|L_t-\mu t|]\leq\mathbb{E}[|L_t-\mu t|^{1+\varepsilon}]^{1/(1+\varepsilon)}\leq(c_{1+\varepsilon}t\mathbb{E}[|W_1-\mu|^{1+\varepsilon}])^{1/(1+\varepsilon)}\leq ct^{1/(1+\varepsilon)}, \tag{2.52}$$

since W_i has finite moment of order $1+\varepsilon$ by assumption, where, without loss of generality, we can assume that $\varepsilon \le 1$. Hence, we have shown that the supremum over k of the absolute value of (2.48) is bounded from above by a constant divided by $t^{1-\beta}$, where $\beta=1/(1+\varepsilon)$. That the same is true for the term (2.49) can be seen analogously. This completes the proof of (2.45).

To prove (2.43), we note that, by convention, $\varepsilon_0(t-1)=0$, so that we only need to prove that $\sup_{k\ge1} |k+\delta|\,|\varepsilon_k(t-1)| \le cL_{t-1}$. For this, note from (2.6), the bound $p_k \le ck^{-\gamma}$, $\gamma>2$, and from the lower bound $L_t \ge t$ that

$$\begin{aligned}\sup_{k\ge1} |k+\delta|\,|\varepsilon_k(t-1)| &\le \sum_{k\ge1}(k+|\delta|)|\varepsilon_k(t-1)| \le \sum_{k\ge1}(k+|\delta|)\overline{N}_k(t-1)+t\sum_{k\ge1}(k+|\delta|)p_k \\ &\le L_{t-1}+|\delta|(t-1)+t\sum_{k\ge1}(k+|\delta|)p_k \le cL_{t-1}, \end{aligned} \tag{2.53}$$

for some constant c. This completes the proof of (2.43).

To complete the proof of Lemma 2.4(b), we first show that (2.45) implies, for every $1\le n\le t$ and all $k\ge1$, that

$$\mathbb{E}[\mathbf{1}_{\{k\le k_t\}}(T_t^n\varepsilon(t-1))_k] \le \left(1-\frac{C_\varepsilon^{(1)}}{t}\right)\|\mathbb{E}[\mathbf{1}_{(-\infty,k_t]}(\,\cdot\,)\varepsilon(t-1)]\|+\frac{nC_\varepsilon^{(3)}}{t^{1-\beta}}. \tag{2.54}$$

To see (2.54), we use induction on n. We note that (2.54) for $n=1$ is precisely equal to (2.43), and this initializes the induction hypothesis. To advance the induction hypothesis, we note that

$$\mathbf{1}_{\{k\le k_t\}}(T_t^n\varepsilon(t-1))_k = \mathbf{1}_{\{k\le k_t\}}T_t(Q(n-1))_k, \tag{2.55}$$

where $Q_k(n-1)=\mathbf{1}_{\{k\le k_t\}}(T_t^{n-1}\varepsilon(t-1))_k$. We wish to use (2.45), and we first check the assumptions (i) and (ii). By definition, $Q_0(n-1)=0$, which establishes (i). For assumption (ii), we need to do some more work. According to (2.29), and using that $2L_{t-1}+t\delta > L_{t-1} \ge t-1$, for t sufficiently large,

$$\sum_{k=1}^{\infty}(k+|\delta|)(T_tQ)_k \le \left(1+\frac{1}{t}\right)\sum_{k=1}^{\infty}(k+|\delta|)Q_k,$$

and hence, by induction,

$$\sum_{k=1}^{\infty}(k+|\delta|)(T_t^{n-1}Q)_k \le \left(1+\frac{1}{t}\right)^{n-1}\sum_{k=1}^{\infty}(k+|\delta|)Q_k.$$

Substituting $Q_k = \varepsilon_k(t-1)$ and using that $|\varepsilon_k(t-1)| \le N_k(t-1) + tp_k$, yields that

$$\begin{aligned}
(2.56) \qquad & \sum_{k\le k_t}(k+|\delta|)(T_t^{n-1}N(t-1))_k + t\sum_{k\le k_t}(k+|\delta|)(T_t^{n-1}p)_k \\
&\le \left(1+\frac{1}{t}\right)^{n-1}\sum_{k=1}^{\infty}(k+|\delta|)N_k(t-1) + \left(1+\frac{1}{t}\right)^{n-1} t\sum_{k=1}^{\infty}(k+|\delta|)p_k \\
&\le \left(1+\frac{1}{t}\right)^{n-1} cL_{t-1},
\end{aligned}$$

according to (2.53). Using the inequality $1+x \le e^x$, $x \ge 0$, together with $n \le t$, this in turn yields that

$$\sup_{k\ge 1} |k+\delta|\, |Q_k(n-1)| \le ecL_{t-1}, \tag{2.57}$$

which implies assumption (ii).

By the induction hypothesis, we have that, for $k \le k_t$,

$$\mathbb{E}[Q_k(n-1)] \le \left(1-\frac{C_\varepsilon^{(1)}}{t}\right)\|\mathbb{E}[\mathbf{1}_{(-\infty,k_t]}(\,\cdot\,)\varepsilon(t-1)]\| + \frac{(n-1)C_\varepsilon^{(3)}}{t^{1-\beta}}, \tag{2.58}$$

so that we obtain, from (2.45), with $Q = \mathbf{1}_{(-\infty,k_t]}(\,\cdot\,)T_t\varepsilon(t-1)$,

(2.59)

$$\mathbb{E}[\mathbf{1}_{\{k\le k_t\}}(T_t^n\varepsilon(t-1))_k] \le \left(1-\frac{C_\varepsilon^{(1)}}{t}\right)\|\mathbb{E}[\mathbf{1}_{(-\infty,k_t]}(\,\cdot\,)\varepsilon(t-1)]\| + \frac{(n-1)C_\varepsilon^{(3)} + cC_Q}{t^{1-\beta}},$$

which advances the induction hypothesis when $C_\varepsilon^{(3)} > cC_Q$.

By (2.59), we obtain that, for $W_t \le t$,

$$\begin{aligned}
\mathbb{E}[\mathbf{1}_{\{k\le k_t\}}(T_t'\varepsilon(t-1))_k \mid W_t] &\le \left(1-\frac{C_\varepsilon^{(1)}}{t}\right)\|\mathbb{E}[\varepsilon(t-1)\mid W_t]\| + \frac{W_tC_\varepsilon^{(3)}}{t^{1-\beta}} \\
&= \left(1-\frac{C_\varepsilon^{(1)}}{t}\right)\|\mathbb{E}[\varepsilon(t-1)]\| + \frac{W_tC_\varepsilon^{(3)}}{t^{1-\beta}},
\end{aligned}$$

where we use that $\varepsilon(t-1)$ is independent of W_t. In the case that $W_t > t$, we bound, similarly as in (2.53),

$$\sup_{k\le k_t} |(T_t'\varepsilon(t-1))_k| \le cL_t, \tag{2.60}$$

so that

$$\mathbb{E}[\mathbf{1}_{\{k\leq k_t\}}(T'_t\varepsilon(t-1))_k \mid W_t] \leq \left(1-\frac{C_\varepsilon^{(1)}}{t}\right)\|\mathbb{E}[\varepsilon(t-1)]\|+\frac{W_tC_\varepsilon^{(3)}}{t^{1-\beta}}+c\mathbb{E}[L_t\mathbf{1}_{\{W_t>t\}} \mid W_t]. \tag{2.61}$$

The bound in (b) follows from this by taking expectations on both sides, using that

$$\mathbb{E}[L_t\mathbf{1}_{\{W_t>t\}}]=\mu(t-1)\mathbb{P}(W_t>t)+\mathbb{E}[W_t\mathbf{1}_{\{W_t>t\}}]\leq\left(\frac{\mu}{t^\varepsilon}+\frac{1}{t^\varepsilon}\right)\mathbb{E}[W_t^{1+\varepsilon}], \tag{2.62}$$

after which we use that $\beta=1/(1+\varepsilon)\geq 1-\varepsilon$ and choose the constants appropriately. This completes the proof of Lemma 2.4(b). $\square$

Proof of Lemma 2.4(c). Recall that

$$\widetilde{\varkappa}_k(t)=\varkappa_k(t)\mathbf{1}_{\{k\leq k_t\}} \quad \text{with } \varkappa_k(t)=t((T'_t-I)p)_k-(Sp)_k, \tag{2.63}$$

where T_t is defined in (2.29), $T'_t=T_t^{W_t}$, S is defined in (2.31), and where I denotes the identity operator. In what follows, we will assume that $k\leq k_t$, so that $\widetilde{\varkappa}_k(t)=\varkappa_k(t)$. We start by proving a trivial bound on $\varkappa_k(t)$. By (2.34), we have that

$$\varkappa_k(t)=\varepsilon_k(t)-(T'_t\varepsilon(t-1))_k-\mathbf{1}_{\{W_t=k\}}+r_k, \tag{2.64}$$

where $\sup_{k\geq 1}|\varepsilon_k(t)|\leq cL_t$ by (2.53) and $\sup_{k\leq k_t}|(T'_t\varepsilon(t-1))_k|\leq cL_t$ by (2.60). Thus

$$\sup_{k\leq k_t}|\varkappa_k(t)|\leq C_\eta L_t \tag{2.65}$$

for some C_η (recall that $k_t=\eta t$, where $\eta\in(\mu,2\mu+\delta)$). For $x\in[0,1]$ and $w\in\mathbb{N}$, we let

$$f_k(x;w)=((I+x(T_t-I))^w p)_k.$$

Then $\varkappa_k(t)=\varkappa_k(t;W_t)$, where

$$\varkappa_k(t;w)=t[f_k(1;w)-f_k(0;w)]-(Sp)_k, \tag{2.66}$$

and $x\mapsto f_k(x;w)$ is a polynomial in x of degree w. By a Taylor expansion around $x=1$,

$$f_k(1;w)=p_k+w((T_t-I)p)_k+\tfrac{1}{2}f''_k(x_k;w) \tag{2.67}$$

for some $x_k\in(0,1)$, and, since $I+x(T_t-I)$ and T_t-I commute,

$$f''_k(x;w)=w(w-1)((I+x(T_t-I))^{w-2}(T_t-I)^2p)_k.$$

We next claim that, on the event $\{k_t \le 2L_{t-1}+(t-1)\delta\}$,

$$\sup_{k\le k_t} |((I+x(T_t-I))Q)_k| \le \sup_{k\le k_t} |Q_k|.$$

Indeed, $I+x(T_t-I)=(1-x)I+xT_t$ and $x\in[0,1]$, so that the claim follows when $\sup_{k\le k_t} |(T_tQ)_k| \le \sup_{k\le k_t} |Q_k|$. The latter is the case, since, on the event that $k+\delta\le 2L_{t-1}+t\delta$, and arguing as in (2.51), we have

$$\begin{aligned}\sup_{k\le k_t} |(T_tQ)_k| &\le \sup_{k\le k_t} \left[\left(1-\frac{k+\delta}{2L_{t-1}+t\delta}\right)|Q_k| + \frac{k-1+\delta}{2L_{t-1}+t\delta}|Q_{k-1}|\right]\\ &\le \left(1-\frac{1}{2L_{t-1}+t\delta}\right)\sup_{k\le k_t} |Q_k|.\end{aligned}$$

Since $k\le k_t$, the inequality $k+\delta\le 2L_{t-1}+t\delta$ follows when $k_t\le 2L_{t-1}+(t-1)\delta$.

As a result, on the event $\{k_t\le 2L_{t-1}+(t-1)\delta\}$, we have that

$$\max_{x\in[0,1]} \sup_{k\le k_t} |f_k''(x;w)| \le w(w-1) \sup_{k\le k_t} |((T_t-I)^2p)_k|. \tag{2.68}$$

Now recall the definition (2.31) of the operator S, and note that, for any sequence $Q=\{Q_k\}_{k=1}^{\infty}$, we can write

$$((T_t-I)Q)_k = \frac{\theta}{2L_{t-1}+t\delta}(SQ)_k = \frac{1}{t\mu}(SQ)_k + (R_tQ)_k, \tag{2.69}$$

where the remainder operator R_t is defined as

$$(R_tQ)_k = \left(\frac{k+\delta}{2t\mu+t\delta} - \frac{k+\delta}{2L_{t-1}+t\delta}\right)Q_k + \left(\frac{k-1+\delta}{2L_{t-1}+t\delta} - \frac{k-1+\delta}{2t\mu+t\delta}\right)Q_{k-1}. \tag{2.70}$$

Combining (2.66), (2.67), (2.68) and (2.69), on the event $\{k_t\le 2L_{t-1}+(t-1)\delta\}$ and uniformly for $k\le k_t$, we obtain that

$$\varkappa_k(t;w) \le \left(\frac{w}{\mu}-1\right)(Sp)_k + wt\sup_{k\le k_t}|(R_tp)_k| + \frac{1}{2}w(w-1)t\sup_{k\le k_t}|((T_t-I)^2p)_k|, \tag{2.71}$$

together with a similar lower bound with minus signs in front of the last two terms. Indeed,

$$\begin{aligned}\varkappa_k(t;w) &= t[f_k(1;w)-f_k(0;w)]-(Sp)_k\\ &= tw((T_t-I)p)_k + \tfrac{1}{2}tf_k''(x_k;w) - (Sp)_k\\ &= \frac{wt}{\mu t}(Sp)_k + wt(Rp)_k - (Sp)_k + \tfrac{1}{2}tf_k''(x_k;w),\end{aligned}$$

and (2.71) follows from this identity and (2.68).

With (2.71) at hand, we are now ready to complete the proof of (c). We start by treating the case where $r_m=1$ for some integer $m\geq 1$. In this case, with $w=W_t=m=\mu$, we have that $(w/\mu-1)(Sp)_k\equiv 0$. Furthermore, the inequality $k_t\leq 2L_{t-1}+(t-1)\delta$ is true almost surely when t is sufficiently large. Hence, we are done if we can bound the last two terms in (2.71) with $w=W_t$. To do this, note that, by the definition (2.29) of T_t and the fact that $2L_{t-1}+t\delta\geq k_t=\eta t$, with $\eta>\mu$,

$$\sup_{k\geq 1}|((T_t-I)Q)_k|\leq \frac{2}{\eta t}\sup_{k\geq 1}(k+|\delta|)|Q_k|. \tag{2.72}$$

Applying (2.72) twice yields that

$$|((T_t-I)^2p)_k|\leq \frac{4}{\eta^2t^2}\sup_{k\geq 1}(k+|\delta|)^2p_k,$$

and, since by Proposition 1.3, $p_k\leq ck^{-\gamma}$ for some $\gamma>2$, there is a constant $\widetilde{C}_p$ such that

$$\sup_{k\geq 1}(k+|\delta|)^2p_k\leq \widetilde{C}_p. \tag{2.73}$$

Finally, since $L_t=mt$, we have that

$$|(R_tp)_k|\leq \frac{2}{m(t-1)t}\sup_{k\geq 1}(k+|\delta|)p_k\leq \frac{2\widetilde{C}_p}{m(t-1)t}.$$

Summarizing, we arrive at the statement that there exists $c_{m,\delta}$ such that

$$\sup_{k\leq k_t}|\varkappa_k(t;m)|\leq \frac{c_{m,\delta}}{t},$$

which proves the claim in (c) with $\beta=0$ when $r_m=1$.

We now move to random initial degrees. For any $a\in(0,1)$, we can split

$$\varkappa_k(t)=\varkappa_k(t)\mathbf{1}_{\{W_t\leq t^a\}}+\varkappa_k(t)\mathbf{1}_{\{W_t>t^a\}}. \tag{2.74}$$

On the event $\{k_t\leq 2L_{t-1}+(t-1)\delta\}$, the first term of (2.74) can be bounded by the right-hand side of (2.71), i.e.,

$$\varkappa_k(t)\mathbf{1}_{\{W_t\leq t^a\}}$$
$$\leq \left(\left(\frac{W_t}{\mu}-1\right)(Sp)_k+tW_t\sup_{k\leq k_t}|(R_tp)_k|+\frac{W_t(W_t-1)}{2}t\sup_{k\leq k_t}|((T_t-I)^2p)_k|\right)\mathbf{1}_{\{W_t\leq t^a\}},$$

with a similar lower bound where the last two terms have a minus sign. From (2.65), we obtain the upper bound

$$\varkappa_k(t)\mathbf{1}_{\{W_t>t^a\}}\leq C_\eta L_t\mathbf{1}_{\{W_t>t^a\}}.$$

Combining these two upper bounds with the identity (2.74), and adding the term $(W_t/\mu-1)(Sp)_k\mathbf{1}_{\{W_t>t^a\}}$ to the right-hand side, yields that on the event that $\{k_t\leq 2L_{t-1}+(t-1)\delta\}$,

$$\varkappa_k(t)\leq\Big(\frac{W_t}{\mu}-1\Big)(Sp)_k+tW_t\mathbf{1}_{\{W_t\leq t^a\}}\sup_{k\leq k_t}|(R_tp)_k| + tW_t^2\mathbf{1}_{\{W_t\leq t^a\}}\sup_{k\leq k_t}|((T_t-I)^2p)_k|+\mathbf{1}_{\{W_t>t^a\}}C_\eta L_t, \tag{2.75}$$

and similarly we get as a lower bound, using that $|W_t/\mu-1|\leq W_t$,

$$\varkappa_k(t)\geq\Big(\frac{W_t}{\mu}-1\Big)(Sp)_k-tW_t\mathbf{1}_{\{W_t\leq t^a\}}\sup_{k\leq k_t}|(R_tp)_k| - tW_t^2\mathbf{1}_{\{W_t\leq t^a\}}\sup_{k\leq k_t}|((T_t-I)^2p)_k|-\mathbf{1}_{\{W_t>t^a\}}(C_sW_t+C_\eta L_t), \tag{2.76}$$

where we used that $\sup_{k\geq 1}|(Sp)_k|\leq C_s$. We use (2.75) and (2.76) on the event $\{k_t\leq 2L_{t-1}+(t-1)\delta\}$, and (2.65) on the event $\{k_t>2L_{t-1}+(t-1)\delta\}$ to arrive at

$$\begin{aligned}\varkappa_k(t)\leq&\Big(\frac{W_t}{\mu}-1\Big)(Sp)_k+tW_t\mathbf{1}_{\{W_t\leq t^a\}}\sup_{k\leq k_t}|(R_tp)_k|\\ &+tW_t^2\mathbf{1}_{\{W_t\leq t^a\}}\sup_{k\leq k_t}|((T_t-I)^2p)_k|\\ &+(\mathbf{1}_{\{W_t>t^a\}}+\mathbf{1}_{\{k_t>2L_{t-1}+(t-1)\delta\}})((C_s+C_\eta)W_t+C_\eta L_{t-1}),\end{aligned} \tag{2.77}$$

with a similar lower bound where the last three terms have minus signs. We now take expectations on both sides of (2.77) and take advantage of the equality $\mathbb{E}[W_t/\mu]=1$ and the property that $(Sp)_k$ is deterministic, so that the first term in the right-hand side drops out. Moreover, using that W_t and L_{t-1} are independent, as well as that $k_t>2L_{t-1}+(t-1)\delta$ implies that $L_{t-1}\leq k_t$, we arrive at

$$|\mathbb{E}[\varkappa_k(t)]|\leq\mathbb{E}[\mathbf{1}_{\{W_t>t^a\}}((C_s+C_\eta)W_t+C_\eta\mu t)] \tag{2.78}$$

$$+(C_\eta k_t+(C_s+C_\eta)\mu)\mathbb{P}(k_t>2L_{t-1}+(t-1)\delta) \tag{2.79}$$

$$+t\mathbb{E}\Big[\sup_{k\leq k_t}|(R_tp)_k|\Big]\mathbb{E}[W_t\mathbf{1}_{\{W_t>t^a\}}] \tag{2.80}$$

$$+t\mathbb{E}[W_t^2\mathbf{1}_{\{W_t\leq t^a\}}]\mathbb{E}\Big[\sup_{k\leq k_t}|((T_t-I)^2p)_k|\Big]. \tag{2.81}$$

We now bound each of these four terms one by one. To bound (2.78), we use that W_t has finite $(1+\varepsilon)$-moment, to obtain that

$$\mathbb{E}[\mathbf{1}_{\{W_t>t^a\}}W_t]=\mathbb{E}[\mathbf{1}_{\{W_t>t^a\}}W_t^{-\varepsilon}W_t^{1+\varepsilon}]\leq t^{-a\varepsilon}\mathbb{E}[W_t^{1+\varepsilon}]=O(t^{-a\varepsilon})$$

and

$$t\mathbb{E}[\mathbf{1}_{\{W_t>t^a\}}]=t\mathbb{P}(W_t^{1+\varepsilon}>t^{a(1+\varepsilon)})\leq t^{1-a(1+\varepsilon)}\mathbb{E}[W_t^{1+\varepsilon}]=O(t^{1-a(1+\varepsilon)}),$$

which bounds (2.78) as

$$\mathbb{E}[\mathbf{1}_{\{W_t>t^a\}}((C_s+C_\eta)W_t+C_\eta\mu t)]=O(t^b), \tag{2.82}$$

with $b=\max\{-a\varepsilon, 1-a(1+\varepsilon)\}$.

To bound (2.79), we use that $L_{t-1}<\frac{1}{2}(\eta t-\delta(t-1))=\frac{1}{2}(\eta-\delta)(t-1)+\frac{1}{2}\eta$ when $k_t>2L_{t-1}+(t-1)\delta$. Now, since $\eta\in(\mu,2\mu+\delta)$, we have that $\frac{1}{2}(\eta-\delta)<\mu$. Standard large deviation theory and the fact that the initial degrees W_i are non-negative give that the probability that $L_{t-1}<\sigma(t-1)$, with $\sigma<\mu$, is exponentially small in t. As a result, we obtain that

$$(C_\eta k_t+(C_s+C_\eta)\mu)\mathbb{P}(k_t>2L_{t-1}+(t-1)\delta)=O(t^{-1}). \tag{2.83}$$

To bound (2.80), we use that $2L_{t-1}+t\delta\geq L_{t-1}\geq t-1\geq t/2$, and also use (2.73), to obtain that

$$\mathbb{E}\Big[\sup_{k\leq k_t}|(R_tp)_k|\Big]\leq\frac{c}{t^2}\mathbb{E}|L_{t-1}-t\mu|\sup_{k\geq 1}(k+|\delta|)p_k\leq\frac{c}{t^2}\mathbb{E}|L_{t-1}-t\mu|.$$

Thus,

$$t\mathbb{E}\Big[\sup_{k\leq k_t}|(R_tp)_k|\Big]\mathbb{E}[W_t\mathbf{1}_{\{W_t>t^a\}}]\leq\frac{c}{t}\mathbb{E}|L_{t-1}-t\mu|t^{-a\varepsilon}\leq O(t^{-a\varepsilon-\varepsilon/(1+\varepsilon)}), \tag{2.84}$$

where the final bound follows from (2.52).

Finally, to bound (2.81), note that

$$\mathbb{E}[W_t^2\mathbf{1}_{\{W_t\leq t^a\}}]=\mathbb{E}[W_t^{1-\varepsilon}W_t^{1+\varepsilon}\mathbf{1}_{\{W_t\leq t^a\}}]\leq t^{a(1-\varepsilon)}\mathbb{E}[W_t^{1+\varepsilon}]=O(t^{a(1-\varepsilon)}),$$

and, by (2.29) and the fact that $2L_{t-1}+t\delta\geq\eta t$ for some $\eta>0$, we have that

$$\mathbb{E}\Big[\sup_{k\leq k_t}|((T_t-I)^2p)_k|\Big]\leq\frac{c}{t^2}\sup_{k\geq 1}(k+|\delta|)^2p_k. \tag{2.85}$$

This leads to the bound that

$$t\mathbb{E}[W_t^2\mathbf{1}_{\{W_t\leq t^a\}}]\mathbb{E}\Big[\sup_{k\leq k_t}|((T_t-I)^2p)_k|\Big]\leq O(t^{a(1-\varepsilon)-1}). \tag{2.86}$$

Combining the bounds in (2.82), (2.83), (2.84) and (2.86) completes the proof of part (c) of Lemma 2.4, for any a such that $1/(\varepsilon+1)<a<1$. □

3. Proof of Theorem 1.5

In this section, we write $F(x)=\mathbb{P}(W_1\le x)$, and assume that $1-F(x)=x^{1-\tau}L(x)$ for some slowly varying function $x\mapsto L(x)$. Throughout this section, we write $\tau=\tau_{\mathrm{W}}$.

From (1.1) it is immediate that

$$d_i(t)=d_i(t-1)+X_{i,t}\quad \text{for } i=0,1,2,\dots,t-1, \tag{3.1}$$

where, conditionally on $d_i(t-1)$ and $\{W_j\}_{j=1}^t$, the distribution of $X_{i,t}$ is binomial with parameters W_t and success probability

$$q_i(t)=\frac{d_i(t-1)+\delta}{2L_{t-1}+t\delta}. \tag{3.2}$$

Hence, for $t>i$,

(3.3)

$$\begin{aligned}\mathbb{E}[(d_i(t)+\delta)^s\mid\{W_j\}_{j=1}^t]&=\mathbb{E}[\mathbb{E}[(d_i(t-1)+\delta+X_{i,t})^s\mid d_i(t-1),\{W_j\}_{j=1}^t]\mid\{W_j\}_{j=1}^t]\\&\le\mathbb{E}[(d_i(t-1)+\delta+\mathbb{E}[X_{i,t}\mid d_i(t-1),\{W_j\}_{j=1}^t])^s],\end{aligned}$$

where we have used the Jensen inequality $\mathbb{E}[(a+X)^s]\le(a+\mathbb{E}[X])^s$, which follows from the concavity of $t\mapsto(a+t)^s$ for $0<s<1$. Next, we make the substitution $\mathbb{E}[X_{i,t}|d_i(t-1),\{W_j\}_{j=1}^t]=W_tq_i(t)$ and use the inequality $2L_{t-1}+t\delta\ge L_{t-1}+\delta$, to obtain that

$$\begin{aligned}\mathbb{E}[(d_i(t)+\delta)^s\mid\{W_j\}_{j=1}^t]&\le\mathbb{E}[(d_i(t-1)+\delta)^s\mid\{W_j\}_{j=1}^t]\left(1+\frac{W_t}{2L_{t-1}+t\delta}\right)^s\\&\le\mathbb{E}[(d_i(t-1)+\delta)^s\mid\{W_j\}_{j=1}^t]\left(\frac{L_t+\delta}{L_{t-1}+\delta}\right)^s.\end{aligned}$$

Thus, by induction, and because $d_i(i)=W_i$, we get that, for all $t>i\ge1$,

$$\mathbb{E}[(d_i(t)+\delta)^s\mid\{W_j\}_{j=1}^t]\le(W_i+\delta)^s\prod_{n=i+1}^t\left(\frac{L_n+\delta}{L_{n-1}+\delta}\right)^s=(W_i+\delta)^s\left(\frac{L_t+\delta}{L_i+\delta}\right)^s. \tag{3.4}$$

The case $i=0$ can be treated by $(d_0(t)+\delta)^s=(d_1(t)+\delta)^s$, which is immediate from the definition of $G(1)$. Thus,

$$\mathbb{E}[(d_i(t)+\delta)^s]\le\mathbb{E}\left[(W_i+\delta)^s\left(\frac{L_t+\delta}{L_i+\delta}\right)^s\right]. \tag{3.5}$$

Define $f(W_i)=(W_i+\delta)^s$ and

$$g(W_i)=\left(\frac{L_t+\delta}{L_i+\delta}\right)^s=\left(1+\frac{W_{i+1}+W_{i+2}+\dots+W_t}{W_1+W_2+\dots+W_i+\delta}\right)^s,$$

and notice that when we condition on all W_j, $1\le j\le t$, except W_i, then the map $W_i\mapsto f(W_i)$ is increasing in its argument, whereas $W_i\mapsto g(W_i)$ is decreasing. This implies that

$$\mathbb{E}[f(W_i)g(W_i)]\le \mathbb{E}[f(W_i)]\mathbb{E}[g(W_i)]. \tag{3.6}$$

Hence,

$$\mathbb{E}[(d_i(t)+\delta)^s]\le \mathbb{E}[(W_i+\delta)^s]\mathbb{E}\left[\left(\frac{L_t+\delta}{L_i+\delta}\right)^s\right]\le \mathbb{E}[(W_i+\delta)^s]\mathbb{E}[(L_t+\delta)^s]\mathbb{E}[(L_i+\delta)^{-s}], \tag{3.7}$$

where in the final step we have applied the inequality (3.6) once more.

For $i,t\to\infty$,

$$\mathbb{E}[(L_i+\delta)^{-s}]=(1+o(1))\mathbb{E}[L_i^{-s}] \quad\text{and}\quad \mathbb{E}[(L_t+\delta)^s]=(1+o(1))\mathbb{E}[L_t^s]. \tag{3.8}$$

The moment of order s of $W_i+\delta$ can be bounded by

$$\mathbb{E}[(W_i+\delta)^s]\le \mathbb{E}\left[W_i^s\left(1+\frac{|\delta|}{W_i}\right)^s\right]\le (1+|\delta|)^s\mathbb{E}[W_i^s]=(1+|\delta|)^s\mathbb{E}[W_1^s], \tag{3.9}$$

since $W_i\ge 1$. Combining (3.7), (3.8) and (3.9) gives for i sufficiently large and $t>i$,

$$\mathbb{E}[(d_i(t)+\delta)^s]\le (1+|\delta|)^s\mathbb{E}[W_1^s]\mathbb{E}[L_i^{-s}]\mathbb{E}[L_t^s](1+o(1)). \tag{3.10}$$

We will bound each of the terms $\mathbb{E}[W_1^s]$, $\mathbb{E}[L_t^s]$ and $\mathbb{E}[L_i^{-s}]$ separately.

Evidently, $\mathbb{E}[W_1^s]$ can be bounded by some constant, since all moments smaller than $\tau-1$ are finite. We will show that, for some constant C_s,

$$\mathbb{E}[L_t^s]\le C_s t^{s/(\tau-1)}l(t)^s \tag{3.11}$$

and that, for i sufficiently large,

$$\mathbb{E}[L_i^{-s}]\le C_s i^{-s/(\tau-1)}l(i)^{-s}. \tag{3.12}$$

We will first show claim (3.12) and then (3.11). For claim (3.12), we define the norming sequence $\{a_n\}_{n\ge 1}$ by

$$a_n=\sup\{x:1-F(x)\ge n^{-1}\}, \tag{3.13}$$

so that it is immediate that $a_n=n^{1/(\tau-1)}l(n)$, where $n\mapsto l(n)$ is slowly varying. We use that $L_i\ge W_{(i)}=\max_{1\le j\le i}W_j$, so that

$$\mathbb{E}[L_i^{-s}]\le \mathbb{E}[W_{(i)}^{-s}]=-\mathbb{E}[(-Y_{(i)})^s], \tag{3.14}$$

where $Y_j=-W_j^{-1}$ and $Y_{(i)}=\max_{1\le j\le i} Y_j$. Clearly, we have that $Y_j\in[-1,0]$, so that $\mathbb{E}[(-Y_1)^s]<\infty$. Also, $a_iY_{(i)}=-a_i/W_{(i)}$ converges in distribution to $-E^{-1/(\tau_{\mathrm{W}}-1)}$, where E is exponential with mean 1, so it follows from [28, Theorem 2.1] that, as $i\to\infty$,

$$\mathbb{E}\left[\left(\frac{a_i}{L_i}\right)^s\right]\le -\mathbb{E}[(-a_iY_{(i)})^s]\to\mathbb{E}[E^{-1/(\tau-1)}]<\infty, \tag{3.15}$$

which proves the claim (3.12).

We now turn to claim (3.11). The discussion in [21, p. 565 and Corollary 1] yields that, for $s<\tau-1$, $\mathbb{E}[L_t^s]=\mathbb{E}[|L_t|^s]\le 2^{s/2}\lambda_s(t)$, for some function $\lambda_s(t)$ depending on s, t and F. Using the discussion in [21, p. 564], we have that $\lambda_s(t)\le C_st^{s/(\tau-1)}M^*(t^{1/(\tau-1)})^s$, where $M^*(\,\cdot\,)$ is a slowly varying function. With some more effort, it can be shown that we can replace $M^*(t^{1/(\tau-1)})$ by $l(t)$, which gives (3.11).

Combining (3.10), (3.11) and (3.12), we obtain that

$$\mathbb{E}[(d_i(t)+\delta)^s]\le C\left(\frac{t}{i\vee 1}\right)^{s/(\tau-1)}\left(\frac{l(t)}{l(i)}\right)^s. \tag{3.16}$$

Finally, we note that, since $d_i(t)\ge\min\{x:x\in S_{\mathrm{W}}\}\equiv\delta+\nu$ where $\nu>0$, and using (1.2), we can bound $\mathbb{E}[d_i(t)^s]\le(1\vee\nu^{-1})^s\mathbb{E}[(d_i(t)+\delta)^s]$, which together with (3.16) establishes the proof of Theorem 1.5.

Acknowledgements. The work of MD, HvdE and RvdH was supported in part by the Netherlands Organisation for Scientific Research (NWO). We thank D. Denisov for pointing us to [21].

References

1. Aiello, W., Chung, F. and Lu, L., Random evolution in massive graphs, in *Handbook of Massive Data Sets*, Massive Comput. **4**, pp. 97–122, Kluwer, Dordrecht, 2002.
2. Barabási, A. L. and Albert, R., Emergence of scaling in random networks, *Science* **286** (1999), 509–512.
3. Berestycki, N. and Durrett, R., A phase transition in the random transposition random walk, *Probab. Theory Related Fields* **136** (2006), 203–233.
4. Bianconi, G. and Barabási, A.-L., Bose–Einstein condensation in complex networks, *Phys. Rev. Lett.* **86**:24 (2001), 5632–5635.
5. Bianconi, G. and Barabási, A.-L., Competition and multiscaling in evolving networks, *Europhys. Lett.* **54**:4 (2001), 436–442.

6. Bollobs, B., Borgs, C., Chayes, J. and Riordan, O., Directed scale-free graphs, in *Proceedings of the Fourteenth Annual ACM-SIAM Symposium on Discrete Algorithms (Baltimore, MD, 2003)*, pp. 132–139, ACM, New York, 2003.
7. Bollobás, B., Janson, S. and Riordan, O., The phase transition in inhomogeneous random graphs, *Random Structures Algorithms* **31** (2007), 3–122.
8. Bollobás, B. and Riordan, O., Mathematical results on scale-free random graphs, in *Handbook of Graphs and Networks*, pp. 1–34, Wiley, Weinheim, 2003.
9. Bollobás, B., Riordan, O., Spencer, J. and Tusnády, G., The degree sequence of a scale-free random graph process, *Random Structures Algorithms* **18** (2001), 279–290.
10. Borgs, C., Chayes, J., Daskalakis, D. and Roch, C., First to market is not everything: an analysis of preferential attachment with fitness, in *Proceedings of the 39th annual ACM symposium on the theory of computation*, pp. 135–144, 2007.
11. Britton, T., Deijfen, M. and Martin-Löf, A., Generating simple random graphs with prescribed degree distribution, *J. Stat. Phys.* **124** (2006), 1377–1397.
12. Chung, F. and Lu, L., The average distances in random graphs with given expected degrees, *Proc. Natl. Acad. Sci. USA* **99**:25 (2002), 15879–15882.
13. Chung, F. and Lu, L., Connected components in random graphs with given expected degree sequences, *Ann. Comb.* **6** (2002), 125–145.
14. Cooper, C. and Frieze, A., A general model of web graphs, *Random Structures Algorithms* **22** (2003), 311–335.
15. Ergün, G. and Rodgers, G. J., Growing random networks with fitness, *Phys. A* **303** (2002), 261–272.
16. Faloutsos, C., Faloutsos, P. and Faloutsos, M., On power-law relationships of the internet topology, *Computer Communications Reviews* **29** (1999), 251–262.
17. Feller, W., *An Introduction to Probability Theory and its Applications*. Vol. II, Wiley, New York, 1971.
18. Grimmett, G. R. and Stirzaker, D. R., *Probability and Random Processes*, Oxford University Press, New York, 2001.
19. Gut, A., *Probability: A Graduate Course*, Springer Texts in Statistics, Springer, New York, 2005.
20. Hagberg, O. and Wiuf, C., Convergence properties of the degree distribution of some growing network models, *Bull. Math. Biol.* **68**:6 (2006), 1275–1291.
21. Hall, P., Order of magnitude of moments of sums of random variables, *J. London Math. Soc.* **24** (1981), 562–568.
22. van der Hofstad, R. and Hooghiemstra, G., Diameters in preferential attachment models, *Preprint*, 2007.
23. Jordan, J., The degree sequences and spectra of scale-free random graphs, *Random Structures Algorithms* **29** (2006), 226–242.
24. Katona, Z. and Móri, T. F., A new class of scale free random graphs, *Statist. Probab. Lett.* **76**:15 (2006), 1587–1593.
25. Kumar, R., Raghavan, P., Rajagopalan, S., Sivakumar, D., Tomkins, A. and Upfal, E., Stochastic models for the web graph, in *41st Annual Symposium on Foundations of Computer Science (Redondo Beach, CA, 2000)*, pp. 57–65, IEEE Comput. Soc. Press, Los Alamitos, CA, 2000.

26. Kumar, R., Raghavan, P., Rajagopalan, S. and Tomkins, A., Trawling the web for emerging cyber communities, *Computer Networks* **31** (1999), 1481–1493.
27. Newman, M. E. J., The structure and function of complex networks, *SIAM Rev.* **45** (2003), 167–256.
28. Pickands, III, J., Moment convergence of sample extremes, *Ann. Math. Statist.* **39** (1968), 881–889.
29. Söderberg, B., General formalism for inhomogeneous random graphs, *Phys. Rev. E* **66**:6 (2002), 6 pp.
30. Szymański, J., Concentration of vertex degrees in a scale-free random graph process, *Random Structures Algorithms* **26** (2005), 224–236.

Maria Deijfen
Stockholm University
Department of Mathematics
SE-106 91 Stockholm
Sweden
mia@math.su.se

Henri van den Esker
Electrical Engineering, Mathematics
and Computer Science
Delft University of Technology
P.O. Box 5031
NL-2600 GA Delft
The Netherlands
H.vandenEsker@tudelft.nl

Remco van der Hofstad
Department of Mathematics
and Computer Science
Eindhoven University of Technology
P.O. Box 513
NL-5600 MB Eindhoven
The Netherlands
rhofstad@win.tue.nl

Gerard Hooghiemstra
Electrical Engineering, Mathematics
and Computer Science
Delft University of Technology
P.O. Box 5031
NL-2600 GA Delft
The Netherlands
G.Hooghiemstra@tudelft.nl

Received March 1, 2007
published online March 21, 2008