Internet Mathematics Vol. 1, No. 4: 463-483

Random Deletion in a Scale-Free Random Graph Process

Colin Cooper, Alan Frieze, and Juan Vera

Abstract. We study a dynamically evolving random graph which adds vertices and edges using preferential attachment and deletes vertices randomly. At time t, with probability $\alpha_1 > 0$ we add a new vertex u_t and m random edges incident with u_t. The neighbours of u_t are chosen with probability proportional to degree. With probability $\alpha - \alpha_1 \geq 0$ we add m random edges to existing vertices where the endpoints are chosen with probability proportional to degree. With probability $1 - \alpha - \alpha_0$ we delete a random vertex, if there are vertices left to delete. With probability α_0 we delete m random edges. Assuming that $\alpha + \alpha_1 + \alpha_0 > 1$ and α_0 is sufficently small, we show that for large k, t, the expected number of vertices of degree k is approximately $d_k t$ where as $k \to \infty$, $d_k \sim Ck^{-1-\beta}$ where $\beta = \frac{2(\alpha-\alpha_0)}{3\alpha-1-\alpha_1-\alpha_0}$ and $C > 0$ is a constant. Note that β can take any value greater than 1.

1. Introduction

Recently there has been much interest in understanding the properties of real-world, large-scale networks such as the structure of the Internet and the World Wide Web. For a general introduction to this topic, see Bollobás and Riordan [Bollobás and Riordan 02], Hayes [Hayes 00], Watts [Watts 99], or Aiello, Chung and Lu [Aiello et al. 02]. One approach is to model these networks by random graphs. Experimental studies by Albert, Barabási, and Jeong [Albert et al. 99]; Broder et al. [Broder et al. 00]; and Faloutsos, Faloutsos, and Faloutsos [Faloutsos et al. 99] have demonstrated that in the World Wide Web/Internet, the proportion of vertices of a given degree follows an approximate inverse power law, i.e., the proportion of vertices of degree k is approximately $Ck^{-\alpha}$ for some

1542-7951/04 $0.50 per page

constants C, α. The classical models of random graphs introduced by Erdős and Renyi [Erdös and Rényi 59] do not have power law degree sequences, so they are not suitable for modeling these networks. This has driven the development of various alternative models for random graphs.

One approach to remedy this situation is to study graphs with a prescribed degree sequence (or prescribed expected degree sequence). This is proposed as a model for the web graph by Aiello, Chung, and Lu in [Aiello et al. 00]. Mihail and Papadimitriou also use this model [Mihail and Papadimitriou 02] in their study of large eigenvalues, as do Chung, Lu, and Vu in [Chung et al. 03, Chung et al. 02].

An alternative approach, which we will follow in this paper, is to sample graphs via some generative procedure which yields a power law distribution. There is a long history of such models, outlined in the survey by Mitzenmacher [Mitzenmacher 04]. We will use an extension of the preferential attachment model to generate our random graph. The preferential attachment model has been the subject of recently revived interest, and dates back to Yule [Yule 25] and Simon [Simon 55]. It was proposed as a random graph model for the web by Barabási and Albert [Barabási and Albert 99], and their description was elaborated by Bollobás and Riordan [Bollobás and Riordan 04a] who showed that whp the diameter of a graph constructed in this way was $\sim \frac{\log \log n}{\log n}$. Subsequently, Bollobás, Riordan, Spencer, and Tusnády [Bollobás et al. 01] proved that the degree sequence of such graphs does follow a power law distribution.

An evolving network, such as a P2P network, sometimes loses vertices. Bollobás and Riordan [Bollobás and Riordan 04b] consider the effect of deleting vertices from the basic preferential attachment model of [Barabási and Albert 99], [Bollobás and Riordan 04a], *after* all vertices have been generated. In this paper, we study the deletion of vertices within a dynamic setting. Chung and Lu [Chung and Lu 04] have independently considered a model with vertex and edge deletions.

We will study here the following model and examine its likely degree sequence: we consider a process which generates a sequence of *simple* graphs $G_t, t = 1, 2, \ldots$. The graph $G_t = (V_t, E_t)$ has v_t vertices and e_t edges.

Time-Step 1.
To initialise the process, we start with G_1 consisting of an isolated vertex x_1.

Time-Step $t \geq 2$.

1. With probability $1 - \alpha - \alpha_0$ we delete a randomly chosen vertex x from V_{t-1}. If $V_{t-1} = \emptyset$, we do nothing in this case.

2. With probability α_0 we delete $\min\{m, |E_{t-1}|\}$ randomly chosen edges from E_{t-1}.

3. With probability α_1 we add a vertex x_t to G_{t-1}. We then add m random edges incident with x_t. Assume first that $e_{t-1} > 0$: the m random neighbours $w_1, w_2, \ldots, w_m$ are chosen independently. For $1 \leq i \leq m$ and $w \in V_{t-1}$,
$$\mathbf{Pr}(w_i = w) = \frac{d(w, t-1)}{2e_{t-1}}, \tag{1.1}$$
where $d(w, t-1)$ denotes the degree of vertex w at the beginning of substep t. Thus neighbours are chosen by *preferential attachment*.

   After adding these edges, we coalesce multiple edges into a single copy.

   Special cases. If $e_{t-1} = v_{t-1} = 0$, then we start again as in Time-Step 1. If $e_{t-1} = 0$ and $v_{t-1} > 0$, then we add a new vertex x_t and join it to a randomly chosen vertex in V_{t-1}.

4. With probability $\alpha - \alpha_1$ we add m random edges to existing vertices. Both endpoints are chosen independently with the same probabilities as in (1.1).

   After adding these edges, we coalesce multiple edges into a single copy and delete any loops.

   Special cases. If $e_{t-1} = 0$ or $v_{t-1} = 0$, then we do nothing.

Remark 1.1. We have dealt with the case $e_{t-1} = 0$ in a somewhat arbitrary manner. Since this will be shown to happen very rarely, it should not matter what we do here. We have chosen a process that avoids technicalities in our proofs.

Remark 1.2. There are perhaps more natural choices for a deletion process. For example, one might want to make the deletion of high-degree vertices less likely than low-degree vertices. Our choice of deleting a uniform random vertex leads naturally to a recurrence relation for expected degrees. The analysis becomes more difficult if we deviate from this approach. We leave it as an interesting open question to study other models of deletion.

Let $D_k(t)$ be the number of vertices of degree $k \geq 0$ in G_t and let $\overline{D}_k(t)$ be the expectation of this variable. Let
$$\beta = \frac{2(\alpha - \alpha_0)}{3\alpha - 1 - \alpha_1 - \alpha_0} \tag{1.2}$$
and
$$\rho = \frac{\alpha_1}{\alpha}.$$

There are certain inequalities which must hold for the results to be meaningful:

$$\alpha > \alpha_1 \qquad \alpha + \alpha_0 < 1 \tag{1.3}$$
$$\alpha > 1/2 \qquad \alpha + \alpha_0 + \alpha_1 > 1 \tag{1.4}$$
$$\alpha - \alpha_0 > 0 \qquad 3\alpha - \alpha_0 - \alpha_1 > 1. \tag{1.5}$$

In addition we have to add the following condition, which is only here for the proof to be valid. It is tempting to conjecture that (1.3)–(1.5) are sufficient.

$$\alpha_0 \text{ is sufficently small} \tag{1.6}$$

The main result of this paper is

Theorem 1.3. *Assume that* (1.3)–(1.6) *hold. Then there exists a constant* $C = C(m, \alpha, \alpha_0, \alpha_1)$ *such that for* $k \geq 1$

$$\left| \frac{\overline{D}_k(t)}{t} - Ck^{-1-\beta} \right| = O\left(t^{-\rho/8}\right) + O\left(k^{-2-\beta}\right).$$

In general, by suitable choice of $\alpha, \alpha_0, \alpha_1$, we find that β can take any value greater than 1, and thus the model exhibits a power law with exponent $1+\beta > 2$.

It is perhaps worth considering two simple instances before launching into the proof of Theorem 1.3:

- no edges added between old vertices and no edge deletions. In this case, $\alpha_1 = \alpha$ and $\beta = \frac{2\alpha}{2\alpha-1} \geq 2$.

- no deletions. In this case, $\alpha = 1$ and $\beta = \frac{2}{2-\alpha_1} > 1$.

Remark 1.4. The exact range of α_0 for which this theorem is true is likely to be such that both the expected number of vertices and edges grow linearly with t, i.e., $\alpha + \alpha_0 + \alpha_1 - 1 > 0$ and $\alpha - \alpha_0 > 0$.

1.1. Proof Outline

Our proof methodology is similar to that described in [Cooper and Frieze 03]. We first find a recurrence for the $\overline{D}_k(t)$; see (5.2). We approximate this by a recurrence in k only; see (5.3). We show that the latter recurrence gives a good approximation, see Lemma 5.1. We then solve the recurrence in k using Laplace's method [Jordan 39].

2. Number of Vertices in G_t

Let

$$\nu = \alpha + \alpha_1 + \alpha_0 - 1 > 0$$

and let v_t be the number of vertices in G_t.

Then

$$v_t = X_1 + X_2 + \cdots + X_t,$$

where $X_1 = 1$. For $i \geq 2$ we have $X_i = +1$ with probability α_1, $X_i = 0$ with probability $\alpha + \alpha_0 - \alpha_1$, and $X_i = -1$ with probability $1 - \alpha - \alpha_0$ unless $v_{i-1} = 0$, in which case $X_i = 0$. The distribution of v_t dominates that of $Y_1 + Y_2 + \cdots + Y_t$, where the Y_i are independent and $Y_i = +1$ with probability α_1, $Y_i = 0$ with probability $\alpha + \alpha_0 - \alpha_1$, and $Y_i = -1$ with probability $1 - \alpha - \alpha_0$. The Y_i are independent and so it is appropriate to use Hoeffding's Theorem for the sum of bounded random variables. This gives us

$$\mathbf{Pr}(|Y_1 + Y_2 + \cdots + Y_t - \nu t| \geq u) \leq 2 \exp\left\{-\frac{u^2}{2t}\right\} \tag{2.1}$$

for any $u > 0$.

It follows immediately that

$$\mathbf{Pr}(v_t = 0) \leq 2e^{-\nu^2 t/2}$$

and so qs[1]

$$\nexists \tau \in [(\log t)^2, t] \text{ such that } v_\tau = 0. \tag{2.2}$$

From (2.1) and (2.2) we obtain

$$|v_t - (\alpha + \alpha_0 + \alpha_1 - 1)t| \leq ct^{1/2} \log t, \qquad \text{qs}, \tag{2.3}$$

for any constant $c > 0$.

3. Bounding the Maximum Degree

Let $d(s,t)$ be the total number of edges created in time-steps $s, s+1, \ldots, t$ that are incident with vertex x_s. This could be greater than the degree of x_s in G_t since we include in this count edges that are generated and later deleted. We prove

Lemma 3.1. $$d(s,t) < (t/s)^{1-\rho/2} (\log t)^5 \qquad \text{qs.}$$

[1] An event happens *quite surely* qs if the probability it fails to occur is $O(t^{-K})$ for any $K > 0$.

Proof. Fix s, t. Let $\phi(s) = 1 + |\{i < s : x_i \in V_t\}|$ and let D_s be the $s - \phi(s)$ vertices which appear before x_s but which get deleted. Note that qs $\phi(s) \geq s/(\log t)^2$, as can be seen from (2.3). Our approach is to show that $d(s,t)$ is dominated by m plus the degree of $x_{\phi(s)}$ in the equivalent process where $\alpha = 1$ and no deletions occur. Intuitively, vertex deletion should not make degrees larger.

Let Process 1 denote our process G_τ, but with the understanding that we do not delete the selected vertex x at a deletion step, rather we just remove its edges.

Process 2 is as Process 1 except that edges not incident with x_s which are deleted at or before step t are never inserted. Thus X_τ the number of inserted edges at any insertion step τ satisfies $0 \leq X_\tau \leq m$, where $m - X_\tau$ counts deleted edges not incident with x_s.

Process 3 is as Process 2, except that those edges incident with x_s deleted at steps τ, $s < \tau \leq t$ are no longer deleted.

Condition on the set A of vertices which are deleted and assume that $s \notin A$. We can study our degree question via the placing of balls into two bins. Bin 1 has a ball for every edge incident with vertex x_s and Bin 2 has a ball for every edge-vertex incidence of vertices other then x_s. The balls-in-bins process follows G_τ for $\tau > s$. It starts with Bin 1 containing $a_1 = m$ balls and Bin 2 containing a_2 balls. We break the addition and deletion of vertices, edges into substeps involving a single edge. At an addition substep we add a ball to Bin 2 and then we randomly add a further ball to one of the bins, this bin being chosen with probability proportional to its contents at the time corresponding to the first substep associated with the addition of a block of m edges. At each deletion substep we delete two balls from the bins.

Consider an edge u which is incident with $x_k \in A$ but not incident with x_s. At some time $\tau > k$ u will disappear and this will mean the deletion of two balls from Bin 2. Suppose instead we never bothered to add edge u. We compare the distribution of the number of balls in Bin 1 in Process 1 where u is added and in Process 2 where it is not. Process 2 starts with a_1 balls in Bin 1 and $a_2 - 2$ balls in Bin 2. It is now easy to couple the processes so that the second process always has at least as many balls in Bin 1 as does the first process. (While this is the case, the random ball is always at least as likely to land in Bin 1 in the second process as it is in the first). We continue both processes until u disappears in the first process and then we continue with the coupling.

In this way, we will eliminate all vertices in A which do not have x_s as a neighbour. Now consider an edge u which is incident with $x_k \in A$ and also incident with x_s. The situation is now reversed. After the deletion of this edge in Process 2 we start with $a_1 - 1, a_2 - 1$ balls in each bin, whereas in Process 3 we start with a_1, a_2 balls. Since there are no loops allowed in our model, Bin 1 can never have more balls than Bin 2. We can couple the two processes so that the first bin always has strictly more balls in Process 3 than it does in Process 2. If the bin contents are b_1, b_2 in Process 3 then they will be $b_1' - 1, b_2' - 1$ in Process 2, where $b_1 \geq b_1'$ and $b_1 + b_2 = b_1' + b_2'$.

The upshot of all this is that $d(s,t)$ is stochastically dominated by m plus the degree of $x_{\phi(s)}$ in the process where vertices are never deleted. The m accounts for the edges A added at Time-Step s and in the coupled processes all the other edges incident with vertices not in V_t are never added, making x_s the $\phi(s)$th vertex in the process. So we can think of deleting the vertices in D_s and giving the m balls corresponding to A to Bin 2. Thus we now estimate degree bounds when we add new edges between old vertices with probability $1-\rho$ and add a new vertex with probability ρ.

Some technicalities arise from the case where vertex x_u is deleted and some of its edges are directed to x_s, and from the erasure of loops and parallel edges. In the above coupling, x_u will effectively generate less than m edges. However, since we seek an upper bound on $d(s,t)$ this explains the $X_{\tau+1} \leq X_\tau + \cdots$ in place of $X_{\tau+1} = X_\tau + \cdots$ in (3.1).

Fix $s \leq t$ and let $X_\tau = deg(s,\tau)$ for $\tau = s, s+1, \ldots, t$ and let $\lambda = \frac{(s/t)^{(1+\rho)/2}}{20m^2 \ln t}$. Let Y be a 0,1 random variable with $\mathbf{Pr}(Y=1) = 1-\rho$. Then conditional on $X_\tau = x$, we have

$$X_{\tau+1} \leq X_\tau + YB\left(2m, \frac{x}{2m\tau}\right) + (1-Y)B\left(m, \frac{x}{2m\tau}\right) \tag{3.1}$$

and so

$$\begin{aligned}
\mathbf{E}(e^{\lambda X_{\tau+1}} \mid X_\tau = x) &\leq e^{\lambda x}\left((1-\rho)\left(1 - \frac{x}{2m\tau} + \frac{x}{2m\tau}e^\lambda\right)^{2m}\right. \\
&\quad \left. + \rho\left(1 - \frac{x}{2m\tau} + \frac{x}{2m\tau}e^\lambda\right)^m\right) \\
&\leq e^{\lambda x}\left((1-\rho)\exp\left\{\frac{x}{\tau}(e^\lambda - 1)\right\} + \rho \exp\left\{\frac{x}{2\tau}(e^\lambda - 1)\right\}\right) \\
&= \exp\left\{\lambda x\left(1 + \frac{(2-\rho)(1+2m^2\lambda)}{2\tau}\right)\right\},
\end{aligned}$$

after using $x \leq m\tau$.

Thus

$$\mathbf{E}(e^{\lambda X_{\tau+1}}) \leq \mathbf{E}\exp\left\{X_\tau \lambda\left(1 + \frac{(2-\rho)(1+2m^2\lambda)}{2\tau}\right)\right\}.$$

If we put $\lambda_t = \lambda$ and $\lambda_{\tau-1} = \lambda_\tau\left(1 + \frac{(2-\rho)(1+2m^2\lambda_\tau)}{2\tau}\right)$, then provided $\lambda_s \leq 1$ we will have

$$\mathbf{E}(e^{\lambda X_t}) \leq e^{m\lambda_s}.$$

Now provided $\lambda_\tau \leq \Lambda = \frac{1}{2m^2 \ln t}$, we can write

$$\lambda_{\tau-1} \leq \lambda_\tau \left(1 + \frac{(2-\rho)(1+2m^2\Lambda)}{2\tau}\right)$$

and then

$$\begin{aligned} \lambda_s &\leq \lambda \prod_{\tau=s}^{t} \left(1 + \frac{(2-\rho)(1+2m^2\Lambda)}{2\tau}\right) \\ &\leq 10\lambda(t/s)^{1-\rho/2} \end{aligned}$$

which is $\leq \Lambda$ by the definition of λ.

Putting $u = (t/s)^{1-\rho/2}(\ln t)^3$, we get

$$\begin{aligned} \mathbf{Pr}(X_t \geq u) &\leq e^{m\lambda_s - \lambda u} \\ &\leq \exp\{\lambda(10m(t/s)^{1-\rho/2} - u)\} \\ &= O(t^{-K}) \end{aligned}$$

for any constant $K > 0$ and the lemma follows. □

4. Number of Edges in G_t

Let

$$\eta = \frac{m(\alpha - \alpha_0)\nu}{1 + \alpha_1 - \alpha - \alpha_0}$$

and let e_t be the number of edges in G_t.

We first prove a crude lower bound on e_t,

$$e_t \geq t/(\log t)^\xi \qquad \text{qs}, \tag{4.1}$$

for some $\xi = \xi(\alpha, \alpha_1)$.

We note first that, except when $v_t = 0$, an increase in v_t creates at least one edge. We cannot assert that $e_t \geq (1 - o(1))v_t$ because some of these edges will be deleted. So let $t_0 = t/(\log t)^{8/\rho}$. It follows from Lemma 3.1 that qs $x_1, x_2, \ldots, x_{t_0}$ together are incident with at most

$$t^{1-\rho/2}(\log t)^5 \sum_{s=1}^{t_0} s^{\rho/2-1} = O(t/\log t)$$

of the edges ever created.

Now by the use of Chernoff bounds, we see that the number of executions of Step 2 is qs at most $(1 + o(1))\alpha_0 t$. So if $\nu' = \alpha + \alpha_1 - (m-1)\alpha_0 - 1 > 0$, then qs there are at least $\nu' t/2$ edges created which (i) have both endpoints in $\{x_{t_0+1}, x_{t_0+2}, \ldots, x_t\}$ and (ii) are not deleted in any Step 2. From Lemma 3.1, these vertices qs have degree at most $(\log t)^{12/\rho}$. It follows, say by Vizing's theorem, that among these edges we can find a matching $\{e_1, e_2, \ldots, e_p\}$, $p \geq p_0 = \lceil \frac{\nu t}{3(\log t)^{12/\rho}} \rceil$. Choose a submatching M of size exactly p_0. We show that qs a fraction $(\log t)^{-O(1)}$ of these edges survive the deletion process, thus proving (4.1).

We condition on the event

$$\mathcal{E} = \{|v_\tau - \nu\tau| \leq \tau^{1/2} \log \tau \text{ for } t_1 \leq \tau \leq t\}.$$

Then for any $T \subseteq M$ and $e_i \in M \setminus T$, we have

$$\mathbf{Pr}(e_i \text{ survives} \mid e_j, j \neq i, \text{ survives iff } j \in T, \mathcal{E}) \geq \left(\prod_{\tau=t_0}^{t} \left(1 - \frac{2}{\nu'\tau} \right) \right)^2 \geq e^{-5\log(t/t_0)/\nu'} \geq (\log t)^{-O(1)}.$$

To verify this, condition on sequence of step types $i_1, i_2, \ldots, i_t \in \{1, 2, 3, 4\}$, i.e., require that at Time-Step s we carry out a step of type i_s. This is enough to determine whether $\mathcal{E}$ occurs and whether the number of executions of Step 2 is as expected. Now observe that at time $\tau \geq t_0$, there are at least $(1-o(1))\nu'\tau - p_0 = (1-o(1))\nu'\tau$ vertices to choose from, if one has to be deleted.

So, conditional on $\mathcal{E}$, the number of surviving edges dominates

$$Bin\left(\frac{\nu' t}{10(\log t)^{12/\rho}}, (\log t)^{-O(1)} \right)$$

and the Chernoff bound completes the proof of (4.1).

Let Δ_t denote the maximum degree in G_t.

$$\begin{aligned}
\mathbf{E}(e_{t+1} \mid e_t, v_t > 0) &= e_t + \alpha m - \alpha_0 \min\{m, e_t\} \\
&\quad - (1 - \alpha - \alpha_0) \sum_{k \geq 0} k \frac{D_k(t)}{v_t} - O\left(\frac{\Delta_t}{e_t} \right) \\
&= e_t + m(\alpha - \alpha_0) - (1 - \alpha - \alpha_0) \frac{2e_t}{v_t} \\
&\quad - O\left(\frac{\Delta_t}{e_t} \right) + \alpha_0 (m - e_t) 1_{e_t \leq m}.
\end{aligned}$$

The $O\left(\frac{\Delta_t}{e_t}\right)$ term accounts for the probability that we create fewer than m edges in Steps 3 and 4 due to loop deletion and edge coalescence.

Using (2.3) and (4.1) and Lemma 3.1, we see that

$$\begin{aligned}\mathbf{E}(e_{t+1}) &= \mathbf{E}(e_t) + m(\alpha-\alpha_0) - \frac{2(1-\alpha-\alpha_0)}{\nu t}\mathbf{E}(e_t) + \alpha_0\mathbf{E}((m-e_t)1_{e_t\le m}) \\ &\quad + O(t^{-\rho/2}(\log t)^{\xi+5}).\end{aligned}$$

Now let $f_t = e_t - \eta t$. Then we have

$$\begin{aligned}\mathbf{E}(f_{t+1}) &= \mathbf{E}(f_t)\left(1 - \frac{2(1-\alpha-\alpha_0)}{\nu t}\right) + \alpha_0\mathbf{E}((m-e_t)1_{e_t\le m}) \\ &\quad + O(t^{-\rho/2}(\log t)^{\xi+5})\end{aligned}$$

and

$$|\mathbf{E}(f_{t+1})| \le |\mathbf{E}(f_t)| + O(t^{-\rho/2}(\log t)^{\xi+5}) + 1_{f_t\le m}.$$

So we have

$$|\mathbf{E}(f_t)| = O(t^{1-\rho/2}(\log t)^{\xi+5})$$

and

$$\mathbf{E}(e_t) = \eta t + O(t^{1-\rho/2}(\log t)^{\xi+5}).$$

Our next task is to prove a concentration inequality for e_t. We have found this surprisingly difficult and we only obtained something weaker than might be expected. The main problem is that e_t can change by a large amount in one step, if a high degree vertex is deleted. This prohibits the use of standard concentration inequalities.

Lemma 4.1.

$$\mathbf{Pr}(|e_t - \eta t| \ge t^{1-\rho/8}) = O(t^{-\rho/4}(\log t)^{\xi+5}). \tag{4.2}$$

Proof. We will use the Chebychev inequality. Writing $e_{t+1} = e_t + Y_t$, we have

$$\mathbf{Var}(e_{t+1}) \le \mathbf{Var}(e_t) + \mathbf{E}(Y_t^2) + 2(\mathbf{E}(e_tY_t) - \mathbf{E}(e_t)\mathbf{E}(Y_t)). \tag{4.3}$$

Now

$$\mathbf{E}(Y_t^2 \mid G_t) \le \alpha m^2 + \alpha_0 m^2 + (1-\alpha-\alpha_0)v_t^{-1}\sum k^2 D_k(t).$$

It follows from Lemma 3.1 that

$$\mathbf{E}(Y_t^2) = O(t^{1-\rho}(\log t)^{10}). \tag{4.4}$$

We write, with $O(\Delta_t/e_t)$ coming from coalescence of edges,

$$\begin{aligned}\mathbf{E}(e_tY_t \mid G_t) &= e_t((\alpha-\alpha_0)m + \alpha_0(m-e_t)1_{e_t\le m} \\ &\quad - O(\Delta_t/e_t) - 2(1-\alpha-\alpha_0)v_t^{-1}e_t)\end{aligned}$$

and so

$$\begin{aligned}
\mathbf{E}(e_tY_t) &\le m(\alpha-\alpha_0)\mathbf{E}(e_t)+m-2(1-\alpha-\alpha_0)\mathbf{E}(v_t^{-1}e_t^2)\\
&\le m(\alpha-\alpha_0)\mathbf{E}(e_t)+m-2(1-\alpha-\alpha_0)\frac{\mathbf{E}(e_t^2)}{\nu t+t^{1/2}\log t}\\
&\quad +O(t^{-10}) \qquad (4.5)\\
\mathbf{E}(e_t)\mathbf{E}(Y_t) &\ge \mathbf{E}(e_t)(m(\alpha-\alpha_0)-O(t^{-\rho/2}(\log t)^{\xi+5}))\\
&\quad -2(1-\alpha-\alpha_0)\mathbf{E}(e_t)\mathbf{E}(v_t^{-1}e_t)\\
&\ge m(\alpha-\alpha_0)\mathbf{E}(e_t)-2(1-\alpha-\alpha_0)\frac{\mathbf{E}(e_t)^2}{\nu t-t^{1/2}\log t}\\
&\quad -O(t^{1-\rho/2}(\log t)^{\xi+5})). \qquad (4.6)
\end{aligned}$$

It follows from (4.3), (4.4), (4.5), and (4.6) that

$$\mathbf{Var}(e_{t+1})\le \mathbf{Var}(e_t)+O(t^{1-\rho/2}(\log t)^{\xi+5})$$

and so

$$\mathbf{Var}(e_t)=O(t^{2-\rho/2}(\log t)^{\xi+5}).$$

Applying the Chebychev inequality, we obtain

$$\mathbf{Pr}(|e_t-\eta t|\ge t^{1-\rho/8})\le 2\frac{\mathbf{Var}(e_t)}{t^{2-\rho/4}}=O(t^{-\rho/4}(\log t)^{\xi+5}).$$

(The factor 2 following the first inequality accounts for the mean of e_t only being known to be $\eta t\pm O(t^{1-\rho/2}(\log t)^{\xi+5})$.) □

5. Establishing a Recurrence for $\overline{D}_k(t)$

We can now find an approximate recurrence for the $\overline{D}_k(t)$ when $k\le k_0$ where

$$k_0=k_0(t)=\lfloor t^{1-\rho/2}(\log t)^5\rfloor.$$

$D_{-1}(t)=0$ for all $t>0$ and for $k\ge 0$

$$\begin{aligned}
\overline{D}_k(t+1) &= \overline{D}_k(t)+(2\alpha-\alpha_1)m\mathbf{E}\left(-\frac{kD_k(t)}{2e_t}+\frac{(k-1)D_{k-1}(t)}{2e_t}\right.\\
&\qquad\left.-O\left(\frac{\Delta_t}{e_t}\right)\,\middle|\, e_t>0\right)\mathbf{Pr}(e_t>0)\\
&\quad+(1-\alpha-\alpha_0)(k+1)\mathbf{E}\left(\frac{D_{k+1}(t)}{v_t}-\frac{D_k(t)}{v_t}\,\middle|\, v_t>0\right)\mathbf{Pr}(v_t>0)\\
&\quad+\alpha_0 m\mathbf{E}\left(\frac{(k+1)D_{k+1}(t)}{e_t}-\frac{kD_k(t)}{e_t}\,\middle|\, e_t\ge m\right)\mathbf{Pr}(e_t\ge m)
\end{aligned}$$

$$+\alpha_1 1_{k=m}\left(1-O\left(\mathbf{E}\left(\frac{\Delta_t}{e_t}\;\middle|\; e_t>0\right)\right)\right)\mathbf{Pr}(e_t>0)$$
$$+1_{k<m}O\left(\mathbf{Pr}(e_t<m)+\mathbf{E}\left(\frac{\Delta_t}{e_t}\;\middle|\; e_t>0\right)\right)+O(\mathbf{Pr}(e_t=0)). \tag{5.1}$$

A term such as $\mathbf{E}\left(k\frac{D_k(t)}{e_t}\;\middle|\; e_t>0\right)$ is expressed as

$$\begin{aligned}
&\mathbf{E}\left(\frac{kD_k(t)}{e_t}\;\middle|\; e_t>0\right)\\
&\quad=\mathbf{E}\left(\frac{kD_k(t)}{e_t}\;\middle|\; |e_t-\eta t|\le t^{1-\rho/8}\right)\mathbf{Pr}(|e_t-\eta t|\le t^{1-\rho/8}\mid e_t>0)\\
&\qquad+\mathbf{E}\left(\frac{kD_k(t)}{e_t}\;\middle|\; |e_t-\eta t|> t^{1-\rho/8}, e_t>0\right)\mathbf{Pr}(|e_t-\eta t|> t^{1-\rho/8}\mid e_t>0)\\
&\quad=\frac{\mathbf{E}(kD_k(t)\mid |e_t-\eta t|\le t^{1-\rho/8})\mathbf{Pr}(|e_t-\eta t|\le t^{1-\rho/8}\mid e_t>0)}{\eta t}(1+O(t^{-\rho/8}))\\
&\qquad+O(\mathbf{Pr}(|e_t-\eta t|> t^{1-\rho/8}\mid e_t>0)),
\end{aligned}$$

where we used the fact that $kD_k(t)/e_t\le 2$ to handle the second term.

For $k\ge 1$, we have $\overline{D}_k(t)=\mathbf{E}(D_k(t)\mid e_t>0)\mathbf{Pr}(e_t>0)$ and so

$$\begin{aligned}
\mathbf{E}(kD_k(t)\mid |e_t-\eta t|&\le t^{1-\rho/8})\mathbf{Pr}(|e_t-\eta t|\le t^{1-\rho/8}\mid e_t>0)\\
&= k\overline{D}_k(t)-\mathbf{E}(kD_k(t)\mid |e_t-\eta t|> t^{1-\rho/8})\\
&\quad\times\mathbf{Pr}(|e_t-\eta t|> t^{1-\rho/8}\mid e_t>0)\\
&= k\overline{D}_k(t)+O(t^{1-\rho/4}\log^5 t).
\end{aligned}$$

Thus we find, for $k\ge 0$,

$$\mathbf{E}\left(\frac{kD_k(t)}{e_t}\;\middle|\; e_t>0\right)=\frac{k\overline{D}_k(t)}{\eta t}+O(t^{-\rho/8}).$$

Similarly,

$$\mathbf{E}\left(\frac{D_k(t)}{v_t}\;\middle|\; v_t>0\right)=\frac{\overline{D}_k(t)}{\nu t}+O(t^{-1/2}\log t).$$

Now consider the term

$$\begin{aligned}
&\mathbf{E}\left(\frac{\Delta_t}{e_t}\;\middle|\; e_t>0\right)=\\
&\mathbf{E}\left(\frac{\Delta_t}{e_t}\;\middle|\; |e_t-\eta t|\le t^{1-\rho/8}\right)\mathbf{Pr}(|e_t-\eta t|\le t^{1-\rho/8},\Delta_t\le t^{1-\rho/2}(\log t)^2\mid e_t>0)
\end{aligned}$$

$$
\begin{aligned}
& +\mathbf{E}\left(\frac{\Delta_t}{e_t} \,\middle|\, (|e_t-\eta t|>t^{1-\rho/8} \text{ or } \Delta_t>t^{1-\rho/2}(\log t)^2), e_t>0\right) \\
& \times \mathbf{Pr}(|e_t-\eta t|>t^{1-\rho/8} \text{ or } \Delta_t \geq t^{1-\rho/2}(\log t)^2 \mid e_t>0) \\
= \; & O(t^{-\rho/4}(\log t)^{\xi+5})
\end{aligned}
$$

since $\Delta_t \leq e_t$.

Going back to (5.1), we get

$$
\begin{aligned}
\overline{D}_k(t+1) \;=\; & \overline{D}_k(t)+(2\alpha-\alpha_1)m\left(-\frac{k\overline{D}_k(t)}{2\eta t}+\frac{(k-1)\overline{D}_{k-1}(t)}{2\eta t}\right) \\
& +(1-\alpha-\alpha_0)(k+1)\left(\frac{\overline{D}_{k+1}(t)}{\nu t}-\frac{\overline{D}_k(t)}{\nu t}\right) \\
& +m\alpha_0\left(\frac{(k+1)\overline{D}_{k+1}(t)}{\eta t}-\frac{k\overline{D}_k(t)}{\eta t}\right)+\alpha_1 1_{k=m}+O(t^{-\rho/8}) \\
=\; & \overline{D}_k(t)+\left(\left(\frac{1-\alpha-\alpha_0}{\nu}+\frac{m\alpha_0}{\eta}\right)(k+1)\right)\frac{\overline{D}_{k+1}(t)}{t} \\
& -\left(\left(\frac{(2\alpha-\alpha_1+2\alpha_0)m}{2\eta}\right)+\frac{1-\alpha-\alpha_0}{\nu}\right)k \\
& +\left(\frac{1-\alpha-\alpha_0}{\nu}\right)\frac{\overline{D}_k(t)}{t}+\frac{(2\alpha-\alpha_1)m}{2\eta}(k-1)\frac{\overline{D}_{k-1}(t)}{t} \\
& +\alpha_1 1_{k=m}+O(t^{-\rho/8}). \qquad (5.2)
\end{aligned}
$$

To motivate the recurrence (5.3) below, we (heuristically) put $\bar{d}_k=\frac{\overline{D}_k(t)}{t}$, assume it is a constant, and get

$$
\begin{aligned}
\bar{d}_k \;=\; & \left(\frac{1-\alpha-\alpha_0}{\nu}+\frac{\alpha_0}{\eta}\right)(k+1)\bar{d}_{k+1} \\
& -\left(\left(\frac{(2\alpha-\alpha_1+2\alpha_0)m}{2\eta}+\frac{1-\alpha-\alpha_0}{\nu}\right)k+\left(\frac{1-\alpha-\alpha_0}{\nu}\right)\right)\bar{d}_k \\
& +\frac{(2\alpha-\alpha_1)m}{2\eta}(k-1)\bar{d}_{k-1}+\alpha_1 1_{k=m}+O(t^{-\rho/8}).
\end{aligned}
$$

This leads to the consideration of the recurrence: $d_{-1}=0$ and for $k \geq -1$,

$$
(A_2(k+2)+B_2)d_{k+2}+(A_1(k+1)+B_1)d_{k+1}+(A_0k+B_0)d_k=-\alpha_1 1_{k=m-1}, \quad (5.3)
$$

where

$$A_2 = \frac{1-\alpha-\alpha_0}{\nu} + \frac{m\alpha_0}{\eta} \qquad B_2 = 0$$
$$A_1 = -\frac{(2\alpha-\alpha_1+2\alpha_0)m}{2\eta} - \frac{1-\alpha-\alpha_0}{\nu} \qquad B_1 = -1-\frac{1-\alpha-\alpha_0}{\nu} .$$
$$A_0 = \frac{(2\alpha-\alpha_1)m}{2\eta} \qquad B_0 = 0$$

Having motivated the recurrence (5.3) we will now show rigorously that it is indeed a good approximation for $\frac{\overline{D}_k(t)}{t}$. We first note that Lemma 3.1 implies

$$0 \leq \overline{D}_k(t) \leq t^{-10} \text{ for } k \geq k_0(t). \tag{5.4}$$

Lemma 5.1. *Let d_k be a solution to (5.3) such that $|d_k| \leq \frac{C}{k}$ for $k > 0$ and for some C. Then there exists a constant $M > 0$ such that for $t \geq 2,\ k \geq -1$,*

$$|\overline{D}_k(t) - td_k| \leq Mt^{1-\rho/8}. \tag{5.5}$$

Proof. Equation (5.4) and $|d_k| \leq C/k$ imply that (5.5) holds for $k \geq k_0$ when t is sufficiently large. So assume now that $k \leq k_0$. Let $\Theta_k(t) = \overline{D}_k(t) - td_k$. Note that by definition $\Theta_{-1}(t) = 0$ and that for $k \geq k_0$, $\Theta_k(t) = O(t^{1-\rho/2}(\log t)^5)$. Assume $0 \leq k < k_0$. It follows from (5.2) and (5.3) that

$$\begin{aligned}\Theta_k(t+1) &= \Theta_k(t) + A_2(k+1)\frac{\Theta_{k+1}(t)}{t} + (A_1k+B_1+1)\frac{\Theta_k(t)}{t}\\ &\quad + A_0(k-1)\frac{\Theta_{k-1}(t)}{t} + O(t^{-\rho/8}).\end{aligned} \tag{5.6}$$

When t is large, $k \leq k_0$ implies that $t + A_1k + B_1 + 1 \geq 0$. We can adjust M to deal with small t and so assume this is true. Let L denote the hidden constant in $O(t^{-\rho/8})$ of (5.2). Our inductive hypothesis $\mathcal{H}_t$ is that $\Theta_k(t) \leq Mt^{1-\rho/8}$ for every $k \geq 0$. It is trivially true for small t. Since $k \leq k_0(t+1)$, (5.6) implies that

$$\begin{aligned}|\Theta_k(t+1)| &\leq A_2(k+1)\frac{|\Theta_{k+1}(t)|}{t} + (t+A_1k+B_1+1)\,\frac{|\Theta_k(t)|}{t}\\ &\quad + A_0(k-1)\frac{|\Theta_{k-1}(t)|}{t} + Lt^{-\rho/8}\\ &\leq (t + A_2(k+1) + A_1k + B_1 + 1 + A_0(k-1))\,Mt^{-\rho/8} + Lt^{-\rho/8}\\ &= (t + A_2 + B_1 + 1 - A_0)\,Mt^{-\rho/8} + Lt^{-\rho/8}\\ &\leq Mt^{1-\rho/8} + Lt^{-\rho/8} \qquad \text{assuming that } \alpha_0 \leq \alpha - \alpha_1/2\\ &\leq M(t+1)^{1-\rho/8}\end{aligned}$$

provided $M \geq 2L$. This verifies $\mathcal{H}_{t+1}$ and completes the proof by induction. □

Notice that Lemma 5.1 implies that if there is a solution for (5.3) such that $|d_k| \leq \frac{C}{k}$ for some C, then

$$\lim_{t\to\infty} \frac{\overline{D}_k(t)}{t} = d_k.$$

Remark 5.2. The theorem implies that if there is a solution for (5.3) such that $d_k \leq \frac{C}{k}$, then $\lim_{t\to\infty} \frac{\overline{D}_k(t)}{t}$ exists and is equal to d_k. In particular, this shows that if a solution to (5.3) exists such that $d_k \leq \frac{C}{k}$, then it must be a unique such solution to (5.3).

6. Solving (5.3)

To solve (5.3) we first solve the homogeneous equation

$$(A_2(k+2)+B_2)f_{k+2}+(A_1(k+1)+B_1)f_{k+1}+(A_0k+B_0)f_k = 0, \qquad k \geq 1 \tag{6.1}$$

using Laplace's method as explained in [Jordan 39]. (Notice that $k \geq 1$ as opposed to $k \geq -1$ in (5.3). This will be useful in dealing with the inhomogeneity.)

For $k \geq 1$, we make the substitution

$$f_k = \int_{t=a}^{t=b} t^{k-1} v(t)dt \tag{6.2}$$

for $a, b, v(t)$ to be determined.

Integrating by parts

$$kf_k = [t^k v(t)]_a^b - \int_{t=a}^{t=b} t^k v'(t)dt. \tag{6.3}$$

Let

$$\begin{aligned} \phi_1(t) &= A_2t^2 + A_1t + A_0 \\ \phi_0(t) &= B_2t^2 + B_1t + B_0. \end{aligned}$$

Substituting (6.2) and (6.3) into (6.1), we obtain

$$[t^k\phi_1(t)v(t)]_a^b - \int_a^b t^k\phi_1(t)v'(t)dt + \int_a^b t^{k-1}\phi_0(t)v(t)dt = 0. \tag{6.4}$$

Equation (6.1) will be satisfied if we ensure that

$$\frac{v'(t)}{v(t)} = \frac{\phi_0(t)}{t\phi_1(t)}, \tag{6.5}$$

and

$$[t^k v(t)\phi_1(t)]_a^b = 0. \tag{6.6}$$

We satisfy (6.6) by taking $a = 0$ and b equal to a root of $v(t)\phi_1(t) = 0$. Going back to (6.1), we have

$$\begin{aligned} \phi_0(t) &= -At \\ \phi_1(t) &= Bt^2 - (B+C)\,t + C \\ &= B(t-1)\,(t - C/B)\,, \end{aligned}$$

where

$$\begin{aligned} A &= 1 + \frac{1-\alpha-\alpha_0}{\nu} = \frac{\alpha_1}{\nu} \\ B &= \frac{1-\alpha-\alpha_0}{\nu} + \frac{m\alpha_0}{\eta} = \frac{\alpha(1-\alpha-\alpha_0)+\alpha_0\alpha_1}{(\alpha-\alpha_0)\nu} \\ C &= \frac{(2\alpha-\alpha_1)m}{2\eta} = B + \frac{\alpha_1(3\alpha-1-\alpha_1-\alpha_0)}{2(\alpha-\alpha_0)\nu} > B. \end{aligned}$$

The differential equation is homogeneous and can be integrated to give

$$v(t) = D_0(t-1)^\beta\,(t - C/B)^{-\beta}\,,$$

where $D_0 = D_0(\alpha, \alpha_0, \alpha_1) \neq 0$ and

$$\beta = \frac{A}{C-B}$$

is given by (1.2).

Since $C > B$ we see that $v(t)$ is continuous in $[0, 1]$ and this allows us to take $b = 1$ and satisfy (6.6).

Substituting into (6.2) and removing the constant multiplicative factor C_1, we define a solution u_1 to (6.1) for $k \geq 1$: let $\gamma = B/C$,

$$u_1(k) = \int_0^1 t^{k-1}\left(\frac{1-t}{1-\gamma t}\right)^\beta dt. \tag{6.7}$$

Notice that u_1 does not satisfy Equation (6.1) when $k = 0$; in particular, using (6.3) and (6.5) we get

$$\begin{aligned} 2A_2u_1(2) + (A_1+B_1)u_1(1) &= 2A_2\int_0^1 tv(t)dt + (A_1+B_1)\int_0^1 v(t)dt \\ &= A_2\left(\left[t^2v(t)\right]_0^1 - \int_0^1 t^2v'(t)dt\right) \\ &\quad + A_1\left([tv(t)]_0^1 - \int_0^1 tv'(t)dt\right) + B_1\int_0^1 v(t)dt \end{aligned}$$

$$
\begin{aligned}
&= \left[(A_2t^2 + A_1t)v(t)\right]_0^1 \\
&\quad - \int_0^1 (A_2t^2 + A_1t)v'(t)dt + B_1 \int_0^1 v(t)dt \\
&= [(\phi_1(t) - A_0)v(t)]_0^1 \\
&\quad - \int_0^1 (\phi_1(t) - A_0)v'(t)dt + B_1 \int_0^1 v(t)dt \\
&= [(\phi_1(t) - A_0)v(t)]_0^1 + \int_0^1 A_0 v'(t)dt \\
&\quad - \int_0^1 (\phi_1(t)v'(t) + B_1 v(t))dt \\
&= [\phi_1(t)v(t)]_0^1 \\
&= -\phi_1(0)v(0) \neq 0. \qquad (6.8)
\end{aligned}
$$

Lemma 6.1. *Let $k \geq 1$. Then*

$$u_1(k) = (1 + O(k^{-1}))C_1 k^{-(1+\beta)}$$

for $C_1 = C_1(\alpha)$ a fixed constant.

Proof.

$$
\begin{aligned}
u_1(k) &= \int_0^1 t^{k-1}(1-t)^\beta \frac{1}{(1-\gamma t)^\beta} dt \\
&= \int_0^1 t^{k-1}(1-t)^\beta \sum_{j=0}^{\infty} \binom{\beta + j - 1}{j} (\gamma t)^j dt \\
&= \sum_{j=0}^{\infty} \binom{\beta + j - 1}{j} \gamma^j \int_0^1 t^{k+j-1}(1-t)^\beta dt \\
&= \sum_{j=0}^{\infty} \binom{\beta + j - 1}{j} \gamma^j \frac{\Gamma(k+j)\Gamma(\beta+1)}{\Gamma(k+j+\beta+1)} \\
&= \sum_{j=0}^{\infty} \gamma^j \frac{\Gamma(\beta+j)}{\Gamma(j+1)\Gamma(\beta)} \frac{\Gamma(k+j)\Gamma(\beta+1)}{\Gamma(k+j+\beta+1)}
\end{aligned}
$$

assuming k is large, using Stirling for $\Gamma(k+j), \Gamma(k+j+\beta+1)$, we get

$$
\begin{aligned}
&= (1 + O(k^{-1}))\beta \sum_{j=0}^{\infty} \gamma^j \frac{\Gamma(j+\beta)}{\Gamma(j+1)} (k+\beta+j)^{-\beta-1} \\
&= (1 + O(k^{-1}))C_1 k^{-1-\beta}.
\end{aligned}
$$

□

Up until now, we have only found one solution of (6.1). The set of solutions to (6.1) is two-dimensional. One can show that any solution, independent of u_1, must have $|f_k| \to \infty$ as $k \to \infty$ and so u_1 is the only relevant solution. We do not actually need to prove this assertion because Remark 5.2 implies that our choice of solution d_k to the recurrence (5.3) well approximates $\overline{D}_k(t)$ as long as $d_k = O(k^{-1})$, which is what we will prove now.

Now we are going to prove that we can choose d_0 such that the solution of (5.3) satisfies the requirements of Lemma 5.1.

Lemma 6.2. *There exist C and a choice of $d_0 = d_0(C)$ such that $d_0 = d$ implies $d_k = Cu_1(k)$ for $k \geq m$.*

Proof. Assume first $m > 1$. Define $w_k = 0$ for $k \geq m$, $w_{m-1} = -\frac{\alpha_1}{(m-1)A_0}$ and for $j = m-2, m-3, \ldots 1$, let w_j be such that

$$A_2(j+2)w_{j+2} + (A_1(j+1) + B_1)w_{j+1} + A_0 j w_j = 0.$$

Then w_k satisfies Equation (5.3) for $k \geq 1$.

Now, let $C = -\frac{2A_2w_2+(A_1+B_1)w_1}{2A_2u_1(2)+(A_1+B_1)u_1(1)}$ and $d = -\frac{A_2(Cu_1(1)+w_1)}{B_1}$. Notice that (6.8) implies that C is well-defined.

Claim 6.3.

$$\hat{d}_k = \begin{cases} 0 & \text{if } k = -1 \\ d & \text{if } k = 0 \\ Cu_1(k) + w_k & \text{otherwise} \end{cases}$$

is a solution to (5.3).

For $k \geq 1$, w_k satisfies Equation (5.3) and $u_1(k)$ satisfies Equation (6.1), therefore $\hat{d}_k = Cu_1(k) + w_k$ satisfies (5.3).

For $k = 0$, we have

$$\begin{aligned} 2A_2\hat{d}_2 + (A_1 + B_1)\hat{d}_1 &= 2A_2(Cu_1(2) + w_2) + (A_1 + B_1)(Cu_1(1) + w_1) \\ &= C(2A_2u_1(2) + (A_1 + B_1)u_1(1)) + 2A_2w_2 + (A_1 + B_1)w_1 \\ &= 0. \end{aligned}$$

For $k = -1$: $A_2\hat{d}_1 + B_1\hat{d}_0 = A_2(Cu_1(1) + w_1) + B_1 d = 0$. This finishes the proof of the claim.

Now, if $m = 1$, take $C = -\frac{\alpha_1}{2A_2u_1(2)+(A_1+B_1)u_1(1)}$, $d = -\frac{CA_2u_1(1)}{B_1}$ and define

$$\hat{d}_k = \begin{cases} 0 & \text{if } k = -1 \\ d & \text{if } k = 0 \\ Cu_1(k) & \text{otherwise.} \end{cases}$$

A similar argument to the one above shows that in this case $\hat{d}_k$ is also a solution to (5.3). Lemma 6.1 implies that $\hat{d}_k = O(k^{-1})$, as required by Lemma 5.1. □

7. The Degree Sequence

Now we can prove our main theorem. We repeat its statement here:

Theorem 7.1. (Theorem 3.1 restated.) *There exists a constant $C = C(m, \alpha)$ such that for $k > 0$,*

$$\left| \frac{\overline{D}_k(t)}{t} - Ck^{-1-\beta} \right| = O\left(t^{-\rho/8}\right) + O\left(k^{-2-\beta}\right).$$

Proof. Notice that $u_1(k) \leq \frac{A}{k}$ for some A. By Lemma 6.2 there are C' and d such that $d_0 = d$ implies for every $k > m$, $d_k = C'u_1(k) \leq \frac{AC'}{k}$. By Lemma 5.1,

$$\left| \frac{\overline{D}_k(t)}{t} - C'u_1(k) \right| = O\left(t^{-\rho/8}\right).$$

But from Lemma 6.1,

$$u_1(k) = C_1 k^{-1-\beta} + O\left(k^{-2-\beta}\right),$$

and therefore

$$\left| \frac{\overline{D}_k(t)}{t} - C'C_1 k^{-1-\beta} \right| = O\left(t^{-\rho/8}\right) + O\left(k^{-2-\beta}\right).$$

□

Acknowledgments. We thank one of the referees for pointing out an error in our proof of Lemma 4.1. The second author was supported in part by NSF grant CCR-0200945. The third author was supported in part by NSF grant CCR-0200945.

References

[Albert et al. 99] R. Albert, A. Barabási, and H. Jeong. "Diameter of the World Wide Web." *Nature* 401 (1999), 103–131.

[Aiello et al. 00] W. Aiello, F. R. K. Chung, and L. Lu. "A Random Graph Model for Massive Graphs." In *Proc. of the 32nd Annual ACM Symposium on the Theory of Computing*, pp. 171–180. New York: ACM Press, 2000.

[Aiello et al. 02] W. Aiello, F. R. K. Chung, and L. Lu. "Random Evolution in Massive Graphs." In *Handbook on Massive Data Sets*, edited by James Abello et al., pp. 510–519. Norwood, MA: Kluwer Academic Publishers, 2002.

[Barabási and Albert 99] A. Barabási and R. Albert. "Emergence of Scaling in Random Networks." *Science* 286 (1999), 509–512.

[Broder et al. 00] A. Broder, R. Kumar, F. Maghoul, P. Raghavan, S. Rajagopalan, R. Stata, A. Tomkins, and J. Wiener. "Graph Structure in the Web." In *Proceedings of the 9th International World Wide Web Conference on Computer Networks*, pp. 309–320. Amsterdam: North-Holland Publishing Co., 2000.

[Bollobás and Riordan 02] B. Bollobás and O. Riordan. "Mathematical Results on Scale-Free Random Graphs." In *Handbook of Graphs and Networks*, pp. 1–34. Berlin: Wiley-VCH, 2002.

[Bollobás and Riordan 04a] B. Bollobás and O. Riordan. "The Diameter of a Scale-Free Random Graph." *Combinatorica* 24:1 (2004), 5–34.

[Bollobás and Riordan 04b] B. Bollobás and O. Riordan. "Coupling Scale-Free and Classical Random Graphs." *Internet Mathematics* 1:2 (2004), 215–225.

[Bollobás et al. 01] B. Bollobás, O. Riordan, J. Spencer, and G. Tusanády. "The Degree Sequence of a Scale-Free Random Graph Process." *Random Structures and Algorithms* 18 (2001), 279–290.

[Cooper and Frieze 03] C. Cooper and A. M. Frieze. "A General Model of Undirected Web Graphs." *Random Structures and Algorithms* 22 (2003), 311-335.

[Chung and Lu 04] F. R. K. Chung and L. Lu. "Coupling On-Line and Off-Line Analyses for Random Power Law Graphs." *Internet Mathematics* 1:4 (2004), 409–461.

[Chung et al. 03] F. R. K. Chung, L. Lu, and V. Vu. "Eigenvalues of Power Law Graphs." *Annals of Combinatorics* 7 (2003), 21–33.

[Chung et al. 02] F. R. K. Chung, L. Lu, and V. Vu. "The Spectra of Random Graphs with Given Expected Degrees." *Proceedings of the National Academy of Sciences* 100 (2003), 6313–6318.

[Erdös and Rényi 59] P. Erdös and A. Rényi. "On Random Graphs I." *Publicationes Mathematicae Debrecen* 6 (1959), 290–297.

[Faloutsos et al. 99] M. Faloutsos, P. Faloutsos, and C. Faloutsos. "On Power-Law Relationships of the Internet Topology." In *Proceedings of the Conference on Applications, Technologies, Architectures, and Protocols for Computer Communication*, pp. 251–262. New York: ACM Press, 1999.

[Hayes 00] B. Hayes. "Graph Theory in Practice: Part II." *American Scientist* 88 (2000), 104–109.

[Jordan 39] C. Jordan. *Calculus of Finite Differences.* Sopron, Hungary: Rottig and Romwalter, 1939.

[Mitzenmacher 04] M. Mitzenmacher. "A Brief History of Generative Models for Power Law and Lognormal Distributions." *Internet Mathematics* 1:2 (2004), 226–251.

[Mihail and Papadimitriou 02] M. Mihail and C. H. Papadimitriou. "On the Eigenvalue Power Law." In *Proceedings of 6th International Workshop on Randomization and Approximation Techniques*, pp. 254–262. New York: Springer-Verlag, 2002.

[Simon 55] H. A. Simon. "On a Class of Skew Distribution Functions." *Biometrika* 42:3/4 (1955), 425–440.

[Watts 99] D. J. Watts. *Small Worlds: The Dynamics of Networks Between Order and Randomness.* Princeton, NJ: Princeton University Press, 1999.

[Yule 25] G. Yule. "A Mathematical Theory of Evolution Based on the Conclusions of Dr. J.C. Willis." *Philosophical Transactions of the Royal Society of London (Series B)* 213 (1925), 21–87.

Colin Cooper, Department of Computer Science, King's College, University of London, London WC2R 2LS, UK (ccooper@dcs.kcl.ac.uk)

Alan Frieze, Department of Mathematical Sciences, Carnegie Mellon University, Pittsburgh, PA 15213 (alan@random.math.cmu.edu)

Juan Vera, Department of Mathematical Sciences, Carnegie Mellon University, Pittsburgh, PA, 15213 (jvera@andrew.cmu.edu)

Received October 2, 2003; accepted January 6, 2004.