A general model of web graphs

Colin Cooper* Alan Frieze†

July 31, 2003

Abstract

We describe a very general model of a random graph process whose proportional degree sequence obeys a power law. Such laws have recently been observed in graphs associated with the world wide web.

1 Introduction

At the present moment there is considerable research into the structure of large-scale real networks, and in modeling these networks as the outcomes of discrete random processes. A general introduction to this topic can be found in Hayes [16] or Watts [26]. In particular, there is a strong interest in the structure of the Internet and World Wide Web (**www**). Experimental studies by, Albert, Barabasi and Jeong [1], Broder et al [7] and Faloutsos, Faloutsos and Faloutsos [15] of the structure of the **www** have demonstrated an inverse power law for the proportion of vertices with a given degree.

To model such structures, we require a graph process which (a) evolves randomly by the addition of new vertices and/or edges at each time step t and (b) whose expected proportional degree sequence follows a power law. Such random graph process are referred to as web graphs, or scale-free graphs. These processes differ from the more traditional models of random graphs introduced by Erdős and Rényi [13], [14] where the number of vertices remains fixed, and the proportion of vertices of a given degree is Poisson distributed, and hence the degree sequence drops off exponentially in the upper tail.

One method of producing graph processes with a power law degree sequence is to introduce an element of preferential attachment (or copying) into the way that a new vertex attaches its edges to the existing graph. There is a long history of such models, outlined in the survey by Mitzenmacher [23]. We will use the preferential attachment model to generate our random graph. The preferential attachment random graph has been the subject of recently revived interest. It dates back to Yule [27], and Simon [25]. It was proposed as a model for the web by Barabási and Albert [2], and their description was elaborated by Bollobas, Riordan, Spencer, and Tusnády [3] who proved that the degree sequence does follow a power law distribution. Bollobas and Riordan [4] obtained several additional results regarding the diameter and connectivity of such graphs.

If we may briefly summarize models which generalize the work of [2], then they are of the following form. A new vertex is added at each step, which directs a fixed number of edges to the existing

*Department of Computer Science, King's College, University of London, London WC2R 2LS, UK. E-mail: ccooper@dcs.kcl.ac.uk. Research supported in part by the STORM Research Centre, UNL.

†Department of Mathematical Sciences, Carnegie Mellon University, Pittsburgh PA15213, U.S.A., Supported in part by NSF grant CCR-9818411. E-mail: alan@random.math.cmu.edu.

graph. The terminal vertices of these edges are chosen either by copying or by a mixture of copying and uniform (uar) selection. There are two main methods for choosing a terminal vertex by copying:
(a) Select a vertex directly with probability proportional to degree,
(b) Represent each vertex degree by points in a configuration model and select uniformly over points.
Details of the configuration model approach can be found in [3], [5]. The processes associated with these two methods are called respectively, web graphs and scale-free graphs. Processes studied by Cooper and Frieze [9], Kumar et al [20], [21], are web graphs. Those studied by Bollobás and Riordan [3], [4], Buckley and Osthus [8], Dogorovtsev, Mendes and Samukhin [11], and Drinea, Enachescu and Mitzenmacher [12] are scale-free graphs. See the recent survey by Bollobás and Riordan [5] for further details.

The broad brush structure of these models can be summarized by their power law parameter x; that is, the proportion of vertices of degree k is asymptotically proportional to k^{-x}. Here x is a fixed constant whose precise value is dependent on the precise mixture of copying and uniform selection when choosing a terminal vertex. The sampling method (a) or (b), makes no difference to the parameter x once the exact mixture of copying and uniform selection has been decided.

The aim of this paper is to obtain the degree distribution of a general web graph process which allows as much choice as possible at each step. In particular we allow the insertion of additional edges between existing vertices. Without this development, it seems that the power law parameter x must always be at least 3, whereas if edges can be inserted between existing vertices, the parameter x can take any value greater than 2. We were also interested to see, if by varying the details of the model sufficiently, we could move away from a power law result for the degree sequence. However, as we show, this is not the case, as long as some element of copying (however small) is permitted.

Our results are presented in terms of $D_k(t)$, the number of vertices of degree k. We prove the following: For all $k \geq 1$ and all $t \geq 1$ the random variable $D_k(t)$ is concentrated about the expected value $\overline{D}_k(t)$. As $t \to \infty$ this expected value is itself well approximated by td_k, where d_k is the solution of a certain difference equation. We show that, with some minor exceptions, the solution of this difference equation is of the form $d_k \sim Ck^{-x}$, as $k \to \infty$, for an explicitly determined value x.

We refer to our model as *undirected*, although, in fact, directed variants are easily obtained. We consider the general undirected web model to be intrinsically interesting, aside from applications to the **www**. Moreover, although the edges of a typical **www** graph are directed, the idea of an undirected model has many attractions, not least its simplicity. One particular example which supports the use of an undirected model is as follows.

The Google search engine [6] holds a partial model of the **www** which it is continuously updating. Once a node is added to the search engine database, a list is maintained of pages in the database with forward links to this node. For a given node with url ***node-url***, these links can be found by entering the ***link: node-url*** query to Google. Thus the model of the **www** held by this search engine is equivalent to an undirected web graph in the sense that one can find all links pointing to a node.

Results for directed web graphs corresponding to the *Hub-Authority* model of the **www**, as proposed by Kleinberg [19] are given in [10]. They are similar to the results given here.

We now pass to a more detailed description of our general web graph model.

2 Undirected web graph model: Definitions and results

We describe the evolution of a random multi-graph $G(t)$ which is an example of the type of model referred to as a web graph. Our model is more general than previous models, but its degree sequence is still amenable to analysis (in most cases) and we find the seemingly ubiquitous power law. By specialising the parameters we obtain models equivalent to many of the previously defined models. In the model of [2], as clarified by various authors, a new vertex is added at time t and this vertex chooses m random neighbours, with probability proportional to their current degree. We generalise this in the following ways: we allow (a) new edges to be inserted between existing vertices, (b) a variable number of edges to be added at each step and (c) endpoint vertices are chosen by a mixture of uniform selection and copying. This results in a large number of parameters, which we will describe below. We first give a precise description of the process.

Initially, at step $t = 0$, there is a single vertex v_0. At any step $t = 1, 2, ..., T, ...$ there is a birth process in which either new vertices or new edges are added. Specifically, either a procedure NEW is followed with probability $1 - \alpha$, or a procedure OLD is followed with probability α. In procedure NEW, a new vertex v is added to $G(t-1)$ with one or more edges added between v and $G(t-1)$. In procedure OLD, an existing vertex v is selected and extra edges are added at v.

The recipe for adding edges at step t typically permits the choice of initial vertex v (in the case of OLD) and of terminal vertices (in both cases) to be made from $G(t-1)$ either u.a.r or according to vertex degree, or a mixture of these two based on further sampling. The number of edges added to vertex v at step t by the procedures (NEW, OLD) is given by distributions specific to the procedure.

At this point a question arises about our model: Should we regard the edges as directed or undirected in relation to the sampling procedures used in NEW, OLD? We note that the edges have an intrinsic direction, arising from the way they are inserted, which we can ignore or not as we please. We consider the following specific models:

(i) Undirected model: Sampling procedure based on vertex degree.

(ii) Directed out-model: Sampling procedure for out-edges based on out-degree.

(iii) Directed in-model: Sampling procedure for in-edges based on in-degree.

The process allows multiple directed edges, and self-loops can arise from the OLD procedure. The NEW procedure, as described, does not generate self-loops, although this could easily be modified.

We prove for these models, that provided some copying occurs the proportion of vertices of degree k is **whp** asymptotic (for large k) to Ck^{-x}, where $x > 2$ is an explicit function of the parameters of the model. See (3, 40, 41) for the precise functional form of x. We devote most of the paper to the analysis of the undirected model. The other models can easily be formulated as variants of the undirected case, and are covered briefly in Section 6.

2.1 The parameters of the undirected model

Our undirected model $G(t)$ has sampling parameters $\alpha, \beta, \gamma, \delta, \boldsymbol{p}, \boldsymbol{q}$ whose meaning is given below:

Choice of procedure at step t.

α: Probability that an OLD node generates edges.

$1 - \alpha$: Probability that a NEW node is created.

Procedure NEW

$\boldsymbol{p} = (p_i : i \geq 1)$: Probability that the new node generates i new edges.

β: Probability that choices of terminal vertices are made uniformly.
$1-\beta$: Probability that choices of terminal vertices are made according to degree.

Procedure OLD

$\boldsymbol{q} = (q_i : i \geq 1)$: Probability that the old node generates i new edges.
δ: Probability that the initial node is selected uniformly.
$1-\delta$: Probability that the initial node is selected according to degree.
γ: Probability that choices of terminal vertices are made uniformly.
$1-\gamma$: Probability that choices of terminal vertices are made according to degree.

The models we study here require $\alpha < 1$ and $p_0 = q_0 = 0$. It is convenient to assume a *finiteness* condition for the distributions $\{p_j\},\ \{q_j\}$. This means that there exist j_0, j_1 such that $p_j = 0,\ j > j_0$ and $q_j = 0,\ j > j_1$ Imposing the finiteness condition helps simplify the difference equations used in the analysis.

The model creates edges in the following way: An initial vertex v is selected. If the terminal vertex w is chosen u.a.r, we say v is *assigned uniformly* to w. If the terminal vertex w is chosen according to its vertex degree, we say v is *copied* to w. In either case the edge has an intrinsic direction (v, w), which we may choose to ignore. We note that sampling according to vertex degree is equivalent to selecting an edge u.a.r and then selecting an endpoint u.a.r.

At this point in the discussion, it is appropriate to say a few words about alternative definitions of copying. The papers [20], [21] introduce a copying model in which a new vertex v chooses an old vertex w and selects (copies) a randomly chosen set of out-neighbours of w to be its own out-neighbours. The occurrence of a large number of small complete bipartite subgraphs is a feature found in trawls of the web. The above copying method leads to a larger number of small complete bipartite subgraphs than would be obtained by other models.

Our focus is on degree sequence and, as we now show, the construction of [20], [21] does not lead to fundamentally different results on degree sequence. The precise definition of copying in [20] now follows. A new vertex u of out-degree i is added at each step. The choice of out-edges of u is made as follows. Firstly i provisional vertices are selected u.a.r. Now, independently for each of these i provisional vertices the following choice is made. With probability β vertex v is retained and the edge (u, v) inserted. Or, with probability $1-\beta$ a copied edge (u, w) is inserted instead, where w is the terminal vertex of a uniformly selected out-edge of v. This process of copying is equivalent in terms of expected degree sequence, to the version of copying we propose above for (the in-directed variant of) our model, namely selecting the terminal vertex of a random edge. For, in our model

$$\mathbf{Pr}(w \text{ is the terminal vertex of a u.a.r edge}) = \frac{d^-(w)}{|E|},$$

and in the copying model

$$\mathbf{Pr}(w \text{ is selected by copying an edge }) = \frac{d^-(w)}{|V|}\frac{1}{i},$$

where $|E| = i|V|$.

For directed copying models (in the sense of [20], [21]), it is a result of those papers that the expected proportion of vertices of in-degree k is asymptotic to $Ck^{-\frac{2-\beta}{1-\beta}}$. The equivalent in-directed variant of our model gives the exactly the same result for the expected proportion of vertices of in-degree k. The parameter $(2-\beta)/(1-\beta)$ in our model is obtained from x^- in (41), by putting $\alpha = 0$ and $\mu_p = i$, as a new vertex with exactly i new edges is added at each step.

We remark that the copying model of [20] does not generalize naturally to the case where the out-degree is not a constant value i. For

$$\mathbf{Pr}(w \text{ is selected on copying from } u) = \frac{1}{|V|} \sum_{v \in N^-(w)} \frac{1}{d^+(v)}$$

which seems a difficult quantity to deal with, especially if there is correlation between the in-degree of w and the out-degree of v. However, our approach of selecting the terminal vertex of a random edge remains an easily accessible sampling procedure of an equivalent nature.

Notation
Let $\mu_p = \sum_{j=1}^{j_0} jp_j$, $\mu_q = \sum_{j=1}^{j_1} jq_j$ and let $\theta = 2((1-\alpha)\mu_p + \alpha\mu_q)$. To simplify subsequent notation, we introduce new parameters as follows:

$$\begin{aligned}
a &= 1 + \beta\mu_p + \frac{\alpha\gamma\mu_q}{1-\alpha} + \frac{\alpha\delta}{1-\alpha}, \\
b &= \frac{(1-\alpha)(1-\beta)\mu_p}{\theta} + \frac{\alpha(1-\gamma)\mu_q}{\theta} + \frac{\alpha(1-\delta)}{\theta}, \\
c &= \beta\mu_p + \frac{\alpha\gamma\mu_q}{1-\alpha}, \\
d &= \frac{(1-\alpha)(1-\beta)\mu_p}{\theta} + \frac{\alpha(1-\gamma)\mu_q}{\theta}, \\
e &= \frac{\alpha\delta}{1-\alpha}, \\
f &= \frac{\alpha(1-\delta)}{\theta}.
\end{aligned}$$

We note that

$$c + e = a - 1 \text{ and } b = d + f. \tag{1}$$

Now define the sequence $(d_0, d_1, ..., d_k, ...)$ by $d_0 = 0$, and for $k \geq 1$

$$d_k(a + bk) = (1-\alpha)p_k + (c + d(k-1))d_{k-1} + \sum_{j=1}^{k-1}(e + f(k-j))q_j d_{k-j}. \tag{2}$$

For convenience we define $d_k = 0$ for $k < 0$. Since $a \geq 1$, this system of equations has a unique solution.

Statement of results
The main quantity we study is the random variable $D_k(t)$, the number of vertices of degree k at step t. We let $\overline{D}_k(t) = \mathbf{E}(D_k(t))$. We prove that, as $t \to \infty$, for small k, $\overline{D}_k(t) \approx d_k t$.

Theorem 1. *There exists a constant $M > 0$ such that almost surely for all t, $k \geq 1$*

$$|\overline{D}_k(t) - td_k| \leq Mt^{1/2} \log t.$$

We show in (5), that the number of vertices $\nu(t)$ at step t is **whp** asymptotic to $(1-\alpha)t$. It follows that the proportion of vertices of degree k is **whp** asymptotic to

$$\bar{d}_k = \frac{d_k}{1-\alpha}.$$

The next theorem summarizes what we know about the sequence (d_k) defined by (2).

Theorem 2. *There exist constants $C_1, C_2, C_3, C_4 > 0$ such that*

(i) *$C_1 k^{-\zeta} \le d_k \le C_2 \min\{k^{-1}, k^{-\zeta/j_1}\}$ where $\zeta = (1+d+f\mu_q)/(d+f)$.*

(ii) *If $j_1 = 1$ then $d_k \sim C_3 k^{-(1+1/(d+f))}$.*

(iii) *If $f = 0$ then $d_k \sim C_4 k^{-(1+1/d)}$.*

(iv) *If the* SOLUTION CONDITIONS *given below hold then*

$$d_k = C\left(1 + O\left(\frac{1}{k}\right)\right) k^{-x},$$

where C is constant and

$$x = 1 + \frac{1}{d + f\mu_q}. \tag{3}$$

We say that $\{q_j : j = 1, ..., j_1\}$ is *periodic* if there exists $m > 1$ such that $q_j = 0$ unless $j \in \{m, 2m, 3m, \ldots\}$.

Let

$$\phi_1(y) = y^{j_1} - \left(\frac{d + q_1 f}{b} y^{j_1 - 1} + \frac{q_2 f}{b} y^{j_1 - 2} + \cdots + \frac{q_{j_1} f}{b}\right).$$

Our SOLUTION CONDITIONS are:

S(i) $f > 0$ and either (a) $d + q_1 f > 0$ or (b) $\{q_j\}$ is not periodic.

S(ii) The polynomial $\phi_1(y)$ has no repeated roots. □

We do not suppose the SOLUTION CONDITIONS are necessary. Also, **S(i)** is not very restrictive, as the case $f = 0$ is given by (iii) above.

We also prove some concentration results.

Theorem 3. *For $1 \le k \le k_1 = t^{1/21}$ and for some sufficiently large M,*

$$\mathbf{Pr}(|D_k(t) - \overline{D}_k(t)| \ge M t^{3/4}) \le t^{-\Omega(\log t)}.$$

In a *restricted* process, the decisions as to whether or not to add a new node and how many edges to add at each step are not random, although they may depend on t. Thus the number of vertices and edges in $G(t)$ is deterministic. The precise vertex or vertices chosen by an edge is still random, and thus the structure of the process is not fully determined. The processes considered in [2], [3], [21] and [8] are all of the restricted type.

Theorem 4. *If the process is restricted, then for any $u > 0$,*

$$\mathbf{Pr}(|D_k(t) - \overline{D}_k(t)| \ge u) \le \exp\left\{-\frac{u^2}{8T}\right\}$$

where T is the (deterministic) number of edges in $G(t)$.

Theorem 2 suggests that the maximum degree $\Delta(t)$ of $G(t)$ should satisfy $\Delta(t) = O(t^{1/x})$. This is certainly not the case in some circumstances and indeed may never be the case.

Theorem 5. *If $d + f\mu_q > \frac{1}{2}$ then* **whp**

$$C_4 t^{d+f\mu_q} \le \Delta(t) \le C_5 t^{d+f\mu_q} \tag{4}$$

for some constants $C_4, C_5 > 0$.

3 Evolution of the degree sequence of $G(t)$

Let $\nu(t) = |V(t)|$ be the number of vertices and let $\eta(t) = |2E(t)|$ be the total degree of the graph at the end of step t. $\mathbf{E}\nu(t) = (1-\alpha)t$ and $\mathbf{E}\eta(t) = \theta t$. The random variables $\nu(t)$, $\eta(t)$ are sharply concentrated provided $t \to \infty$. Indeed $\nu(t)$ has binomial distribution $B(t, 1-\alpha)$ and so by the Chernoff bounds,

$$\mathbf{Pr}(|\nu(t) - (1-\alpha)t| \geq t^{1/2}\log t) = O(t^{-K}) \tag{5}$$

for any constant $K > 0$.

Similarly, $\eta(t)$ has expectation θt and is the sum of t independent random variables, each bounded by $\max\{j_0, j_1\}$. Hence, by Hoeffding's theorem [17],

$$\mathbf{Pr}(|\eta(t) - \theta t| \geq t^{1/2}\log t) = O(t^{-K}) \tag{6}$$

for any constant $K > 0$.

These results are almost sure in the sense that they hold for all $t \geq t_0$ with probability $1-O(t_0^{-K+1})$. Thus we can focus on processes such that this is true.

We remind the reader that $D_k(t)$ is the number of vertices of degree k at step t and that $\overline{D}_k(t)$ is its expectation. Here $\overline{D}_j(t) = 0$ for all $j \leq 0, t \geq 0$, $\overline{D}_1(0) = 1$, $\overline{D}_k(0) = 0,\ k \geq 2$.

Using (5) and (6) we see that

$$\overline{D}_k(t+1) = \overline{D}_k(t) + (1-\alpha)p_k + O(t^{-1/2}\log t) \tag{7}$$

$$+ (1-\alpha)\sum_{j=1}^{j_0} p_j \left(\frac{\beta j \overline{D}_{k-1}(t)}{(1-\alpha)t} - \frac{\beta j \overline{D}_k(t)}{(1-\alpha)t} + (1-\beta)\left(\frac{j(k-1)\overline{D}_{k-1}(t)}{\theta t} - \frac{jk\overline{D}_k(t)}{\theta t} \right) \right) \tag{8}$$

$$- \alpha \left(\frac{\delta \overline{D}_k(t)}{(1-\alpha)t} + \frac{(1-\delta)k\overline{D}_k(t)}{\theta t} \right) + \alpha \sum_{j=1}^{j_1} q_j \left(\frac{\delta \overline{D}_{k-j}(t)}{(1-\alpha)t} + \frac{(1-\delta)(k-j)\overline{D}_{k-j}(t)}{\theta t} \right) \tag{9}$$

$$+ \alpha \sum_{j=1}^{j_1} j q_j \left(\gamma \left(\frac{\overline{D}_{k-1}(t)}{(1-\alpha)t} - \frac{\overline{D}_k(t)}{(1-\alpha)t} \right) + (1-\gamma)\left(\frac{(k-1)\overline{D}_{k-1}(t)}{\theta t} - \frac{k\overline{D}_k(t)}{\theta t} \right) \right). \tag{10}$$

Here (8), (9), (10) are (respectively) the main terms of the change in the expected number of vertices of degree k due to the effect on: terminal vertices in NEW, the initial vertex in OLD and the terminal vertices in OLD. Rearranging the right hand side, we find:

$$\begin{aligned}
&\overline{D}_k(t+1) = \overline{D}_k(t) + (1-\alpha)p_k + O(t^{-1/2}\log t) \\
&- \frac{\overline{D}_k(t)}{t}\left(\beta\mu_p + \frac{\alpha\gamma\mu_q}{1-\alpha} + \frac{\alpha\delta}{1-\alpha} + \frac{(1-\alpha)(1-\beta)\mu_p k}{\theta} + \frac{\alpha(1-\gamma)\mu_q k}{\theta} + \frac{\alpha(1-\delta)k}{\theta} \right) \\
&+ \frac{\overline{D}_{k-1(t)}}{t}\left(\beta\mu_p + \frac{\alpha\gamma\mu_q}{1-\alpha} + \frac{(1-\alpha)(1-\beta)\mu_p(k-1)}{\theta} + \frac{\alpha(1-\gamma)\mu_q(k-1)}{\theta} \right) \\
&+ \sum_{j=1}^{j_1} q_j \frac{\overline{D}_{k-j}(t)}{t}\left(\frac{\alpha\delta}{1-\alpha} + \frac{\alpha(1-\delta)(k-j)}{\theta} \right).
\end{aligned}$$

Thus for all $k \geq 1$ and almost surely for all $t \geq 1$,

$$\overline{D}_k(t+1) = \overline{D}_k(t) + (1-\alpha)p_k + O(t^{-1/2}\log t) \qquad (11)$$
$$+ \frac{1}{t}\left((1-(a+bk))\overline{D}_k(t) + (c+d(k-1))\overline{D}_{k-1}(t) + \sum_{j=1}^{j_1} q_j(e+f(k-j))\overline{D}_{k-j}(t)\right).$$

The following Lemma establishes an upper bound on d_k given in Theorem 2(i).

Lemma 1. *The solution of (2) satisfies $d_k \leq \frac{C_2}{k}$.*

Proof We assume that $k > k_0$, and thus $p_k = 0$. Smaller values of k can be dealt with by adjusting C_2. We proceed by induction on k. From (2),

$$\begin{aligned}(a+bk)d_k &\leq (c+d(k-1))\frac{C_2}{k-1} + \sum_{j=1}^{j_1}(e+f(k-j))q_j\frac{C_2}{k-j} \\ &\leq C_2(d+f) + \frac{C_2(c+e)}{k-j_1} \\ &= C_2 b + \frac{C_2(a-1)}{k-j_1},\end{aligned}$$

from (1). So

$$\begin{aligned}d_k - \frac{C_2}{k} &\leq \frac{C_2 b}{a+bk} + \frac{C_2(a-1)}{(k-j_1)(a+bk)} - \frac{C_2}{k} \\ &= \frac{C_2(a-1)}{(k-j_1)(a+bk)} - \frac{C_2 a}{k(a+bk)} \\ &\leq 0.\end{aligned}$$

□

We can now prove Theorem 1, which is restated here for convenience.

Theorem 6. *There exists a constant $M > 0$ such that almost surely for $t,\ k \geq 1$,*

$$|\overline{D}_k(t) - td_k| \leq Mt^{1/2}\log t. \qquad (12)$$

Proof Let $\Delta_k(t) = \overline{D}_k(t) - td_k$. It follows from (2) and (11) that

$$\Delta_k(t+1) = \Delta_k(t)\left(1 - \frac{a+bk-1}{t}\right) + O(t^{-1/2}\log t)$$
$$+ \frac{1}{t}\left((c+d(k-1))\Delta_{k-1}(t) + \sum_{j=1}^{j_1}(e+f(k-j))q_j\Delta_{k-j}(t)\right). \qquad (13)$$

Let L denote the hidden constant in $O(t^{-1/2}\log t)$. We can adjust M to deal with small values of t, so we assume that t is sufficiently large. Let $k_0(t) = \lfloor\frac{t+1-b}{a}\rfloor$. If $k > k_0(t)$ then we observe that (i) $D_k(t) \leq \frac{t\max\{j_0,j_1\}}{k_0(t)} = O(1)$ and (ii) $td_k \leq t\frac{C_2}{k_0(t)} = O(1)$ follows from Lemma 1, and so (12) holds trivially.

Assume inductively that $\Delta_\kappa(\tau) \leq M\tau^{1/2}\log\tau$ for $\kappa + \tau \leq k + t$ and that $k \leq k_0(t)$. Then (13) and $k \leq k_0$ implies that for M large,

$$\begin{aligned} |\Delta_k(t+1)| &\leq L\frac{\log t}{t^{1/2}} + Mt^{1/2}\log t\left(1 + \frac{1}{t}\left(c + dk + \sum_{j=1}^{j_1}(e+fk)q_j - (a+bk-1)\right)\right) \\ &= L\frac{\log t}{t^{1/2}} + Mt^{1/2}\log t \\ &\leq M(t+1)^{1/2}\log(t+1) \end{aligned}$$

provided $M \gg 2L$. We have used (1) to obtain the second line.

This completes the proof by induction. □

3.1 Analysis of the difference equation (2)

Re-writing (2) we see that for $k > j_0$, $p_k = 0$ and then d_k satisfies

$$d_k = d_{k-1}\frac{c + d(k-1)}{a+bk} + \sum_{j=1}^{j_1} d_{k-j}q_j\frac{e + f(k-j)}{a+bk}, \tag{14}$$

which is a linear difference equation with rational coefficients [22].

In the cases where $j_1 = 1$ (an old vertex generates a single edge) or $f = 0$ (old initial vertices are chosen u.a.r) a direct solution to (14) can easily be found, see Sections 3.3 and 3.4.

In general however, when $d > 0$ or $d = 0$ and $\{q_j\}$ is non-periodic, we use classical results on the solution of Laplace's difference equation, (of which (2) is an example) given in [22].

3.2 A general power law bound for d_k

The following lemma completes the proof of Theorem 2(i).

Lemma 2. *For $k > j_0$ we have,*

(i) $d_k > 0$.

(ii)

$$C_1 k^{-(1+d+f\mu_q)/b} \leq d_k \leq C_2 k^{-(1+d+f\mu_q)/bj_1}.$$

Proof

(i) Let κ be the first index such that $p_\kappa > 0$, so that, from (2), $d_\kappa > 0$. It is not possible for both c and d to be zero. Therefore the coefficient of d_{k-1} in (2) is non-zero and thus $d_k > 0$ for $k \geq \kappa$.

(ii) For $k > j_0$ the recurrence (2) satisfies (14), that is

$$d_k = d_{k-1}\frac{c + d(k-1)}{a+bk} + \sum_{j=1}^{j_1} d_{k-j}q_j\frac{e + f(k-j)}{a+bk}.$$

We let $d_i = 0$ for $i < 0$ to handle the cases where $k - j < 0$ in the above sum.

Let $y = 1 + d + f\mu_q$, then

$$\frac{c + d(k-1)}{a + bk} + \sum_{j=1}^{j_1} q_j \frac{e + f(k-j)}{a + bk} = 1 - \frac{y}{a + bk} \geq 0$$

and thus

$$\left(1 - \frac{y}{a + bk}\right) \min\{d_{k-1}, ..., d_{k-j_1}\} \leq d_k \leq \left(1 - \frac{y}{a + bk}\right) \max\{d_{k-1}, ..., d_{k-j_1}\}. \tag{15}$$

It follows that

$$d_{j_0} \prod_{j=j_0+1}^{k} \left(1 - \frac{y}{a + bj}\right) \leq d_k \leq \max\{d_1, d_2, \ldots, d_{j_0}\} \prod_{s=0}^{\lfloor (k-j_0)/j_1 \rfloor} \left(1 - \frac{y}{a + b(k - sj_1)}\right). \tag{16}$$

The LHS of (16) is proved by induction on k. It is trivial for $k = j_0$ and for the inductive step we have

$$\begin{aligned} d_k &\geq d_{j_0} \left(1 - \frac{y}{a + bk}\right) \min_{i=j_0,\ldots,k-1} \left\{ \prod_{j=j_0+1}^{i} \left(1 - \frac{y}{a + bj}\right) \right\} \\ &= d_{j_0} \prod_{j=j_0+1}^{k} \left(1 - \frac{y}{a + bj}\right). \end{aligned}$$

The RHS of (16) is proved as follows: Let $d_{i_1} = \max\{d_{k-1}, ..., d_{k-j_1}\}$, and in general, let $d_{i_{t+1}} = \max\{d_{i_t-1}, ..., d_{i_t-j_1}\}$. Using (15) we see there is a sequence $k - 1 \geq i_1 > i_2 > \cdots i_p > j_0 \geq i_{p+1}$ such that $|i_t - i_{t-1}| \leq j_1$ for all t, and $p \geq \lfloor (k - j_0)/j_1 \rfloor$. Thus

$$d_k \leq d_{i_{p+1}} \prod_{t=0}^{p} \left(1 - \frac{y}{a + bi_t}\right),$$

and the RHS of (16) now follows.

We now consider the product in the LHS of (16).

$$\begin{aligned} \prod_{j=j_0+1}^{k} \left(1 - \frac{y}{a + bj}\right) &= \exp\left\{ \sum_{j=j_0+1}^{k} \left(-\frac{y}{a + bj} - \frac{1}{2}\left(\frac{y}{a + bj}\right)^2 - \cdots \right) \right\} \\ &= \exp\left\{ O(1) - \sum_{j=j_0+1}^{k} \frac{y}{a + bj} \right\} \\ &= C_1 k^{-y/b}. \end{aligned}$$

This establishes the lower bound of the lemma. The upper bound follows similarly, from the upper bound in (16). $\square$

3.3 The case $j_1 = 1$

We prove Theorem 2(ii). When $q_1 = 1$, $p_j = 0, j > j_0 = \Theta(1)$, the general value of d_k, $k > j_0$ can be found directly, by iterating the recurrence (2). Thus

$$\begin{aligned} d_k &= \frac{1}{a+bk}\left(d_{k-1}\left((a-1)+b(k-1)\right)\right) \\ &= d_{k-1}\left(1-\frac{1+b}{a+bk}\right) \\ &= d_{j_0}\prod_{j=j_0+1}^{k}\left(1-\frac{1+b}{a+jb}\right). \end{aligned}$$

Thus, for some constant $C_6 > 0$,

$$d_k \sim C_6(a+bk)^{-x}$$

where

$$x = 1+\frac{1}{b} = 1+\frac{2}{\alpha(1-\delta)+(1-\alpha)(1-\beta)+\alpha(1-\gamma)}.$$

3.4 The case $f = 0$

We prove Theorem 2(iii). The case $(f = 0)$ arises in two ways. Firstly if $\alpha = 0$ so that a new vertex is added at each step. Secondly, if $\alpha \neq 0$ but $\delta = 1$ so that the initial vertex of an OLD choice is sampled u.a.r.

Observe that $b = d$ now, see (1).

We first prove that for a sufficiently large absolute constant $A_2 > 0$ and for all sufficiently large k, that

$$\frac{d_k}{d_{k-1}} = 1-\frac{1+d}{a+dk}+\frac{\xi(k)}{k^2} \tag{17}$$

where $|\xi(k)| \leq A_2$.

We first re-write (2) as

$$\frac{d_k}{d_{k-1}} = \frac{c+d(k-1)}{a+dk}+\sum_{j=1}^{j_1}\frac{eq_j}{a+dk}\prod_{t=k-j+1}^{k-1}\frac{d_{t-1}}{d_t}. \tag{18}$$

(We assume here that $k > j_0$, so that $p_k = 0$.)

Now use induction to write

$$\prod_{t=k-j+1}^{k-1}\frac{d_{t-1}}{d_t} = 1+(j-1)\frac{d+1}{a+dk}+\frac{\xi^*(j,k)}{k^2} \tag{19}$$

where $|\xi^*(j,k)| \leq A_3$ for some constant $A_3 > 0$. (We use the fact that j_1 is constant here.)

Substituting (19) into (18) gives

$$\frac{d_k}{d_{k-1}} = \frac{c+d(k-1)}{a+dk}+\frac{e}{a+dk}+\frac{e(\mu_q-1)(d+1)}{(a+dk)^2}+\frac{\xi^{**}(k)}{(a+dk)k^2}$$

where $|\xi^{**}(k)| \leq eA_3$.

Equation (17) follows immediately from this and $c + e = a - 1$. On iterating (17) we see that for some constant $C_7 > 0$,

$$d_k \sim C_7 k^{-\left(1+\frac{1}{d}\right)}.$$

□

4 Analysis of the general undirected model

4.1 Linear difference equations with rational coefficients: The method of Laplace

This section summarizes Chapter XV (pages 478-503) of *The Calculus of Finite Differences* by I. M. Milne-Thomson [22].

The equation (14) is an example of a linear difference equation with rational coefficients. It can equivalently be written (for $k > j_0$) as,

$$d_k(a + bk) - d_{k-1}(c + d(k-1)) - \sum_{j=1}^{j_1} d_{k-j} q_j (e + f(k-j)) = 0. \tag{20}$$

The values $d_1, ..., d_{j_0}$ are given by (2), and $d_k = 0$ for $k \leq 0$.

Laplace's difference equation is the name given to the equation in an unknown complex function $u : \boldsymbol{C} \to \boldsymbol{C}$ whose coefficients are linear functions of a complex variable w and an integer l. The general form of the homogeneous equation is

$$\sum_{j=0}^{l} [A_j(w + j) + B_j] u(w + j) = 0. \tag{21}$$

Thus (20) is a special case of (21) with $w + l = k,\ l = j_1$ and $d_{k-j_1+j} = u(w+j)$, and with boundary conditions $u(j_0) = d_{j_0}, ..., u(j_0 - j_1) = d_{j_0 - j_1}$.

A method of solving difference equations with rational coefficients in general, and equation (21) in particular is to use the substitution

$$u(w) = \oint_{\mathcal{C}} t^{w-1} v(t) dt.$$

to obtain

$$\oint_{\mathcal{C}} t^{w-1} (t\phi_1(t) v'(t) - \phi_0(t) v(t)) dt = 0,$$

where

$$\begin{aligned} \phi_1(t) &= A_l t^l + A_{l-1} t^{l-1} + \cdots + A_1 t + A_0 \\ \phi_0(t) &= B_l t^l + B_{l-1} t^{l-1} + \cdots + B_1 t + B_0, \end{aligned}$$

and $\mathcal{C}$ is a suitable contour of integration. ($\phi_1(t)$ is the *characteristic equation.*)

The function $v(t)$ is then obtained as the solution of the differential equation

$$t\phi_1(t) v'(t) - \phi_0(t) v(t) = 0. \tag{22}$$

The general method of solution requires (21) to be of the *Normal type*, namely:

N(i) Both A_l and A_0 are non-zero.

N(ii) The differential equation (22) satisfied by $v(t)$ is of the *Fuchsian* type, (defined next).

Let the roots of the characteristic equation be $a_1, ..., a_l$ (with repetition). The condition that $v(t)$ is of the Fuchsian type, requires that $\phi_0(t)/\phi_1(t)$ can be expressed as a convergent power series of t for some $t > 0$. Thus either the roots $a_1, ..., a_l$ of the characteristic equation must be distinct, or if a is repeated ν times, then a is a root of $\phi_0(t)$ at least $\nu - 1$ times.

Assuming the roots are distinct,

$$\begin{aligned}\frac{v'(t)}{v(t)} &= \frac{\phi_0(t)}{t\phi_1(t)} \\ &= -\frac{\alpha_0}{t} + \frac{\beta_1}{t-a_1} + \cdots + \frac{\beta_l}{t-a_l},\end{aligned} \tag{23}$$

and $\phi_0(t)/\phi_1(t)$ has the required series expansion. The general solution is

$$v^*(t) = t^{-\alpha_0}(t-a_1)^{\beta_1}...(t-a_l)^{\beta_l}.$$

As long as there are no repeated roots, a system of fundamental solutions $(u_j(w),\ j = 1, ..., l)$ is given by

$$u_j(w) = \frac{1}{2\pi i}\oint_{\mathcal{C}_j} t^{w-1-\alpha_0}(t-a_1)^{\beta_1}\cdots(t-a_l)^{\beta_l}\,dt,$$

where $\mathcal{C}_j$ is a contour containing 0 and a_j but excluding the other roots. If β_j is an integer the contour integral is replaced by the integral from 0 to a_j.

A specific solution for $u_j(w)$, valid for $\Re(w) > \alpha_0$, can be obtained as

$$u_j(w) = a_j^w \sum_{m=0}^{\infty} C_m B\left(\frac{w-\alpha_0+\theta-1}{\theta}, \beta_j + m + 1\right)$$

where $B(p, q) = \Gamma(p)\Gamma(q)/\Gamma(p+q)$.

The variable $\theta > 1$ measures the angular separation, about the origin, of the root a_j from the other roots in the transformation $t = a_j z^{1/\theta}$ used to expand the transformed integral about $z = 1$ and obtain the above solution.

Now using the fact that $\Gamma(x) = \sqrt{2\pi}e^{-x}x^{x-1/2}(1 + O(x^{-1}))$, we see that

$$u_j(w) = a_j^w w^{-(1+\beta_j)}(1 + O(w^{-1})). \tag{24}$$

4.2 Application of the technique

Considering the equation (20) we see that

$$\begin{aligned}\phi_1(y) &= y^l - \left(\frac{d+q_1 f}{b}y^{l-1} + \frac{q_2 f}{b}y^{l-2} + \cdots + \frac{q_l f}{b}\right) \\ \phi_0(y) &= \frac{a}{b}y^l - \left(\frac{c+q_1 e}{b}y^{l-1} + \frac{q_2 e}{b}y^{l-2} + \cdots + \frac{q_l e}{b}\right).\end{aligned}$$

We assume that $f > 0$ so that N(i) is satisfied. Let the roots of the characteristic equation be ordered in decreasing size so that $|a_1| \geq |a_2| \geq \cdots \geq |a_l|$. Because of the SOLUTION CONDITIONS we see from Lemma 3, given below, that

$$a_1 = 1$$

and all other roots are either negative or complex and satisfy $|a| < 1$. Considering the partial fraction expansion (23) we see that

$$\phi_0(0) = -\alpha_0 \phi_1(0),$$

so that $\alpha_0 = -e/f$. Also

$$\phi_0(1) = \beta_1 \psi(1), \tag{25}$$

where $\psi(y) = \phi_1(y)/(y-1)$ is given by

$$\psi(y) = y^{l-1} + (1-\alpha_1)y^{l-2} + (1-\alpha_1-\alpha_2)y^{l-3} + \cdots + \\ (1-\alpha_1-\cdots-\alpha_{l-2})y + (1-\alpha_1-\cdots-\alpha_{l-1}), \tag{26}$$

where

$$\alpha_1 = \frac{d+q_1 f}{b}, \qquad \alpha_2 = \frac{q_2 f}{b}, \ldots, \alpha_l = \frac{q_l f}{b}.$$

Using $b = d + f$ we see that $\psi(1) = (d + f\mu_q)/b$. Then using $c + e = a - 1$ we see that $\phi_0(1) = 1/b$ and so (25) implies $\beta_1 = 1/(d + f\mu_q)$. The other β_j require detailed knowledge of the roots of $\phi_1(t)$ and are not relevant to the asymptotic solution.

The solutions $u_j(w)$ are valid for $\Re(w) > \alpha_0 = -e/f$ which includes all $k \geq 0$.

Thus considering the root $a_1 = 1$ we see that

$$u_1(k) = k^{-(1+\beta_1)}\left(1 + O\left(\frac{1}{k}\right)\right)$$

where $\beta_1 = \phi_0(1)/\psi(1) = \frac{1}{d+f\mu_q}$, and $1 + \beta_1$ is the parameter x of our degree sequence.

For $j \geq 2$, we use (24), giving

$$u_j(k) \sim a_j^k k^{-(1+\beta_j)} \to 0,$$

faster than $o(1/k)$, if $|a_j| < 1$.

The specific solution for the sequence $(d_1, d_2, ..., d_k, ...)$ is

$$d_k = b_1 u_1(k) + \cdots + b_l u_l(k),$$

where $u_1(w), ..., u_l(w)$ are the fundamental solutions corresponding to the roots $a_1, ..., a_l$. We note that $b_1 \neq 0$. Indeed from Lemma 2, we know d_k obeys a power law, whereas if $b_1 = 0$, then d_k would decay exponentially as $|a_2|^k$.

Thus the error in the approximation of d_k is $O(1/k)$ from the non-asymptotic expansion of $u_1(w)$, and we conclude

$$d_k = Ck^{-\left(1+\frac{1}{d+f\mu_q}\right)}\left(1 + O\left(\tfrac{1}{k}\right)\right).$$

In the case where $\phi_1(t)$ has other solutions $|a_j| = 1,\ j = 2, ..., j', j' \leq l$, then the asymptotic solution d_k will be a linear combination of k-th powers of these roots.

4.3 Roots of the characteristic equation

Lemma 3. *Let $\alpha_1 = (d + q_1 f)/(d + f)$ and for $2 \leq j \leq l$, let $\alpha_j = q_j f/(d + f)$, and let*

$$\phi_1(z) = z^l - \alpha_1 z^{l-1} - \alpha_2 z^{l-2} - \cdots - \alpha_l.$$

Provided $\alpha_1 > 0$ or $\{q_j\}$ is not periodic, then the solutions of $\phi_1(z) = 0$ are

i) *An un-repeated root at* $z = 1$,

ii) $l-1$ *other (possibly repeated) roots* λ *satisfying* $|\lambda| < 1$.

Proof

We note the following (see Pólya & Szegő [24] p106 16,17). A polynomial $f(z)$ of the form

$$f(z) = z^n - p_1 z^{n-1} - p_2 z^{n-2} - \cdots - p_{n-1} z - p_n,$$

where $p_i \geq 0,\ i = 1, ..., n$ and $p_1 + \cdots + p_n > 0$ has just one positive zero ζ. All other zeros z_0 of $f(z)$ satisfy $|z_0| \leq \zeta$.

Now $\alpha_i \geq 0$ and $\sum \alpha_i = 1$, and so $\phi_1(1) = 0$ and all other zeros, z_0, of $\phi_1(z)$ satisfy $|z_0| \leq 1$.

Let $\psi(z) = \phi_1(z)/(z-1)$ be as in (26). Now $\psi(1)$ is given by

$$1 + (1-\alpha_1) + (1-\alpha_1-\alpha_2) + \cdots + (1-\alpha_1-\cdots-\alpha_{l-1}) = \frac{d + f\mu_q}{d+f}, \tag{27}$$

and thus $\psi(1) \neq 0$, so that $z = 1$ is not a repeated root of ϕ_1.

Let z satisfy $\phi_1(z) = 0$, $|z| = 1$, $z \neq 1$, and let $w = 1/z$; then $\phi_1(z) = 0$ is equivalent to $h(w) = 1$, where

$$h(w) = \alpha_1 w + \alpha_2 w^2 + \cdots + \alpha_l w^l.$$

Suppose there exists $w \neq 1$, on the unit circle satisfying $h(w) = 1$. Let $T = \{w, w^2, ..., w^l\}$. Then all elements of T are points on the unit circle. As $w \neq 1$, $\Re(w) < 1$ and $\Re(w^j) \leq 1,\ j = 2, ..., l$.

Now, by assumption S(i) of Theorem 2, either $\alpha_1 > 0$ or $\alpha_1 = 0$ but $\{q_j\}$ is not periodic.

If $\alpha_1 > 0$, then

$$\sum \alpha_j \Re(w^j) \leq \alpha_1 \Re(w) + \alpha_2 + \cdots \alpha_l < 1,$$

and the conclusion, that $h(w) \neq 1$ follows.

Suppose $\alpha_1 = 0$. If $1 \notin T$, then the real part of w^j satisfies $\Re(w^j) < 1$, contradicting $h(w) = 1$. If $1 \in T$ then $w = e^{2\pi i/m}$ for some integer $m > 1$. However, as $\{q_j\}$ is not periodic, the conclusion that $h(w) < 1$ follows as before. □

The proof of Theorem 2 is now complete.

4.4 Concentration

Here we prove Theorems 3 and 4. Fix t and condition on the following for each $\tau \leq t$: the choice of procedure NEW or OLD and the number T_τ of extra edges added at each step τ. Denote this conditional event by $\mathcal{A}$. (For restricted processes, $\mathcal{A}$ will be the whole space). For the moment we will work entirely within the conditional space $\mathcal{A}$. Given $\mathcal{A}$, let $T = \sum_{\tau \leq t} T_\tau$ and let $Y_1, Y_2, \ldots, Y_T$ be the sequence of single choices of edges created. We let

$$Z_i = \mathbf{E}(D_k(t) \mid Y_1, Y_2, \ldots, Y_i, \mathcal{A}) - \mathbf{E}(D_k(t) \mid Y_1, Y_2, \ldots, Y_{i-1}, \mathcal{A})$$

and prove that

$$|Z_i| \leq 4. \tag{28}$$

The Azuma-Hoeffding martingale inequality then implies that

$$\mathbf{Pr}(|D_k(t) - \overline{D}_k(t)| \geq u \mid \mathcal{A}) \leq \exp\left\{-\frac{u^2}{8T}\right\}. \tag{29}$$

This is enough for the proof of Theorem 4.

Fix $Y_1, Y_2, \ldots, Y_i$ and let $Y_i = (x, v)$, $\hat{Y}_i = (\hat{x}, \hat{v})$. Of course $\hat{x} = x$ if Y_i if the choice Y_i is not the beginning of a time step. Then for each complete outcome $\mathbf{Y} = Y_1, Y_2, \ldots, Y_T$ we define a corresponding outcome $\hat{\mathbf{Y}} = Y_1, Y_2, \ldots, Y_{i-1}, \hat{Y}_i, \ldots, \hat{Y}_T$ where for $j > i$, $\hat{Y}_j$ is obtained from Y_j as follows: If Y_j creates a new edge (w, v) by randomly choosing edge (x, v) arising from Y_i, then in $\hat{Y}_j$, (w, v) is replaced by $(w, \hat{v})$, otherwise $\hat{Y}_j = Y_j$.

The map $\mathbf{Y} \to \hat{\mathbf{Y}}$ is measure preserving and in going from $\mathbf{Y}$ to $\hat{\mathbf{Y}}$ only the degrees of $x, \hat{x}, v$ and $\hat{v}$ change and so the number of vertices of degree k changes by at most 4 and (28) follows.

Continuing with the proof of Theorem 3, fix $\mathcal{A} \in \mathcal{A}_0$ and define $\alpha_{\mathcal{A}} = \alpha_{\mathcal{A}}(t)$ to be the indicator for an old node to generate edges at time t. Define $\beta_{\mathcal{A}}, \ldots, \gamma_{\mathcal{A}}$ similarly. Also define $p_{i,\mathcal{A}} = p_{i,\mathcal{A}}(t)$ to be the indicator that exactly i edges are generated from a new node at time t and define the $q_{i,\mathcal{A}}$ similarly. Next let $\nu_{\mathcal{A}}(t)$, $\eta_{\mathcal{A}}(t)$ denote the number of vertices and edges at time t, given $\mathcal{A}$.

Going back to (7) – (10) we can write

$$
\begin{aligned}
\overline{D}_k^{\mathcal{A}}(t+1) &= \overline{D}_k^{\mathcal{A}}(t) + (1-\alpha_{\mathcal{A}})p_{k,\mathcal{A}} + \\
&+ (1-\alpha_{\mathcal{A}})\sum_{j=1}^{j_0} p_{j,\mathcal{A}} \left(\frac{\beta_{\mathcal{A}} j \overline{D}_{k-1}^{\mathcal{A}}(t)}{\nu_{\mathcal{A}}(t)} - \frac{\beta_{\mathcal{A}} j \overline{D}_k^{\mathcal{A}}(t)}{\nu_{\mathcal{A}}(t)} + (1-\beta_{\mathcal{A}}) \left(\frac{j(k-1)\overline{D}_{k-1}^{\mathcal{A}}(t)}{\eta_{\mathcal{A}}(t)} - \frac{jk\overline{D}_k^{\mathcal{A}}(t)}{\eta_{\mathcal{A}}(t)} \right) \right) \\
&- \alpha_{\mathcal{A}} \left(\frac{\delta_{\mathcal{A}} \overline{D}_k^{\mathcal{A}}(t)}{\nu_{\mathcal{A}}(t)} + \frac{(1-\delta_{\mathcal{A}})k\overline{D}_k^{\mathcal{A}}(t)}{\eta_{\mathcal{A}}(t)} \right) + \alpha_{\mathcal{A}} \sum_{j=1}^{j_1} q_{j,\mathcal{A}} \left(\frac{\delta_{\mathcal{A}} \overline{D}_{k-j}^{\mathcal{A}}(t)}{\nu_{\mathcal{A}}(t)} + \frac{(1-\delta_{\mathcal{A}})(k-j)\overline{D}_{k-j}^{\mathcal{A}}(t)}{\eta_{\mathcal{A}}(t)} \right) \\
&+ \alpha_{\mathcal{A}} \sum_{j=1}^{j_1} j q_{j,\mathcal{A}} \left(\gamma_{\mathcal{A}} \left(\frac{\overline{D}_{k-1}^{\mathcal{A}}(t)}{\nu_{\mathcal{A}}(t)} - \frac{\overline{D}_k^{\mathcal{A}}(t)}{\nu_{\mathcal{A}}(t)} \right) + (1-\gamma_{\mathcal{A}}) \left(\frac{(k-1)\overline{D}_{k-1}^{\mathcal{A}}(t)}{\eta_{\mathcal{A}}(t)} - \frac{k\overline{D}_k^{\mathcal{A}}(t)}{\eta_{\mathcal{A}}(t)} \right) \right).
\end{aligned}
$$

where $\overline{\mathcal{D}}_k^{\mathcal{A}}(t) = \mathbf{E}(D_k(t) \mid \mathcal{A})$.

Continuing this line we reach the following in place of (13): $\Delta_k^{\mathcal{A}}(t) = \overline{\mathcal{D}}_k^{\mathcal{A}}(t) - td_k$.

$$
\begin{aligned}
\Delta_k^{\mathcal{A}}(t+1) = \Delta_k^{\mathcal{A}}(t) \left(1 - \frac{a_{\mathcal{A}} + b_{\mathcal{A}}k - 1}{t}\right) + \zeta(t) \\
+ \frac{1}{t} \left((c_{\mathcal{A}} + d_{\mathcal{A}}(k-1))\Delta_{k-1}^{\mathcal{A}}(t) + \sum_{j=1}^{j_1} (e_{\mathcal{A}} + f_{\mathcal{A}}(k-j)) q_{j,\mathcal{A}} \Delta_{k-j}^{\mathcal{A}}(t) \right) \qquad (30)
\end{aligned}
$$

where

$$
\begin{aligned}
a_{\mathcal{A}} &= a_{\mathcal{A}}(t) = 1 + \frac{(1-\alpha_{\mathcal{A}})\beta_{\mathcal{A}}\mu_{p,\mathcal{A}}}{t^{-1}\nu_{\mathcal{A}}(t)} + \frac{\alpha_{\mathcal{A}}\gamma_{\mathcal{A}}\mu_{q,\mathcal{A}}}{t^{-1}\nu_{\mathcal{A}}(t)} + \frac{\alpha_{\mathcal{A}}\delta_{\mathcal{A}}}{t^{-1}\nu_{\mathcal{A}}(t)} \text{ etc} \\
\zeta(t) &= (1-\alpha_{k,\mathcal{A}})p_{k,\mathcal{A}} + (c_{\mathcal{A}} + d_{\mathcal{A}}(k-1))d_{k-1} + \sum_{j=1}^{k-1} (e_{\mathcal{A}} + f_{\mathcal{A}}(k-j)) q_{j,\mathcal{A}} d_{k-j} - d_k(a_{\mathcal{A}} + b_{\mathcal{A}}k)
\end{aligned}
$$

Now let $s = t^{1/21}$ and for $\tau > s$ let

$$\begin{aligned}\Sigma_k^{\mathcal{A}}(\tau) &= \frac{1}{s+1}\sum_{\ell=0}^{s}\Delta_k^{\mathcal{A}}(\tau-\ell)\\ \xi(\tau) &= (s+1)\Delta_k^{\mathcal{A}}(\tau-s) - \sum_{\ell=0}^{s}\Delta_k^{\mathcal{A}}(\tau-\ell)\end{aligned}$$

Then, using $|\Delta_k^{\mathcal{A}}(\tau-\ell) - \Delta_k^{\mathcal{A}}(\tau-s)| = d_k(s-\ell) + O(s)$ for $0 \le \ell \le s$, we see from (30) that

$$\begin{aligned}&(s+1)\Sigma_k^{\mathcal{A}}(\tau+1) = \Delta_k^{\mathcal{A}}(\tau-s)\sum_{\ell=0}^{s}\left(1-\frac{a_{\mathcal{A}}(\tau-\ell)+b_{\mathcal{A}}(\tau-\ell)k-1}{\tau}\right) - \xi(\tau) + \sum_{\ell=0}^{s}\zeta(\tau-\ell) + O(s^2k/\tau)\\ &+\frac{1}{\tau}\sum_{\ell=0}^{s}\left((c_{\mathcal{A}}(\tau-\ell)+(k-1)d_{\mathcal{A}}(\tau-\ell))\Delta_{k-1}^{\mathcal{A}}(\tau-s) + \sum_{j=1}^{j_1}(e_{\mathcal{A}}(\tau-\ell)+(k-j)f_{\mathcal{A}}(\tau-\ell)q_{j,\mathcal{A}})\Delta_{k-j}^{\mathcal{A}}(\tau-s)\right).\end{aligned} \tag{31}$$

Next let

$$\begin{aligned}\mathcal{A}_0 &= \{\mathcal{A} : \ |\nu_{\mathcal{A}}(\tau) - (1-\alpha)\tau| < \tau^{1/2}\log t \text{ and } |\eta_{\mathcal{A}}(\tau) - \theta\tau| < \tau^{1/2}\log t \text{ for } \tau > t^{1/5}\}.\\ \mathcal{A}_1 &= \left\{\mathcal{A} : \left|\sum_{\ell=0}^{s} x_{\mathcal{A}}(\tau-\ell) - (s+1)x\right| \le s^{1/2}\log t,\ x = a, b, \ldots, \text{for } \tau > t^{1/5}\right\}.\end{aligned}$$

It follows easily that

$$\mathbf{Pr}(\mathcal{A} \notin \mathcal{A}_0 \cap \mathcal{A}_1) = O(t^{-\Omega(\log t)}) \tag{32}$$

Assuming that $\mathcal{A} \in \mathcal{A}_0 \cap \mathcal{A}_1$ we can deduce from (31) that

$$\begin{aligned}\Sigma_k^{\mathcal{A}}(\tau+1) &= \Delta_k^{\mathcal{A}}(\tau-s)\left(1-\frac{a+bk-1}{\tau}\right) - \frac{\xi(\tau)}{s+1} + O(\tau^{-1/2}\log t)\\ &\quad+\frac{1}{\tau}\left((c+(k-1)d)\Delta_{k-1}^{\mathcal{A}}(\tau-s) + \sum_{j=1}^{j_1}(e+(k-j)fq_j)\Delta_{k-j}^{\mathcal{A}}(\tau-s)\right)\\ &= \Sigma_k^{\mathcal{A}}(\tau)\left(1-\frac{a+bk-1}{\tau}\right) + O(\tau^{-1/2}\log t) +\\ &\quad+\frac{1}{\tau}\left((c+(k-1)d)\Sigma_{k-1}^{\mathcal{A}}(\tau) + \sum_{j=1}^{j_1}(e+(k-j)fq_j)\Sigma_{k-j}^{\mathcal{A}}(\tau)\right)\end{aligned} \tag{33}$$

We inductively prove the following inequalities for $\kappa \le k_1 = t^{1/21}$ and for some sufficiently large M:

$$|\Sigma_\kappa^{\mathcal{A}}(\tau)| \le \begin{cases} M\tau^{21/40}t^{1/5} & \tau \le t^{8/9}\\ M\tau^{3/4} & t^{8/9} \le \tau \le t.\end{cases} \tag{34}$$

This holds trivially for $\tau \le t^{1/5}$ and so assume inductively that (34) holds for some $\tau \ge t^{1/5}$. We write the RHS of (34) as $M\tau^\rho t^\sigma$ to cover both ranges i.e. $(\rho, \sigma) = (21/40, 1/5)$ or $(3/4, 0)$. Observe that at $\tau = t^{8/9}$ the two expressions coincide. Then we have,

$$\begin{aligned}
\Sigma_k^{\mathcal{A}}(\tau+1) &\leq M\tau^\rho t^\sigma \left(1 - \frac{a+bk-1}{\tau}\right) + O(\tau^{-1/2}\log t) + \\
&\quad + M\tau^\rho t^\sigma \frac{1}{\tau}\left(c + (k-1)d + \sum_{j=1}^{j_1}(e + (k-j)fq_j)\right) \\
&= M\tau^\rho t^\sigma + O(\tau^{-1/2}\log t) \\
&\leq M(\tau+1)^\rho t^\sigma,
\end{aligned}$$

for large enough M, completing the induction.

Thus (34) holds for all $\tau \leq t$, $k \leq k_1$, provided $\mathcal{A} \in \mathcal{A}_0 \cap \mathcal{A}_1$. Now $|\Delta_k^{\mathcal{A}}(\tau) - \Sigma_k^{\mathcal{A}}(\tau)| = O(s)$ and so

$$\max\{\Delta_k^{\mathcal{A}}(\tau) :\ k \leq k_1, \tau \leq t\} \leq \max\{\Sigma_k^{\mathcal{A}}(\tau) :\ k \leq k_1, \tau \leq t\} + O(s)$$

and so Theorem 3 then follows from (29), (32). □

5 Maximum Degree

Here we prove Theorem 5. First observe that the degree of vertex 1 stochasically dominates the degree of every other vertex.

Now let X_t denote the degree of vertex 1 in $G(t)$. We take $X_0 = 1$. Then, after using (5) and (6) we see that for $t \geq 1$,

$$\mathbf{E}(X_{t+1} \mid X_t) = \left(1 + \frac{A}{t}\right) X_t + O(t^{-1/2}\ln t) \tag{35}$$

where $A = d + f\mu_q$. (Note that $A \leq 1$).

Let L be the hidden constant in (35). Then $x_t = \mathbf{E}(X_t)$ satisfies the recurrence: $x_0 = 1$ and for $t \geq 1$

$$0 \leq x_{t+1} \leq \left(1 + \frac{A}{t}\right) x_t + Lt^{-1/2}\ln t. \tag{36}$$

We first argue that we can find K_1, K_2, τ_0 such that if (36) holds

$$K_2\left(A - \frac{1}{2} - \frac{2}{\ln\tau}\right) - L > \frac{K_1}{\tau^{3/2-A}} \qquad \tau \geq \tau_0. \tag{37}$$

$$K_1\tau^A - K_2\tau^{1/2}\ln\tau \geq x_\tau \qquad 1 \leq \tau < \tau_0. \tag{38}$$

To see this, first use induction to show that (36) implies

$$x_\tau \leq (A + 2L)\tau \qquad \tau \geq 1. \tag{39}$$

To satisfy (37,38) we will take

$$\tau_0 = K_1^{1/(3/2-A)}.$$

So if K_1 is sufficiently large and $K_2\left(A - \frac{1}{2}\right) > 2(L+1)$ then (37) will hold. On the other hand to satisfy (38) we have only to choose K_1 large enough so that

$$K_1 \geq K_2 + (A + 2L)K_1^{(1-A)/(3/2-A)}$$

and this is clearly possible.

We next argue inductively that if (37,38) hold for some $t \geq \tau_0$ then they also hold for $t+1$. The verification of (37) is trivial and we have from (36) that

$$\begin{aligned} x_{t+1} &\leq K_1 t^A - K_2 t^{1/2} \ln t + AK_1 t^{A-1} - AK_2 t^{-1/2} \ln t + Lt^{-1/2} \ln t \\ &= K_1(t+1)^A - K_2(t+1)^{1/2} \ln(t+1) - \Theta_t \end{aligned}$$

where

$$\begin{aligned} \Theta_t &\geq K_1((t+1)^A - t^A) - AK_1 t^{A-1} - K_2((t+1)^{1/2} \ln(t+1) - t^{1/2} \ln t) + (AK_2 - L)t^{-1/2} \ln t \\ &\geq -\frac{K_1}{t^{A-2}} + \left(K_2 \left(A - \frac{1}{2}\right) - L\right) t^{-1/2} \ln t - 2K_2 t^{-1/2} \\ &\geq 0 \end{aligned}$$

after using (37). So now we have proved that (38) holds for all $t \geq 1$ and so the RHS of (4) holds in expectation. A similar analysis yields the lower bound on the expectation.

To prove concentration we can use a martingale argument similar to that used in the proof of Theorem 3.

6 Directed variants of the model

A curious phenomena of the directed models is that they are *incomplete* in the sense that the sampling procedure for terminal vertices in the out-model (resp. initial vertices in the in-model) does not need to be specified in order to estimate the proportion of vertices of out-degree k (resp. proportion of vertices of in-degree k). Thus these are not models, but classes of models. For the out-model, for example, terminal vertices can be picked according to *any* rule: assign, copy, direct all edges to vertex 1 etc.

6.1 Sampling based on out-degree

Let $\theta^+ = (1-\alpha)\mu_p + \alpha\mu_q$. The estimate (7-10) is replaced by

$$\overline{D}_k^+(t) = \overline{D}_k^+(t-1) + (1-\alpha)p_k + \alpha \left(\sum_{j=1}^{j_1} q_j \left(\tfrac{\delta}{(1-\alpha)t} \left(\overline{D}_{k-j}^+ - \overline{D}_k^+ \right) + \tfrac{1-\delta}{t\theta^+} \left((k-j)\overline{D}_{k-j}^+ - k\overline{D}_k^+ \right) \right) \right) + O(t^{-1/2} \log t).$$

Then for $k \geq 1$ we obtain

$$d_k^+(1 + e + fk) = (1-\alpha)p_k + \sum_{j=1}^{j_1} d_{k-j}^+ q_j (e + f(k-j)).$$

For large k this is a rational difference equation with characteristic equation

$$\phi_1(y) = y^{j_1} - \left(q_1 y^{j_1 - 1} + \cdots + q_{j_1}\right).$$

Thus provided $\phi_1(y)$ has no repeated roots, and $f > 0$,

$$x^+ = 1 + 1/(f\mu_q) = 1 + \frac{(1-\alpha)\mu_p + \alpha\mu_q}{\alpha(1-\delta)\mu_q}. \tag{40}$$

and

$$d_k^+ \sim Ck^{-x^+}.$$

6.2 Sampling based on in-degree

Let $\theta = \theta^+$ as given above. For $k < 0$ let $D_k^- = 0$. Then we have,

$$
\begin{aligned}
D_k^-(t) &= D_k^-(t-1) + +O(t^{-1/2}\log t) + (1-\alpha)1_{k=0} \\
&+ (1-\alpha)\left(-\sum_{j=1}^{j_0} p_j\left(\tfrac{\beta j \overline{D}_k^-}{(1-\alpha)t} + \tfrac{(1-\beta)jk\overline{D}_k^-}{\theta(t-1)}\right) + \sum_{j=1}^{j_0} p_j\left(\tfrac{\beta j \overline{D}_{k-1}^-}{(1-\alpha)t} + \tfrac{(1-\beta)j(k-1)\overline{D}_{k-1}^-}{\theta t}\right)\right) \\
&+ \alpha\left(\sum_{j=1}^{j_1} q_j\left[\gamma\left(-\tfrac{j\overline{D}_k^-}{(1-\alpha)t} + \tfrac{j\overline{D}_{k-1}^-}{(1-\alpha)t}\right) + (1-\gamma)\left(-\tfrac{jk\overline{D}_k^-}{\theta t} + \tfrac{j(k-1)\overline{D}_{k-1}^-}{\theta t}\right)\right]\right).
\end{aligned}
$$

For $k \geq 0$ we find,

$$
\begin{aligned}
d_0^- &= \frac{(1-\alpha)^2}{(1-\alpha)(1+\beta\mu_p) + \alpha\gamma\mu_q} \\
d_k^- &= \frac{c + d(k-1)}{1+c+dk} d_{k-1}^- \qquad\qquad k \geq 1.
\end{aligned}
$$

The solution of this case is by direct iteration, and is similar in form to Theorem 2 (ii), with $f = 0$, giving

$$
x^- = 1 + \frac{1}{d} = 1 + \frac{(1-\alpha)\mu_p + \alpha\mu_q}{(1-\alpha)(1-\beta)\mu_p + \alpha(1-\gamma)\mu_q}. \tag{41}
$$

References

[1] R. Albert, A. Barabasi and H. Jeong, *Diameter of the world wide web.* Nature 401, 103-131, (1999).

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

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

[4] B. Bollobás and O. Riordan, *The diameter of a scale free random graph*, to appear in Combinatorica.

[5] B. Bollobás and O. Riordan, *Mathematical results on scale-free graphs*, Handbook of Graphs and Networks, S. Bornholdt, H. Schuster (eds), Wiley-VCH, Berlin (2002).

[6] S. Brin and L. Page, *The anatomy of a large scale hypertextual web search engine*, Computer Networks 30, 107-117 (1998).

[7] A. Broder, R. Kumar, F.Maghoul, P. Raghavan, S. Rajagopalan, R. Stata, A. Tomkins and J. Wiener, *Graph structure in the web*, Proceedings of the 9th International World Wide Web Conference, 309-320, (2000).

[8] P. Buckley and D. Osthus, *Popularity based random graph models leading to a scale-free degree sequence*, preprint available from http://www.informatik.hu-berlin.de/~osthus/.

[9] C. Cooper and A. M. Frieze, *Crawling on web graphs*, Proceedings of the 34th Annual ACM Symposiuim on Theory of Computing, 419-427, (2002).

[10] C. Cooper, *Distribution of vertex degree in web graphs*, submitted (2002).

[11] S.N. Dorogovtsev, J.F.F. Mendes and A.N. Samukhin, *Structure of growing networks with preferential linking*, Phys. Rev. Lett. 85, 4633, (2000).

[12] E. Drinea, M. Enachescu and M. Mitzenmacher, *Variations on random graph models for the web*, Technical report, Harvard University, Department of Computer Science, (2001).

[13] P. Erdős and A. Rényi, *On random graphs I*, Publicationes Mathematicae Debrecen 5, 290-297, (1959).

[14] P. Erdős and A. Rényi, *On the evolution of random graphs*, Magyar Tud. Akad. Mat. Kutató Int. Közl. 5, 17-61, (1960).

[15] M. Faloutsos, P. Faloutsos and C. Faloutsos, *On power-law relationships of the internet topology*, SIGCOMM 1999, Comput. Commun. Rev. 29, 251, (1999).

[16] B. Hayes, *Graph theory in practice: Part II.* American Scientist 88, 104-109. (2000). `http://www.sigmaxi.org/amsci/issues/Comsci00/compsci2000-03.html`

[17] W. Hoeffding, *Probability inequalities for sums of bounded random variables*, Journal of the American Statistical Association 58, 13-30 (1963).

[18] S.Janson, T.Łuczak and A.Ruciński, *Random Graphs*, John Wiley and Sons, (2000).

[19] J. Kleinberg. Authoritative sources in a hyperlinked environment. *Proc. 9th ACM-SIAM Symposium on Discrete Algorithms,* (1998), *Journal of the ACM* 46(1999).

[20] R. Kumar, P. Raghavan, S. Rajagopalan, D. Sivakumar, A. Tomkins and E. Upfal, *The web as a graph*, Proceedings of the 19th ACM Symposium on Principles of Database Systems, 1-10, (2000).

[21] R. Kumar, P. Raghavan, S. Rajagopalan, D. Sivakumar, A. Tomkins and E. Upfal, *Stochastic models for the web graph*, Proceedings of the 41th IEEE Symp. on Foundations of Computer Science, 57-65, (2000).

[22] I. M. Milne-Thomson, *The Calculus of Finite Differences.* Macmillian, London (1951).

[23] , M. Mitzenmacher, A brief history of generative models for power law and lognormal distributions, to appear.

[24] G. Pólya and G. Szegő, *Problems and Theorems in Analysis I.* Springer, Berlin (1970).

[25] H. A. Simon, On a class of skew distribution functions, *Biometrika* 42 (1955) 425-440.

[26] D. Watts, *Small Worlds: The Dynamics of Networks between Order and Randomness*, Princeton University Press (1999).

[27] , 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.