The degree sequence of a scale-free random graph process

Béla Bollobás*†‡ Oliver Riordan† Joel Spencer§
Gábor Tusnády¶

October 3, 2002

Abstract

Recently, Barabási and Albert [2] suggested modeling complex real-world networks such as the world-wide web as follows: consider a random graph process in which vertices are added to the graph one at a time and joined to a fixed number of earlier vertices, selected with probabilities proportional to their degrees. In [2] and, with Jeong, in [3], Barabási and Albert suggested that after many steps the proportion $P(d)$ of vertices with degree d should obey a power law $P(d) \; \alpha \; d^{-\gamma}$. They obtained $\gamma = 2.9 \pm 0.1$ by experiment, and gave a simple heuristic argument suggesting that $\gamma = 3$. Here we obtain $P(d)$ asymptotically for all $d \leq n^{1/15}$, where n is the number of vertices, proving as a consequence that $\gamma = 3$.

1 Introduction

Recently there has been considerable interest in using random graphs to model complex real-world networks to gain an insight into their properties. One of the most basic properties of a graph or network is its degree sequence. For the standard random graph model $\mathcal{G}(n, m)$ of all graphs with m edges on a fixed set of n vertices, introduced by Erdős and Rényi in [8] and studied in detail in [9], there is a 'characteristic' degree $2m/n$: the vertex degrees have approximately a Poisson or normal distribution with mean $2m/n$. The same applies to the closely related model $\mathcal{G}(n, p)$ introduced by Gilbert [10], where vertices are joined independently with probability p. In contrast, Barabási and Albert [2], as well as several other groups (see [4, 14] and the references therein), noticed that in many real-world examples the degree sequence has a 'scale-free' power law

*Department of Mathematical Sciences, University of Memphis, Memphis TN 38152, USA
†Trinity College, Cambridge CB2 1TQ, UK
‡Research supported in part by NSF grant no. DSM 9971788
§Courant Institute of Mathematical Sciences, New York University, New York NY 10003, USA
¶Rényi Institute, Budapest, Hungary

distribution: the fraction $P(d)$ of vertices with degree d is proportional over a large range to $d^{-\gamma}$, where γ is a constant independent of the size of the network. To explain this phenomenon, Barabási and Albert [2] suggested the following random graph process as a model.

> ... starting with a small number (m_0) of vertices, at every time step we add a new vertex with $m(\leq m_0)$ edges that link the new vertex to m different vertices already present in the system. To incorporate preferential attachment, we assume that the probability Π that a new vertex will be connected to a vertex i depends on the connectivity k_i of that vertex, so that $\Pi(k_i) = k_i/\sum_j k_j$. After t steps the model leads to a random network with $t+m_0$ vertices and mt edges.

This process is intended as a highly simplified model of the growth of the world-wide web, for example, the vertices representing sites or web pages, and the edges links from sites to earlier sites. The preferential attachment assumption is based on the idea that a new site is more likely to link to existing sites which are 'popular' at the time the site is added. For $m = 1$ this process is very similar to the nonuniform random recursive tree process considered in [18, 17, 15]. An alternative model, replacing the preferential attachment assumption by a notion of 'link copying' is given in [12, 14]. We shall discuss these models briefly in the final section.

In [2, 3] it is stated that computer experiments for the process above suggest that $P(d)\ \alpha\ d^{-\gamma}$ with $\gamma = 2.9 \pm 0.1$. In [3], the following heuristic argument is given to suggest that $\gamma = 3$: consider the degree d_i of the i^{th} new vertex v_i at time t, i.e., when there are $t + m_0$ vertices and mt edges. When a new vertex is added, the probability that it is joined to v_i is md_i over the sum of the degrees, i.e., over $2mt$. This suggests the 'mean-field theory'

$$\frac{\mathrm{d}d_i}{\mathrm{d}t} = \frac{d_i}{2t}.$$

With the initial condition that $d_i = m$ when $t = i$ this gives $d_i = m(t/i)^{1/2}$, which yields $\gamma = 3$.

Here we show how one can calculate the exact distribution of d_i at time t, and obtain an asymptotic formula for $P(d)$, $d \leq t^{1/15}$, which gives $\gamma = 3$ as a simple consequence. The first step is to give an exact definition of a random graph process that fits the rather vague description given above.

2 The model

The description of the random graph process quoted above is rather imprecise. Firstly, as the degrees are initially zero, it is not clear how the process is supposed to get started. More seriously, the expected number of edges linking a new vertex v to earlier vertices is $\sum_i \Pi(k_i) = 1$, rather than m. Also, when choosing in one go a set S of m earlier vertices as the neighbours of v, the distribution of

S is *not* specified by giving the marginal probability that each vertex lies in S. For a trivial example, suppose that $m = 2$ and that the first four vertices form a four-cycle. Then for any $0 \leq \alpha \leq 1/4$ we could join the fifth vertex to each adjacent pair with probability α and to each non-adjacent pair with probability $1/2 - 2\alpha$. This suggests that for $m > 1$ we should choose the neighbours of v one at a time. Once doing so, it is very natural to allow some of these neighbours to be the same, creating multiple edges in the graph. Here we shall consider the precise model introduced in [6], which turns out to be particularly pleasant to work with. This model fits the description above except that it allows multiple edges and also loops—in terms of the interpretation there is no reason to exclude these. Once the process gets started there will in any case not be many loops or multiple edges, so they should have little effect overall. The following definition is essentially as given in [6]; we write $d_G(v)$ for the total (in plus out) degree of the vertex v in the graph G.

We start with the case $m = 1$. Consider a fixed sequence of vertices $v_1, v_2, \ldots$. We shall inductively define a random graph process $(G_1^t)_{t \geq 0}$ so that G_1^t is a directed graph on $\{v_i : 1 \leq i \leq t\}$, as follows. Start with G_1^0 the 'graph' with no vertices, or with G_1^1 the graph with one vertex and one loop. Given G_1^{t-1}, form G_1^t by adding the vertex v_t together with a single edge directed from v_t to v_i, where i is chosen randomly with

$$\mathbb{P}(i = s) = \begin{cases} d_{G_1^{t-1}(v_s)}/(2t-1) & 1 \leq s \leq t-1 \\ 1/(2t-1) & s = t \end{cases} \tag{1}$$

In other words, we send an edge e from v_t to a random vertex v_i, where the probability that a vertex is chosen as v_i is proportional to its (total) degree at the time, counting e as already contributing one to the degree of v_t. For $m > 1$ we add m edges from v_t one at a time, counting the previous edges as well as the 'outward half' of the edge being added as already contributing to the degrees. Equivalently, we define the process $(G_m^t)_{t \geq 0}$ by running the process (G_1^t) on a sequence $v_1', v_2', \ldots$; the graph G_m^t is formed from G_1^{mt} by identifying the vertices $v_1', v_2', \ldots, v_m'$ to form v_1, identifying $v_{m+1}', v_{m+2}', \ldots v_{2m}'$ to form v_2, and so on.

We shall write $\mathcal{G}_m^n$ for the probability space of directed graphs on n vertices $v_1, v_2, \ldots v_n$ where a random $G_m^n \in \mathcal{G}_m^n$ has the distribution derived from the process above. As G_m^n is defined in terms of G_1^{mn}, for most of the time we shall consider the case $m = 1$. As noted in [6], there is an alternative description of the distribution of G_1^n in terms of pairings.

An *n-pairing* $\mathcal{P}$ is a partition of the set $\{1, 2, \ldots 2n\}$ into pairs, so there are $(2n-1)!! = (2n)!/(n!2^n)$ n-pairings. Thinking of the elements $\{1, 2, \ldots 2n\}$ of the ground set as points on the x axis, and the pairs as chords joining them, we shall speak of the left and right endpoint of each pair.

We form a directed graph $\phi(\mathcal{P})$ from an n-pairing $\mathcal{P}$ as follows: starting from the left, merge all endpoints up to and including the first right endpoint reached to form the vertex v_1. Then merge all further endpoints up to the next right endpoint to form v_2, and so on to v_n. For the edges, replace each pair by a directed edge from the vertex corresponding to its right endpoint to that corresponding to its left endpoint. As noted in [6], if $\mathcal{P}$ is chosen uniformly at random from all $(2n-1)!!$ n-pairings, then $\phi(\mathcal{P})$ has the same distribution as a random $G_1^n \in \mathcal{G}_1^n$. This statement is easy to prove by induction on n: thinking in terms of pairings of distinct points on the x axis, one can obtain a random $(n-1)$-pairing from a random n-pairing by deleting the pair containing the rightmost point. The reverse process, starting from an $(n-1)$-pairing $\mathcal{P}$, is to add a new pair with its right endpoint to the right of everything in $\mathcal{P}$, and its left endpoint in one of the $2n-1$ possible places. Now a vertex of degree d in $\phi(\mathcal{P})$ corresponds to d intervals between endpoints in $\mathcal{P}$. The effect of adding a new pair to $\mathcal{P}$ as described is thus to add a new vertex to $\phi(\mathcal{P})$ together with a new edge to a vertex chosen according to (1), with $t = n$.

The advantage of this description from pairings is that it gives us a simple non-recursive definition of the distribution of G_m^n, enabling us to calculate properties of G_m^n directly. We now use this to study the degrees of G_m^n.

3 The degrees of G_m^n

In [2] it was suggested that the fraction of vertices of G_m^n having degree d should fall off as d^{-3} as $d \to \infty$. We shall prove the following precise version of this statement, writing $\#_m^n(d)$ for the number of vertices of G_m^n with *indegree* equal to d, i.e., with (total) degree $m + d$.

Theorem 1. *Let $m \geq 1$ be fixed, and let $(G_m^n)_{n\geq 0}$ be the random graph process defined in section 2. Let*

$$\alpha_{m,d} = \frac{2m(m+1)}{(d+m)(d+m+1)(d+m+2)},$$

and let $\epsilon > 0$ be fixed. Then with probability tending to 1 as $n \to \infty$ we have

$$(1-\epsilon)\alpha_{m,d} \leq \frac{\#_m^n(d)}{n} \leq (1+\epsilon)\alpha_{m,d}$$

for every d in the range $0 \leq d \leq n^{1/15}$.

In turns out that we only need to calculate the expectation of $\#_m^n(d)$; the concentration result is then given by applying the following standard inequality due to Azuma [1] and Hoeffding [11] (see also [5]).

Lemma 2 (Azuma-Hoeffding inequality). *Let $(X_t)_{t=0}^n$ be a martingale with $|X_{t+1} - X_t| \leq c$ for $t = 0, \ldots, n-1$. Then*

$$\mathbb{P}(|X_n - X_0| \geq x) \leq \exp\left(-\frac{x^2}{2c^2 n}\right).$$

The strategy of the proof is as follows. Firstly, as mentioned earlier, the results for general m will follow from those for $m = 1$. We shall use the pairing model to find explicitly the distribution of D_k, the sum of the first k degrees, in this case, and also the distribution of the next degree, $d_{G_1^n}(v_{k+1})$, given D_k. One could combine these formulae to give a rather unilluminating expression for the distribution of $d_{G_1^n}(v_{k+1})$; instead we show that D_k is concentrated about a certain value, and hence find approximately the probability that $d_{G_1^n}(v_{k+1}) = d$. Summing over k gives us the expectation of $\#_1^n(d)$, and concentration follows from Lemma 2.

Before turning to the distributions of the (total) degrees for $m = 1$, we note that their expectations are easy to calculate exactly:

$$\mathbb{E}(d_{G_1^t}(v_t)) = 1 + \frac{1}{2t-1}.$$

Also, for $s < t$,

$$\mathbb{E}(d_{G_1^t}(v_s) \mid d_{G_1^{t-1}}(v_s)) = d_{G_1^{t-1}}(v_s) + \frac{d_{G_1^{t-1}}(v_s)}{2t-1},$$

which implies that

$$\mathbb{E}(d_{G_1^t}(v_s)) = \frac{2t}{2t-1}\,\mathbb{E}(d_{G_1^{t-1}}(v_s)).$$

Thus, for $1 \le s \le n$,

$$\mathbb{E}(d_{G_1^n}(v_s)) = \prod_{i=s}^{n} \frac{2i}{2i-1} = \frac{4^{n-s+1} n!^2 (2s-2)!}{(2n)!(s-1)!^2} = \sqrt{n/s}\,(1 + O(1/s)),$$

using Stirling's formula.

If every degree of G_1^n were equal to its expectation this would give the proposed distribution, but in fact the degrees can be far from their expectations. Indeed we shall see that for almost all vertices the most likely degree is 1!

Let us write d_i for $d_{G_1^n}(v_i)$, i.e., for the (total) degree of the vertex v_i in the graph G_1^n. Our aim is to describe the distributions of the individual d_i. To do this it turns out to be useful to consider their sums $D_k = \sum_{i=1}^{k} d_i$.

Consider first the event $\{D_k - 2k = s\}$, where $0 \le s \le n-k$. This is the event that the last $n-k$ vertices of G_1^n send exactly s edges to the first k vertices. This event corresponds to pairings $\mathcal{P}$ in which the k^{th} right endpoint is $2k+s$. Consider any pairing $\mathcal{P}$ with this property. We shall split $\mathcal{P}$ into two partial pairings, the *left partial pairing* $\mathcal{L}$, and the *right partial pairing* $\mathcal{R}$, each consisting of some number of pairs together with some unpaired elements. For $\mathcal{L}$ we take the partial pairing on $\{1, \dots, 2k+s\}$ induced by $\mathcal{P}$, for $\mathcal{R}$ that on $\{2k+s+1, \dots, 2n\}$. From the restriction on $\mathcal{P}$, in $\mathcal{L}$ the element $2k+s$ must be paired with one of $\{1, \dots, 2k+s-1\}$, precisely s of the remaining $2k+s-2$ elements must be unpaired, and the other $2(k-1)$ elements must be paired off somehow. Any of the

$$(2k+s-1)\binom{2k+s-2}{s}\frac{(2k-2)!}{2^{k-1}(k-1)!}$$

partial pairings obtained in this way may arise as $\mathcal{L}$. Similarly, for $\mathcal{R}$ there are

$$\binom{2n-2k-s}{s}\frac{(2n-2k-2s)!}{2^{n-k-s}(n-k-s)!}$$

possibilities. Any possible $\mathcal{L}$ may be combined with any possible $\mathcal{R}$ to form $\mathcal{P}$ by pairing off the unpaired elements of $\mathcal{L}$ with those of $\mathcal{R}$ in any of $s!$ ways. Multiplying together, and dividing by the total number $(2n)!/(2^n n!)$ of n-pairings we see that for $1 \le k \le n$ and $0 \le s \le n-k$,

$$\mathbb{P}(D_k - 2k = s) = \frac{(2k+s-1)!(2n-2k-s)!n!2^{s+1}}{s!(k-1)!(n-k-s)!(2n)!}. \tag{2}$$

From the expression above it is easy to deduce a concentration result for D_k. For k with $1 \le k \le n$ let us write $p_s = p_{s,k}$ for the probability above, and let

$$r_s = \frac{p_{s+1}}{p_s} = 2\frac{(2k+s)(n-k-s)}{(s+1)(2n-2k-s)}.$$

Note that r_s is a decreasing function of s. Allowing s to be a real number for the moment, the unique positive solution to $r_s = 1$ is given by

$$s = -2k + \sqrt{4kn - 2n + 1/4} + 1/2.$$

Thus $s_0 = \lceil -2k + \sqrt{4kn-2n+1/4} + 1/2 \rceil$ is one of the at most two most likely values of $D_k - 2k$. Also, for n larger than some constant we have

$$\begin{aligned}
\frac{r_{s+1}}{r_s} &= \left(1 - \frac{2k-1}{(s+2)(2k+s)}\right)\left(1 - \frac{n-k}{(2n-2k-s-1)(n-k-s)}\right)\\
&\le \left(1 - \frac{2k-1}{2n^2}\right)\left(1 - \frac{n-k}{2n^2}\right)\\
&\le \exp\left(-\frac{2k-1}{2n^2}\right)\exp\left(-\frac{n-k}{2n^2}\right) \le \exp\left(-\frac{1}{2n}\right).
\end{aligned}$$

As $r_{s_0} \le 1$ it follows that $r_{s_0+x} \le \exp(-x/(2n))$ for $x > 0$, and hence that $p_{s_0+x} \le \exp(-x(x-1)/(4n))$. A similar bound on p_{s_0-x} shows that

$$\mathbb{P}\left(|D_k - (2k+s_0)| \ge 3\sqrt{n\log n}\right) = o(n^{-1}).$$

In fact, as $|s_0 - (2\sqrt{kn} - 2k)| \le 2\sqrt{n}$ for each k, we obtain

$$\mathbb{P}\left(|D_k - 2\sqrt{kn}| \ge 4\sqrt{n\log n}\right) = o(n^{-1}). \tag{3}$$

We now turn to the probability that $d_{k+1} = d+1$, i.e., that the indegree of v_{k+1} is d, given D_k. Suppose that $1 \le k \le n-1$ and $0 \le s \le n-k$, and consider a left partial pairing $\mathcal{L}$ as above. We have already seen that each such $\mathcal{L}$ has

$$s!\binom{2n-2k-s}{s}(2n-2k-2s-1)!!$$

extensions to an n-pairing. Such an extension corresponds to a graph with $d_{k+1} = d+1$ if and only if $2k+s+d+1$ is a right endpoint, and each of $2k+s+1, \ldots, 2k+s+d$ is a left endpoint. Noting that the element paired with $2k+s+d+1$ must be either one of the s unpaired elements in $\mathcal{L}$ or one of $2k+s+1, \ldots, 2k+s+d$, and that $s-1+d$ pairs start before $2k+s+d+1$ and end after this point, each $\mathcal{L}$ has exactly

$$(s+d)(s+d-1)!\binom{2n-2k-s-d-1}{s+d-1}(2n-2k-2s-2d-1)!!$$

such extensions, and for $0 \le d \le n-k-s$ we have, writing $(a)_b$ for $a!/(a-b)!$,

$$\mathbb{P}(d_{k+1} = d+1 \mid D_k - 2k = s) = (s+d)2^d \frac{(n-k-s)_d}{(2n-2k-s)_{d+1}}. \tag{4}$$

It is easy to see that (4) also applies when $k = s = 0$, when we obtain $\mathbb{P}(d_1 = d+1) = d2^d(n)_d/(2n)_{d+1}$. For $k \ge 1$ we can of course combine (2) and (4) to give a rather unilluminating expression for $\mathbb{P}(d_{k+1} = d+1)$. Instead, we shall use (3) and (4) to estimate the expectation of $\#_1^n(d)$, the number of vertices of G_1^n with indegree d. Above and in what follows the functions implied by $o(.)$ or $\sim$ notation are to be interpreted as depending on n only, not on d or k. Also, the constant implied by $O(.)$ notation is absolute.

Let $M = \lfloor n^{4/5}/\log n \rfloor$, let $k = k(n)$ be any function satisfying $M \le k \le n-M$, and let $d = d(n)$ be any function satisfying $0 \le d \le n^{1/15}$. For any D with $|D - 2\sqrt{kn}| \le 4\sqrt{n\log n}$ we can use (4) to write $\mathbb{P}(d_{k+1} = d+1 \mid D_k = D)$ as

$$(2\sqrt{kn} - 2k + O(\sqrt{n\log n}))2^d \frac{(n+k-2\sqrt{kn} + O(\sqrt{n\log n}))^d}{(2n - 2\sqrt{kn} + O(\sqrt{n\log n}))^{d+1}}.$$

Using the bounds on d and k we find that the ratio of $n+k-2\sqrt{kn} = (\sqrt{n}-\sqrt{k})^2$ to $d\sqrt{n\log n}$ tends to infinity as $n \to \infty$, as does $(2n-2\sqrt{kn})/(d\sqrt{n\log n})$. Also, $\sqrt{n\log n} = o(2\sqrt{kn} - 2k)$, so the probability above is equal to

$$(1+o(1))\frac{2\sqrt{kn}-2k}{2n-2\sqrt{kn}}\left(\frac{2(\sqrt{n}-\sqrt{k})^2}{2(n-\sqrt{kn})}\right)^d \sim \sqrt{\kappa}(1-\sqrt{\kappa})^d,$$

where $\kappa = k/n$. As this estimate applies uniformly to $\mathbb{P}(d_{k+1} = d+1 \mid D_k = D)$ for all D with $|D - 2\sqrt{kn}| \le 4\sqrt{n\log n}$, we see from (3) that

$$\mathbb{P}(d_{k+1} = d+1) = o(n^{-1}) + (1+o(1))\sqrt{\kappa}(1-\sqrt{\kappa})^d.$$

In particular, although it is not relevant for the proof, we note that for almost every vertex the most likely indegree is zero.

Keeping n and d fixed and varying k in the range $M \le k \le n-M$, as the estimate above is uniform in k we find that the expected number of vertices v_{k+1}, $M \le k \le n-M$, with degree equal to $d+1$ can be written as

$$o(1) + \sum_{k=M}^{n-M} (1+o(1))\sqrt{k/n}(1-\sqrt{k/n})^d.$$

Thus, as all terms in the sum are positive, we have

$$\mathbb{E}(\#_1^n(d)) = O(M) + o(1) + (1+o(1)) \sum_{k=M}^{n-M} \sqrt{k/n}(1-\sqrt{k/n})^d. \qquad (5)$$

Writing $f = \sqrt{\kappa}(1-\sqrt{\kappa})^d$, we have

$$\frac{1}{f}\frac{\mathrm{d}f}{\mathrm{d}\kappa} = \frac{\kappa^{-1}}{2} - \frac{d}{2}\frac{\kappa^{-1/2}}{1-\kappa^{1/2}}.$$

Provided $n\kappa$ and $n(1-\kappa)$ tend to infinity, the proportional change in f as κ changes by $1/n$ is thus $o(1)$ uniformly in κ. It follows that the sum in (5) can be written as

$$(1+o(1))n \int_{(M+1)/n}^{1-M/n} \sqrt{\kappa}(1-\sqrt{\kappa})^d \mathrm{d}\kappa \sim n \int_0^1 \sqrt{\kappa}(1-\sqrt{\kappa})^d \mathrm{d}\kappa.$$

It is easy to evaluate this integral by substituting $\kappa = (1-u)^2$, and we obtain

$$\mathbb{E}(\#_1^n(d)) = O(M) + (1+o(1))\frac{4n}{(d+1)(d+2)(d+3)} \sim \frac{4n}{(d+1)(d+2)(d+3)},$$

which is the required form of the distribution.

At this point, let us return to the general case $m \geq 1$. Suppose that m is a constant fixed once and for all, and let d_k' be the degree of v_k in the graph G_m^n. We shall estimate $\mathbb{P}(d_{k+1}' = d+m)$, keeping the notation d_K for degrees in the graph G_1^{mn} from which G_m^n is obtained. For the estimate we look at the distributions of $d_{K+1}, \ldots, d_{K+m}$ in G_1^N, where $K = mk$ and $N = mn$. The argument giving the conditional probability estimate (4) actually applies to the conditional probability given the entire sequence of earlier degrees. For $M \leq k \leq n-M$ and $d \leq n^{1/15}$ our earlier estimates show that, provided no $|D_{K'} - 2\sqrt{K'N}|$ is too large,

$$\begin{aligned}\mathbb{P}(d_{K+j+1} = d+1 \mid d_1, d_2, \ldots, d_{K+j}) &\sim \sqrt{(K+j)/N}\left(1-\sqrt{(K+j)/N}\right)^d \\ &\sim \sqrt{\kappa}(1-\sqrt{\kappa})^d,\end{aligned}$$

with $\kappa = k/n = K/N$. Thus, using (3),

$$\begin{aligned}\mathbb{P}(d_{k+1}' = d+m) &= o(n^{-1}) + (1+o(1)) \sum_{a_1+\cdots+a_m=d} \prod_{j=1}^m \sqrt{\kappa}(1-\sqrt{\kappa})^{a_j} \\ &= o(n^{-1}) + (1+o(1))\binom{d+m-1}{m-1}\kappa^{m/2}(1-\sqrt{\kappa})^d.\end{aligned}$$

Proceeding as before we can express the expectation of the number $\#_m^n(d)$ of vertices of G_m^n with indegree d in terms of

$$\int_0^1 \kappa^{m/2}(1-\sqrt{\kappa})^d \mathrm{d}\kappa = 2\int_0^1 (1-u)^{m+1}u^d \mathrm{d}u = 2\frac{(m+1)!d!}{(d+m+2)!},$$

where we have again substituted $\kappa = (1-u)^2$. We find that for $0 \le d \le n^{1/15}$,

$$\mathbb{E}(\#_m^n(d)) \sim \frac{2m(m+1)n}{(d+m)(d+m+1)(d+m+2)}, \tag{6}$$

uniformly in d. We are now ready to prove Theorem 1.

Proof of Theorem 1. We return to considering the graph G_m^n as one graph from the process $(G_m^t)_{t\ge 0}$. Fix $m \ge 1$, $n \ge 1$ and $0 \le d \le n^{1/15}$, and consider the martingale $X_t = \mathbb{E}(\#_m^n(d) \mid G_m^t)$ for $0 \le t \le n$. We have $X_n = \#_m^n(d)$, while $X_0 = \mathbb{E}(\#_m^n(d))$. We claim that the differences $|X_{t+1} - X_t|$ are bounded by two. To see this note that whether at stage t we join v_t to v_i or v_j does not affect the degrees at later times of vertices v_k, $k \notin \{i,j\}$. More precisely, the joint distribution of all other degrees is the same in either case. Since we are just counting vertices with a particular degree, no matter how much the degrees of v_i and v_j are changed in G_m^n, this changes $\#_m^n(d)$ by at most two.

An alternative way of seeing this is to say that at stage t we add a half edge h_{2t-1} directed from $v_{\lceil t/m\rceil}$ paired with a half edge h_{2t} directed to some other vertex, and to consider h_{2t} not as attached to a random vertex, but rather as associated with equal probability to any of $h_1, \ldots, h_{2t-1}$. In the final graph a half edge h_{2t} is attached to $v_{\lceil t/m\rceil}$, while a half edge h_{2t-1} is attached to the vertex the half edge it is associated to is attached to. If we change the choice made at stage t, the effect on the final graph is to move the half edge h_{2t-1} and all later half edges associated directly or indirectly to h_{2t-1} together. This operation only affects two degrees.

Applying Lemma 2, the Azuma-Hoeffding inequality, we find that for each d with $0 \le d \le n^{1/15}$ we have

$$\mathbb{P}\left(|\#_m^n(d) - \mathbb{E}(\#_m^n(d))| \ge \sqrt{n \log n}\right) \le e^{-\log n/8} = o(n^{-1/15}).$$

Noting from (6) that in this range $\mathbb{E}(\#_m^n(d)) \sim \frac{2m(m+1)n}{(d+m)(d+m+1)(d+m+2)}$ and that this is much larger than $\sqrt{n \log n}$, the result follows. $\square$

It is natural to ask how far Theorem 1 can be extended to degrees $d > n^{1/15}$. The bound $d \le n^{1/15}$ was chosen to make the proof as simple as possible, and can certainly be weakened considerably, by choosing a suitable cutoff and considering 'early' and 'late' vertices separately. For large d, equation (6) suggests that the expected number of vertices with degree at least d should be roughly $m(m+1)n/d^2$, and hence that the maximum degree should be $\Theta(\sqrt{n})$. It turns out that this is indeed the case, as could be proved using, for example, the analysis of the pairing model given in [6].

4 Uniform attachment

In [2, 3] it is stated that the preferential attachment assumption of the model is needed to obtain a power-law degree distribution; experimental and heuristic results are given suggesting that with uniform attachment the degrees with be geometrically distributed. It is easy to prove a precise result for this case along the lines of Theorem 1. As the argument is similar to but much simpler than that given above, we only give a rough outline.

Consider a random process in which vertices are added one at a time, starting from any given finite graph G. Suppose that when the vertex v_i is added, it is joined to m earlier vertices, in such a way that the expected number of edges from v_i to v_k is the same for all $k < i$. (It does not matter whether we allow multiple edges or not.) Then the expected indegree of v_k when n vertices have been added is exactly $m \sum_{i=k+1}^{n} \frac{1}{i}$. For $\kappa = k/n$ bounded away from 0 and 1, it is easy to see the degree of v_k has asymptotically a Poisson distribution with mean $\lambda \sim -m \log(\kappa)$. Thus, arguing as before, for any fixed d the proportion of vertices with indegree d is asymptotically

$$\int_0^1 \frac{\lambda^d}{d!} e^{-\lambda} \mathrm{d}\kappa = \frac{m^d}{d!} \int_0^1 (-\log \kappa)^d \kappa^m \mathrm{d}\kappa = \frac{m^d}{(m+1)^{d+1}},$$

giving the expected geometric distribution.

5 Concluding remarks

It is presumably possible to prove a weaker version of Theorem 1 using the following continuous model, which is much more precise than the 'mean-field theory' of [3].

Consider a vertex v born at time t_0 uniformly distributed in $[0, 1]$. When born it has weight (degree) m. If at time t the vertex has weight i then it gets a 'hit' in the infinitesimal time interval $[t, t + \mathrm{d}t]$ with probability $m \frac{i}{2mtn} n \mathrm{d}t = \frac{i}{2t} \mathrm{d}t$. If it does get a hit its weight is incremented by one. The connection is that if vertices are born at time intervals of $1/n$ then at time t the sum of all degrees is $2mnt$, and in an interval of length $\mathrm{d}t$ there are $n\mathrm{d}t$ vertices born, each having m chances to send an edge to v. The differential equations that arise can easily be solved explicitly; one finds that, conditional on t_0, the probability that v has (total) degree $i \geq m$ at time $t > t_0$ is given by

$$\binom{i-1}{m-1} \rho^{m/2} (1 - \sqrt{\rho})^{i-m},$$

where $\rho = t_0/t$. At time 1, writing κ for t_0 and d for $i - m$ as before, this reduces to the expression

$$\binom{d+m-1}{m-1} \kappa^{m/2} (1 - \sqrt{\kappa})^d$$

seen in section 3. For constant d, integrating over κ as before suggests the bounds on $\#_m^n(d)$ given in Theorem 1.

Recall that when defining the process G_1^n above, it was convenient to start from the 'graph' with no vertices, or from the graph with one vertex and no loops. As far as our results are concerned, however, it makes no difference where we start. Given any finite graph G one can define a similar process to G_1^n, or G_m^n, starting from G. Now the joint distribution of the degrees of $v_{s+1}, \dots, v_n$ in G_m^n is independent of G_m^s. If G has s edges then, speaking loosely, for $m = 1$ we can just 'pretend' that G has s vertices as well. As far as all new vertices are concerned, the process starting from G is then indistinguishable from the process $(G_1^t)_{t>s}$, so asymptotic results such as Theorem 1 are unaffected by the starting graph G.

We finish by comparing the model considered here with one much older random process and two new ones. Graphs in which each vertex (apart from the first few) is joined to a fixed number m of randomly chosen earlier vertices are known in the literature as *random recursive dags*, or *random recursive trees* if $m = 1$ (see, e.g., [7]). For $m > 1$ only uniform random recursive dags have been studied significantly. For $m = 1$, however, nonuniform random recursive trees with attachment probabilities proportional to the degrees (also known as random plane-oriented recursive trees) have been studied; see [18, 16, 7], for example. These objects are very close to the random graph G_1^m considered here. The only differences are that loops are not allowed, and that the root vertex is sometimes treated in a slightly different way. The expected number of vertices of degree $d = d(n)$ in these objects was found to within an additive constant by Szymański [18]; a concentration result for d fixed was given by Lu and Feng [15]. For a survey of results on random recursive trees see [17].

Finally, a rather different model for the world-wide web graph was introduced in [12]. Again, vertices are born one at a time, but instead of preferential attachment, each new vertex picks an old vertex to copy from, and copies a randomly selected part of its neighbourhood, as well as choosing (uniformly) new neighbours of its own. In [13, 14] it is shown that such models also give power law degree distributions, as well as explaining the high number of dense bipartite subgraphs found in the web graph.

References

[1] K. Azuma, Weighted sums of certain dependent variables, *Tôhoku Math. J.* **3** (1967), 357–367.

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

[3] A.-L. Barabási, R. Albert and H. Jeong, Mean-field theory for scale-free random networks, *Physica A* **272** (1999), 173-187.

[4] A.-L. Barabási, R. Albert and H. Jeong, Scale-free characteristics of random networks: the topology of the world-wide web, *Physica A* **281** (2000), 69–77.

[5] B. Bollobás, Martingales, isoperimetric inequalities and random graphs. In *Combinatorics* (Eger, 1987), 113–139, Colloq. Math. Soc. János Bolyai, 52, North-Holland, Amsterdam 1988.

[6] B. Bollobás and O.M. Riordan, The diameter of a scale-free random graph, submitted for publication.

[7] L. Devroye and J. Lu, The strong convergence of maximal degrees in uniform random recursive trees and dags, *Random Structures and Algorithms* **7** (1995), 1–14.

[8] P. Erdős and A. Rényi, On random graphs. I, *Publ. Math. Debrecen* **6** (1959), 290–297.

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

[10] E.N. Gilbert, Random graphs, *Ann. Math. Statist.* **30** (1959), 1141–1144.

[11] W. Hoeffding, Probability inequalities for sums of bounded random variables, *J. Amer. Statist. Assoc.* **58** (1963), 13–30.

[12] J. Kleinberg, R. Kumar, P. Raghavan, S. Rajagopalan and A. Tomkins, The web as a graph: measurements, models, and methods, COCOON 1999.

[13] R. Kumar, P. Raghavan, S. Rajagopalan and A. Tomkins, Extracting large scale knowledge bases from the web, VLDB 1999.

[14] R. Kumar, P. Raghavan, S. Rajagopalan, D. Sivakumar, A. Tomkins and E. Upfal, Stochastic models for the web graph, FOCS 2000.

[15] J. Lu and Q. Feng, Strong consistency of the number of vertices of given degrees in nonuniform random recursive trees, *Yokohama Math. J.* **45** (1998), 61–69.

[16] H.M. Mahmoud, R.T. Smythe and J. Szymański, On the structure of random plane-oriented recursive trees and their branches, *Random Structures and Algorithms* **4** (1993), 151–176.

[17] H.M. Mahmoud and R.T. Smythe, A survey of recursive trees, *Th. of Probability and Math. Statistics* **51** (1995), 1–27.

[18] J. Szymański, On a nonuniform random recursive tree, *Annals of Discrete Math.* **33** (1987), 297–306.