A geometric preferential attachment model of networks

Abraham D. Flaxman, Alan M. Frieze*, Juan Vera
Department of Mathematical Sciences,
Carnegie Mellon University,
Pittsburgh PA15213,
U.S.A.

December 6, 2006

Abstract

We study a random graph G_n that combines certain aspects of geometric random graphs and preferential attachment graphs. The vertices of G_n are n sequentially generated points $x_1, x_2, \ldots, x_n$ chosen uniformly at random from the unit sphere in $\boldsymbol{R}^3$. After generating x_t, we randomly connect it to m points from those points in $x_1, x_2, \ldots, x_{t-1}$ which are within distance r. Neighbors are chosen with probability proportional to their current degree and a parameter α biasses the choice towards self loops. We show that if m is sufficiently large, if $r \geq \ln n / n^{1/2-\beta}$ for some constant β, and if $\alpha > 2$, then **whp** at time n the number of vertices of degree k follows a power law with exponent $\alpha + 1$. Unlike the preferential attachment graph, this geometric preferential attachment graph has small separators, similar to experimental observations of [8]. We further show that if $m \geq K \ln n$, K sufficiently large, then G_n is connected and has diameter $O(\ln n / r)$ **whp**.

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 [9], Hayes [23], Watts [34], or Aiello, Chung and Lu [3]. One approach is to model these networks by random graphs. Experimental studies by Albert, Barabási, and Jeong [4], Broder et al [13], and Faloutsos, Faloutsos, and Faloutsos [21] 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 constants C, α. The classical models of random graphs introduced by Erdős and Renyi [19] 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.

*Supported in part by NSF grant CCR-0200945.

One approach is to generate 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 [1]. Mihail and Papadimitriou also use this model [29] in their study of large eigenvalues, as do Chung, Lu, and Vu in [15].

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 [31]. 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. It dates back to Yule [35] and Simon [33]. It was proposed as a random graph model for the web by Barabási and Albert [5], and their description was elaborated by Bollobás and Riordan [10] who showed that at time n, **whp** the diameter of a graph constructed in this way is asymptotic to $\frac{\ln n}{\ln \ln n}$. Subsequently, Bollobás, Riordan, Spencer and Tusnády [12] proved that the degree sequence of such graphs does follow a power law distribution.

The random graph defined in the previous paragraph has good expansion properties. For example, Mihail, Papadimitriou and Saberi [30] showed that **whp** the preferential attachment model has conductance bounded below by a constant. On the other hand, Blandford, Blelloch and Kash [8] found that some WWW related graphs have smaller separators than what would be expected in random graphs with the same average degree. The aim of this paper is to describe a random graph model which has *both* a power-law degree distribution and which has small separators.

We study here the following process which generates a sequence of graphs $G_t, t = 1, 2, \ldots, n$. The graph $G_t = (V_t, E_t)$ has t vertices and mt edges. Here V_t is a subset of S, the surface of the sphere in $\boldsymbol{R}^3$ of radius $\frac{1}{2\sqrt{\pi}}$ (so that $area(S) = 1$).

For $u \in S$ and $r > 0$ we let $B_r(u)$ denote the spherical cap of radius r around u in S. More precisely, $B_r(u) = \{x \in S : ||x - u|| \leq r\}$.

1.1 The random process

The parameters of the process are $m > 0$ the number of edges added in every step and $\alpha \geq 0$ a measure of the bias towards self loops.

Notice that there exists a constant c_0 such that for any $u \in S$, we have

$$A_r = Area(B_r(u)) \sim c_0 r^2.$$

- **Time step** 0**:** To initialize the process, we start with G_0 being the Empty Graph.
- **Time step** $t+1$**:** We choose vertex x_{t+1} uniformly at random in S and add it to G_t. Let $V_t(x_t) = V_t \cap B_r(x_{t+1})$ and let $D_t(x_t) = \sum_{v \in V_t(x_t)} \deg_t(v)$. We add m random edges (x_{t+1}, y_i), $i = 1, 2, \ldots, m$ incident with x_{t+1}. Here, each y_i is chosen independently from $V_t(x_t) \cup \{x_{t+1}\}$ (parallel edges and loops are permitted), such that for each $i = 1, \ldots, m$, for all $v \in V_t(x_{t+1})$,

$$\mathbf{Pr}(y_i = v) = \frac{\deg_t(v)}{\max\left(D_t(x_{t+1}), \alpha m A_r t\right)}$$

and

$$\mathbf{Pr}(y_i = x_{t+1}) = 1 - \frac{D_t(x_{t+1})}{\max\left(D_t(x_{t+1}), \alpha m A_r t\right)}$$

(When $t = 0$ we have $\mathbf{Pr}(y_i = x_1) = 1$.)

Let $d_k(t)$ denote the number of vertices of degree k at time t and let $\overline{d}_k(t)$ denote the expectation of $d_k(t)$.
We will prove the following:

Theorem 1

(a) *If $0 < \beta < 1/2$ and $\alpha > 2$ are constants and $r \sim n^{\beta-1/2}\ln n$ and m is a sufficiently large constant then there exist constants $c, \gamma, \epsilon > 0$ such that for all $k = k(n) \geq m$,*

$$\overline{d}_k(n) = C_k \frac{n}{k^{1+\alpha}} + O(n^{1-\gamma}) \tag{1}$$

where $C_k = C_k(m, \alpha)$ tends to a constant $C_\infty(m, \alpha)$ as $k \to \infty$.

Furthermore, for n sufficiently large, the random variable $d_k(n)$ satisfies the following concentration inequality:

$$\mathbf{Pr}(|d_k(n) - \overline{d}_k(n)| \geq n^{1-\gamma}) \leq e^{-n^{\epsilon}}. \tag{2}$$

(b) *If $\alpha \geq 0$ and $r = o(1)$ then* **whp** *V_n can be partitioned into $T, \bar{T}$ such that $|T|, |\bar{T}| \sim n/2$, and there are at most $4\sqrt{\pi}rnm$ edges between T and $\bar{T}$.*

(c) *If If $\alpha \geq 0$ and $r \geq n^{-1/2}\ln n$ and $m \geq K\ln n$ and K is sufficiently large then* **whp** *G_n is connected.*

(d) *If If $\alpha \geq 0$ and $r \geq n^{-1/2}\ln n$ and $m \geq K\ln n$ and K is sufficiently large then* **whp** *G_n has diameter $O(\ln n/r)$.*

We note that geometric models of trees with power laws have been considered in [20], [6] and [7]. We also note that Gómez-Gardeñes and Moreno [22] have empirically analyzed a one dimensional version of our model when $\alpha = 0$ and their experiments suggest that this yields a power-law exponent of 3.

1.2 Open Questions

In an earlier version of the paper there was no α and we have failed to produce a proof of Theorem 1(a) when $\alpha \leq 2$. This remains a challenge for us at the present moment. We do not think that the $\ln n$ factors are necessary in parts (c),(d).

1.3 Some definitions

Given $U \subseteq S$ and $u \in S$, we define

$$V_t(U) = V_t \cap U \qquad \text{and} \qquad V_t(u) = V_t(B_r(u))$$

and

$$D_t(U) = \sum_{v \in V_t(U)} \deg_t(v) \qquad \text{and} \qquad D_t(u) = D_t(B_r(u)).$$

Given $v \in V_t$, we also define

$$\deg_t^-(v) = \deg_t(v) - m. \tag{3}$$

Notice that $\deg_t^-(v)$ is the number of edges of G_t that are incident to v and were added by vertices that chose v as a neighbor, including loops at v.
Given $U \subseteq S$, let $D_t^-(U) = \sum_{v \in V_t(U)} \deg_t^-(v)$. We also define $D_t^-(u) = D_t^-(B_r(u))$.
Notice that $D_t(U) = m|V_t(U)| + D_t^-(U)$.
We localize some of our notation: given $U \subseteq S$ and $u \in S$ we define $d_k(t, U)$ to be the number of vertices of degree k at time t in U and $d_k(t, u) = d_k(t, B_r(u))$.

2 Outline of the paper

In Section 3 we show that there are small separators. This is easy, since any give great circle can **whp** be used to define a small separator.
We prove a likely power law for the degree sequence in Section 4. We follow a standard practise and prove a recurrence for the expected number of vertices of degree k at time step t. Unfortunatley, this involves the estimation of the expectation of the reciprocal of a random variable and to handle this, we show that this random variable is concentrated. This is quite technical and is done in Section 4.3.
Section 5 proves connectivity when m grows logarithmically with n. The idea is to show that **whp** the sub-graph $G_n(B)$ induced by a ball B of radius $r/2$, center $u \in S$, is connected. This is done by constructing a connected subgraph of $G_n(B)$ via a coupling argument. We then show that the union of the $G_n(B)$'s for $u = x_1, x_2, \ldots, x_n$ is connected and has small diameter.

3 Small separators

Theorem 1(b) is the easiest part to prove. We use the geometry of the instance to obtain a sparse cut. Consider partitioning the vertices using a great circle of S. This will divide V into sets T and $\bar{T}$ which each contain about $n/2$ vertices. More precisely, we have

$$\mathbf{Pr}\left[|T| < (1-\epsilon)n/2\right] = \mathbf{Pr}\left[|\bar{T}| < (1-\epsilon)n/2\right] \leq e^{-\epsilon^2 n/4}.$$

Edges only appear between vertices within distance r, so only vertices appearing in the strip within distance r of the great circle can appear in the cut. Since $r = o(1)$, this strip has area less than $3r\sqrt{\pi}$, and, letting U denote the vertices appearing in this strip, we have

$$\mathbf{Pr}\left[|U| \geq 4\sqrt{\pi}rn\right] \leq e^{-\sqrt{\pi}rn/9}.$$

Even if every one of the vertices chooses its m neighbors on the opposite side of the cut, this will yield at most $4\sqrt{\pi}rnm$ edges **whp**. So the graph has a cut with $\frac{e(T,\bar{T})}{|T||\bar{T}|} \leq \frac{17\sqrt{\pi}rm}{n}$ with probability at least $1 - e^{-\Omega(rn)}$.

4 Proving a power law

4.1 Establishing a recurrence for $\overline{d}_k(t)$: the expected number of vertices of degree k at time t

Our approach to proving Theorem 1(a) is to find a recurrence for $\overline{d}_k(t)$.
We define $\overline{d}_{m-1}(t) = 0$ for all integers t with $t > 0$. Let $\eta_k(G_t, x_{t+1})$ denote the (conditional) probability that a parallel edge to a vertex of degree no more than k is created. Then,

$$\eta_k(G_t, x_{t+1}) = O\left(\sum_{i=m}^{k} \frac{d_i(t, x_{t+1})\, i^2}{\max\{\alpha m A_r t, D_t(x_{t+1})\}^2}\right)$$
$$= O\left(\min\left\{\frac{k^2}{\max\{\alpha m A_r t, D_t(x_{t+1})\}}, 1\right\}\right). \quad (4)$$

Then for $k \geq m$,

$$\mathbf{E}\left[d_k(t+1) \mid G_t, x_{t+1}\right] = d_k(t) + m d_{k-1}(t, x_{t+1}) \frac{k-1}{\max\{\alpha m A_r t, D_t(x_{t+1})\}} - m d_k(t, x_{t+1}) \frac{k}{\max\{\alpha m A_r t, D_t(x_{t+1})\}}$$
$$+ \mathbf{Pr}\left[deg_{t+1}(x_{t+1} = k) \mid G_t, x_{t+1}\right] + O(m\eta_k(G_t, x_{t+1})). \quad (5)$$

Let $\mathcal{A}_t$ be the event

$$\{|D_t(x_{t+1}) - 2mA_r t| \leq C_1 A_r m t^{\gamma} \ln n\}$$

where

$$\max\{2/\alpha, 1/2, 1-2\beta\} < \gamma < 1$$

and C_1 is some sufficiently large constant.
Note that if

$$t \geq (\ln n)^{2/(1-\gamma)}$$

then

$$\mathcal{A}_t \text{ implies } D_t(x_{t+1}) \leq \alpha m A_r t.$$

Then, because $\mathbf{E}[d_k(t,x_{t+1})] \le k^{-1}\mathbf{E}[m|V_t(B_{2r}(x_{t+1}))|] \le k^{-1}m(4A_r t)$ and $d_k(t,x_{t+1}) \le k^{-1}D_t(x_{t+1}) < mt$, we have for $t \ge (\ln n)^{2/(1-\gamma)}$,

$$\begin{aligned}
&\mathbf{E}\left[\frac{d_k(t,x_{t+1})}{\max\{\alpha m A_r t, D_t(x_{t+1})\}}\right] \\
&\quad = \mathbf{E}\left[\frac{d_k(t,x_{t+1})}{\max\{\alpha m A_r t, D_t(x_{t+1})\}} \;\middle|\; \mathcal{A}_t\right]\mathbf{Pr}\left[\mathcal{A}_t\right] + \\
&\qquad\qquad + \mathbf{E}\left[\frac{d_k(t,x_{t+1})}{\max\{\alpha m A_r t, D_t(x_{t+1})\}} \;\middle|\; \neg\mathcal{A}_t\right]\mathbf{Pr}\left[\neg\mathcal{A}_t\right] \\
&\quad = \frac{\mathbf{E}\left[d_k(t,x_{t+1}) \mid \mathcal{A}_t\right]}{\alpha m A_r t}\mathbf{Pr}\left[\mathcal{A}_t\right] + \mathbf{E}\left[O\left(\frac{d_k(t,x_{t+1})}{D_t(x_{t+1})}\right) \;\middle|\; \neg\mathcal{A}_t\right]\mathbf{Pr}\left[\neg\mathcal{A}_t\right] \\
&\quad = \frac{\mathbf{E}\left[d_k(t,x_{t+1}) \mid \mathcal{A}_t\right]}{\alpha m A_r t}\mathbf{Pr}\left[\mathcal{A}_t\right] + O\left(\frac{\mathbf{Pr}\left[\neg\mathcal{A}_t\right]}{k}\right) \\
&\quad = \frac{\mathbf{E}\left[d_k(t,x_{t+1})\right]}{\alpha m A_r t} + \left(O\left(\frac{1}{k}\right) - \frac{\mathbf{E}\left[d_k(t,x_{t+1}) \mid \neg\mathcal{A}_t\right]}{\alpha m A_r t}\right)\mathbf{Pr}\left[\neg\mathcal{A}_t\right] \\
&\quad = \frac{\mathbf{E}\left[d_k(t,x_{t+1})\right]}{\alpha m A_r t} + O\left(\frac{1}{k} + \frac{1}{A_r}\right)\mathbf{Pr}\left[\neg\mathcal{A}_t\right].
\end{aligned}$$

In Lemmas 1 and 3 below we prove that

$$\mathbf{E}\left[d_k(t,x_{t+1})\right] = mA_r\overline{d}_k(t)$$

and that

$$\mathbf{Pr}\left[\neg\mathcal{A}_t\right] = O\left(n^{-2}\right). \tag{6}$$

Thus, if $t \ge (\ln n)^{2/(1-\gamma)}$ then

$$\mathbf{E}\left[\frac{d_k(t,x_{t+1})}{\max\{\alpha m A_r t, D_t(x_{t+1})\}}\right] = \frac{\overline{d}_k(t)}{\alpha m t} + O\left(\frac{1}{n^2}\left(\frac{1}{A_r} + \frac{1}{k}\right)\right). \tag{7}$$

In a similar way

$$\mathbf{E}\left[\frac{d_{k-1}(t,x_{t+1})}{\max\{\alpha m A_r t, D_t(x_{t+1})\}}\right] = \frac{\overline{d}_{k-1}(t)}{\alpha m t} + O\left(\frac{1}{n^2}\left(\frac{1}{A_r} + \frac{1}{k}\right)\right). \tag{8}$$

On the other hand, given G_t, x_{t+1}, if

$$p = 1 - \frac{D_t(x_{t+1})}{\max\left(D_t(x_{t+1}), \alpha m A_r t\right)}$$

then

$$\mathbf{Pr}\left[deg_{t+1}(x_{t+1} = k) \mid G_t, x_{t+1}\right] = \mathbf{Pr}\left[\mathrm{Bi}(m,p) = k-m\right]$$

So, if $t \ge (\ln n)^{2/(1-\gamma)}$,

$$\begin{aligned}
\mathbf{Pr}\left[x_{t+1} = k\right] &= \binom{m}{k-m}\mathbf{E}\left[p^{k-m}(1-p)^{2m-k} \;\middle|\; \mathcal{A}_t\right]\mathbf{Pr}\left[\mathcal{A}_t\right] + O(\mathbf{Pr}\left[\neg\mathcal{A}_t\right]) \\
&= \binom{m}{k-m}\left(1-\frac{2}{\alpha}\right)^{k-m}\left(\frac{2}{\alpha}\right)^{2k-m}(1+O(t^{\gamma-1}\ln n))\mathbf{Pr}\left[\mathcal{A}_t\right] + O(n^{-2}) \\
&= \binom{m}{k-m}\left(1-\frac{2}{\alpha}\right)^{k-m}\left(\frac{2}{\alpha}\right)^{2k-m} + O(t^{\gamma-1}\ln n).
\end{aligned}$$

Now note that from equations (4) and (6) that if

$$t \geq t_0 = n^{(1-2\beta)/\gamma}$$

and

$$k \leq k_0(t) = (mA_r t^\gamma \ln n)^{1/2}$$

then

$$\mathbf{E}(\eta_k(G_t, x_{t+1})) = O(t^{\gamma-1}\ln n). \tag{9}$$

Taking expectations on both sides of (5) and using (7,8,9), we see that if $t \geq t_0$ and $k \leq k_0(t)$ then

$$\begin{aligned} \overline{d}_k(t+1) = \overline{d}_k(t) + \frac{k-1}{\alpha t}\overline{d}_{k-1}(t) - \frac{k}{\alpha t}\overline{d}_k(t) \\ + \binom{m}{k-m}\left(1-\frac{2}{\alpha}\right)^{k-m}\left(\frac{2}{\alpha}\right)^{2m-k} + O\left(t^{\gamma-1}\ln n\right) \end{aligned} \tag{10}$$

We consider the recurrence given by $f_{m-1} = 0$ and for $k \geq m$,

$$f_k = \frac{k-1}{\alpha}f_{k-1} - \frac{k}{\alpha}f_k + \binom{m}{k-m}\left(1-\frac{2}{\alpha}\right)^{k-m}\left(\frac{2}{\alpha}\right)^{2m-k},$$

which, for $k > 2m$, has solution

$$\begin{aligned} f_k &= f_{2m}\prod_{i=m+1}^{k}\frac{i-1}{i+\alpha} \\ &= \phi_k(m,\alpha)\left(\frac{m}{k}\right)^{\alpha+1}, \end{aligned}$$

and has that $\phi_k(m,\alpha)$ tends to a limit $\phi_\infty(m,\alpha)$ depending only on m,α as $k \to \infty$, . We can absorb the values $f_m, f_{m+1}, \ldots, f_{2m}$ into this notation.
We finish the proof of (1) by showing that there exists a constant $M > 0$ such that

$$|\overline{d}_k(t) - f_k t| \leq M(t_0 + t^\gamma \ln n) \tag{11}$$

for all $0 \leq t \leq n$ and $m \leq k \leq k_0(t)$.
This is trivially true for $t < t_0$.
For $k > k_0(t)$ this follows from $\overline{d}_k(t) \leq 2mt/k$.
Let $\Theta_k(t) = \overline{d}_k(t) - f_k t$. Then for $t \geq t_0$ and $m \leq k \leq k_0(t)$,

$$\Theta_k(t+1) = \frac{k-1}{\alpha t}\Theta_{k-1}(t) - \frac{k}{\alpha t}\Theta_k(t) + O(t^{\gamma-1}\ln n). \tag{12}$$

Let L denote the hidden constant in $O(t^{\gamma-1}\ln n)$ of (12). Our inductive hypothesis $\mathcal{H}_t$ is that

$$|\Theta_k(t)| \leq M(t_0 + t^\gamma \ln n)$$

for every $m \leq k \leq k_0(t)$ and M sufficiently large. It is trivially true for $t \leq t_0$. So assume that $t \geq t_0$. Then, from (12),

$$\begin{aligned} |\Theta_k(t+1)| &\leq M(t_0 + t^\gamma \ln n) + Lt^{\gamma-1}\ln n \\ &\leq M(t_0 + (t+1)^\gamma \ln n). \end{aligned}$$

This verifies $\mathcal{H}_{t+1}$ and completes the proof by induction.

4.2 Expected Value of $d_k(t,u)$

Lemma 1 *Let $u \in S$ and let k and t be positive integers. Then $\mathbf{E}[d_k(t,u)] = A_r\overline{d}_k(t)$*

Proof By symmetry, for any $w \in S$, $d_k(t,u)$ has the same distribution as $d_k(t,w)$. Then

$$\begin{aligned}
\mathbf{E}[d_k(t,u)] &= \int_S \mathbf{E}[d_k(t,u)]\,dw = \int_S \mathbf{E}[d_k(t,w)]\,dw \\
&= \mathbf{E}\left[\int_S d_k(t,w)dw\right] = \mathbf{E}\left[\int_S \sum_{v\in V_t} 1_{\deg v=k} 1_{v\in B_r(w)} dw\right] \\
&= \mathbf{E}\left[\sum_{v\in V_t} 1_{\deg v=k} \int_S 1_{w\in B_r(v)} dw\right] = \mathbf{E}\left[\sum_{v\in V_t} 1_{\deg v=k} A_r\right] \\
&= A_r \mathbf{E}[d_k(t)]
\end{aligned}$$

□

Lemma 2 *Let $u \in S$ and $t > 0$ then $\mathbf{E}[D_t(u)] = 2A_r mt$*

Proof

$$\mathbf{E}[D_t(u)] = \sum_{k>0} \mathbf{E}[d_k(t,u)] = A_r \sum_{k>0} \mathbf{E}[d_k(t)] = A_r \mathbf{E}\left[\sum_{k>0} d_k(t)\right] = 2A_r mt$$

□

4.3 Concentration of $D_t(u)$

In this section we prove

Lemma 3 *If $t > 0$ and u is chosen randomly from S then*

$$\mathbf{Pr}\left[|D_t(u) - \mathbf{E}[D_t(u)]| \geq A_r m(t^{2/\alpha} + t^{1/2}\ln t)\ln n\right] = O\left(n^{-2}\right).$$

Proof We think of every edge added as two directed edges. We also think of x_t, the vertex added, as being added with $(\alpha m A_r t - D_t(x_t))^+ = \max\{\alpha m A_r t - D_t(x_t), 0\}$ "phantom" edges pointing to it. Then choosing a vertex is equivalent to choosing one of these directed edges uniformly, and taking the vertex pointed to by this edge as the chosen vertex. So the i-th step of the process is defined by a tuple of random variables $T = (X, Y_1, \ldots, Y_m) \in S \times E_i^m$ where X is the location of the new vertex, a randomly chosen point in S, and Y_j is an edge chosen uniformly at random from among the edges directed into $B_r(X)$ in G_{i-1}. The process G_t is then defined by a sequence $\langle T_1, \ldots, T_t\rangle$, where each $T_i \in S \times E_i^m$.
Let s be a sequence $s = \langle s_1, \ldots, s_t\rangle$ where $s_i = (x_i, y_{(i-1)m+1}, \ldots, y_{im})$ with $x_i \in S$ and $y_j \in E_{\lceil j/m\rceil}$. We say s is *acceptable* if for every j, y_j is an edge entering $B_r(x_{\lceil t/j\rceil})$. Notice that non-acceptable sequences have probability 0 of being realized. Fix $t >$

0. Fix an acceptable sequence $s = \langle s_1, \ldots, s_t \rangle$, and let $A_\tau(s) = \{z \in S \times E_\tau^m : \langle s_1, \ldots, s_{\tau-1}, z\rangle \text{ is acceptable}\}$. For any τ with $1 \le \tau \le t$ and any $z \in A_\tau(s)$ let

$$g_\tau(z) = \mathbf{E}\left[D_t(u) \mid T_1 = s_1, \ldots, T_{\tau-1} = s_{\tau-1}, T_\tau = z\right],$$

let $r_\tau(s) = \sup\{|g_\tau(z) - g_\tau(\hat{z})| : z, \hat{z} \in A_\tau(s)\}$ and let $\hat{r}^2(s) = \sum_{\tau=1}^t (\sup_s r_\tau(s))^2$, where the supremum is taken over all acceptable sequences.
From the Azuma-Hoeffding inequality (see for example [2]) we know that for all $\lambda > 0$,

$$\mathbf{Pr}\left[|D_t(u) - \mathbf{E}\left[D_t(u)\right]| \ge \lambda\right] < 2e^{-\lambda^2/2\hat{r}^2}. \tag{13}$$

Fix τ, with $1 \le \tau \le t$. Our goal now is to bound $r_\tau(s)$ for any acceptable sequence s.
Fix $z, \hat{z} \in A_\tau(s)$. We define $\Omega(G_t, \hat{G}_t)$, a coupling between $G_t = G_t(s_1, \ldots, s_{\tau-1}, z)$ and $\hat{G}_t = G_t(s_1, \ldots, s_{\tau-1}, \hat{z})$

- Step τ: Start with the graph $G_\tau(s_1, \ldots, s_{\tau-1}, z)$ and $\hat{G}_\tau(s_1, \ldots, s_{\tau-1}, \hat{z})$ respectively.

- Step σ ($\sigma > \tau$): Choose the same point $x_\sigma \in S$ in both processes. Let E_σ (resp. $\hat{E}_\sigma$) be the edges pointing to the vertices in $B_r(x_\sigma)$ in $G_{\sigma-1}$ (resp. $\hat{G}_{\sigma-1}$) plus the $(\alpha m A_r \sigma - D_\sigma(x_\sigma))^+$ (resp. $(\alpha m A_r \sigma - \hat{D}_\sigma(x_\sigma))^+$) phantom edges pointing to x_σ. Let $C_\sigma = E_\sigma \cap \hat{E}_\sigma$, $R_\sigma = E_\sigma \setminus \hat{E}_\sigma$, and $L_\sigma = \hat{E}_\sigma \setminus E_\sigma$

  Notice that $|E_\sigma|, |\hat{E}_\sigma| \ge \alpha m A_r \sigma$. Notice also that if $D_\sigma(x_\sigma), D'_\sigma(x_\sigma) \le \alpha m A_r \sigma$, then $|E_\sigma| = |\hat{E}_\sigma|$ and $|R_\sigma| = |L_\sigma|$. Without loss of generality assume that $|E_\sigma| \le |\hat{E}_\sigma|$.

  Now, define $p = 1/|E_\sigma|$ and $\hat{p} = 1/|\hat{E}_\sigma|$. Construct G_σ by choosing m edges uniformly at random $e_1^\sigma, \ldots, e_m^\sigma$ in E_σ, and then joining x_σ to their endpoints, $y_1^\sigma, \ldots, y_m^\sigma$. For each of the m edges $e_i = e_i^\sigma$, we define $\hat{e}_i = \hat{e}_i^\sigma$ by

  - If $e_i \in C_\sigma$ then, with probability $\hat{p}/p$, $\hat{e}_i = e_i$. With probability $1 - \hat{p}/p$, $\hat{e}_i$ is chosen from L_σ uniformly at random.
  - If $e_i \in R_\sigma$, $\hat{e}_i \in L_\sigma$ is chosen uniformly at random.

  Notice that for every $i = 1, \ldots, m$ and every $e \in \hat{E}_\sigma$, $\mathbf{Pr}\left[\hat{e}_i = e\right] = \hat{p}$. To finish, in $\hat{G}_\sigma$ join x_σ to the m vertices pointed to by the edges $\hat{e}_i$.

Now let

$$\Delta_\sigma = \sum_{\rho=\tau}^{\sigma} \sum_{i=1}^{m} 1_{y_i^\sigma \ne \hat{y}_i^\sigma},$$

and for $u \in S$ let

$$\Delta_\sigma(u) = \sum_{\rho=\tau}^{\sigma} \sum_{i=1}^{m} 1_{|\{y_i^\sigma, \hat{y}_i^\sigma\} \cap B_r(u)| = 1}.$$

Lemma 4

$$|g_\tau(z) - g_\tau(\hat{z})| \le \mathbf{E}\left[\Delta_t(u)\right].$$

Proof

$$\begin{aligned}|g_\tau(z) - g_\tau(\hat{z})| &= |\mathbf{E}_{G_t}[D_t(u)] - \mathbf{E}_{\hat{G}_t}[D_t(u)]| \\ &= |\mathbf{E}_{\Omega(G_t, G'_t)}[D_t(u) - D'_t(u)]| \\ &\leq \mathbf{E}_{\Omega(G_t, G'_t)}[\Delta_t(u)]\end{aligned}$$

since only when $|\{y_i^\sigma, \hat{y}_i^\sigma\} \cap B_r(u)| = 1$ do we add ± 1 to the *difference* $D_\rho(u) - D'_\rho(u)$. □

Recall that $A_r = Area(B_r(u)) \sim c_0 n^{2\beta-1}(\ln n)^2$ and we have fixed τ to be an integer with $1 \leq \tau \leq t$.

Lemma 5 *Let $t \geq 1$ and $u \in S$. Then for some constant $C > 0$,*

$$\mathbf{E}[\Delta_t(u)] \leq C m A_r \left(\frac{t}{\tau}\right)^{2/\alpha}.$$

Proof Let $\tau < \sigma \leq t$. We start with

$$\Delta_\sigma = \Delta_{\sigma-1} + \sum_{i=1}^m 1_{y_i^\sigma \neq \hat{y}_i^\sigma}. \tag{14}$$

Now fix $G_{\sigma-1}, \hat{G}_{\sigma-1}$ and x_σ and i. Then taking expectations with respect to our coupling,

$$\begin{aligned}\mathbf{E}\left[1_{y_i^\sigma \neq \hat{y}_i^\sigma}\right] = \mathbf{Pr}(y_i^\sigma \neq \hat{y}_i^\sigma) = \mathbf{Pr}(e_i^\sigma \neq \hat{e}_i^\sigma) = \\ 1 - \frac{|C_\sigma|}{|E_\sigma|}\frac{\hat{p}}{p} = 1 - \frac{|C_\sigma|}{|\hat{E}_\sigma|} = \frac{|L_\sigma|}{|\hat{E}_\sigma|} = \frac{\max\{|L_\sigma|, |R_\sigma|\}}{\max\{|E_\sigma|, |\hat{E}_\sigma|\}} \leq \frac{|L_\sigma| + |R_\sigma|}{\alpha m A_r \sigma}\end{aligned} \tag{15}$$

Therefore

$$\mathbf{E}\left[\Delta_\sigma \,\middle|\, G_{\sigma-1}, \hat{G}_{\sigma-1}, x_\sigma\right] \leq \Delta_{\sigma-1} + m\frac{|L_\sigma| + |R_\sigma|}{\alpha m A_r \sigma} \tag{16}$$

For each $e \in E(\hat{G}_{\sigma-1}) \setminus E(G_{\sigma-1})$, $e \in L_\sigma$ implies x_σ is in the ball of radius r centered at the end point of e. Similarly for $e \in R_\sigma$. Therefore,

$$\mathbf{E}\left[|L_\sigma| + |R_\sigma| \mid G_{\sigma-1}, \hat{G}_{\sigma-1}\right] \leq 2A_r\Delta_{\sigma-1}. \tag{17}$$

Then,

$$\mathbf{E}[\Delta_\sigma] \leq \mathbf{E}[\Delta_{\sigma-1}] + m\frac{\mathbf{E}[|L_\sigma| + |R_\sigma|]}{\alpha m A_r \sigma} \leq \mathbf{E}[\Delta_{\sigma-1}] + \frac{2\mathbf{E}[\Delta_{\sigma-1}]}{\alpha\sigma} = \mathbf{E}[\Delta_{\sigma-1}]\left(1 + \frac{2}{\alpha\sigma}\right),$$

so, $\mathbf{E}[\Delta_t] \leq e^{10/\alpha^2}\left(\frac{t}{\tau}\right)^{2/\alpha}\mathbf{E}[\Delta_\tau]$. Now, $\Delta_\tau \leq m$, because the graphs G_τ and $\hat{G}_\tau$ differ at most in the last m edges. Therefore $\mathbf{E}[\Delta_t] \leq m e^{10/\alpha^2}\left(\frac{t}{\tau}\right)^{2/\alpha}$.
Finally, note that if v is a random point in S then $\mathbf{E}[\Delta_t(v)] = A_r\mathbf{E}[\Delta_t]$. For this, fix u and let ϕ denote a random rotation of S. Let $v = \phi(u)$ and then run Process 1 with $\phi(G_\tau), \phi(\hat{G}_\tau)$ and $x_\sigma, \sigma > \tau$ and then consider Process 2 starting with $G_\tau, \hat{G}_\tau$ and

$\phi^{-1}(x_\sigma), \sigma > \tau$. The mapping ϕ^{-1} does not disturb the distribution of $x_\sigma, \sigma > \tau$ and therefore $\Delta_t(u)$ in Process 2 is equal to $\Delta_t(v)$ in Process 1. □

By applying Lemma 5, we have that for any acceptable sequence

$$R^2(s) = \sum_{\tau=1}^{t} r_\tau(s)^2 \leq (CmA_r)^2 t^{4/\alpha} \sum_{\tau=1}^{t} \tau^{-4/\alpha} = O\left(A_r^2 m^2 (t \ln t + t^{4/\alpha})\right)$$

Therefore, by using Equation (13), we have that there is C_1 such that

$$\mathbf{Pr}\left[|D_t(u) - \mathbf{E}\left[D_t(u)\right]| \geq C_1 A_r m (t^{2/\alpha} + t^{1/2} \ln t)(\ln n)^{1/2}\right] \leq e^{-2\ln n} = n^{-2}.$$

4.4 Concentration of $d_k(t)$

We follow the proof of Lemma 3, replacing $D_t(u)$ by $d_k(t)$ and using the same coupling. When we reach Lemma 4 we find that $|g_\tau(z) - g_\tau(\hat{z})| \leq 2\mathbf{E}[\widehat{D}_t]$ (each edge disrepancy can affect two vertices), the rest is the same.
This proves (1) and completes the proof of Theorem 1(a) .

5 Connectivity

Here we are going to prove that for $r \geq n^{-1/2} \ln n$, $m > K \ln n$, and K sufficiently large, **whp** G_n is connected and has diameter $O(\ln n/r)$. Notice that G_n is a subgraph of the graph $G(n,r)$, the intersection graph of the caps $B_r(x_t)$, $t = 1, 2, \ldots, n$ and therefore it is disconnected for $r = o((n^{-1} \ln n)^{1/2})$ [32]. We denote the diameter of G by $\mathrm{diam}(G)$, and follow the convention of defining $\mathrm{diam}(G) = \infty$, when G is disconnected. In particular, when we say that a graph has finite diameter this implies it is connected.
Let

$$T = K_1 \ln n / A_r = O(n/\ln n)$$

where K_1 is sufficiently large, and $K_1 \ll K$.

Lemma 6 *Let $u \in S$ and let $B = B_{r/2}(u)$. Then*

$$\mathbf{Pr}\left[\mathrm{diam}(G_n(B)) \geq 2(K_1 + 1) \ln n\right] = O(n^{-3})$$

where $G_n(B)$ is the induced subgraph of G_n in B.

Proof
Given τ_0 and N, we consider the following process which generates a sequence of graphs $H_s = (W_s, F_s)$, $s = 1, 2, \ldots, N$. (The meanings of N, τ_0 will become apparant soon).
Time step 1
To initialize the process, we start with H_1 consisting of τ_0 isolated vertices $y_1, \ldots, y_{\tau_0}$.
Time step $s \geq 1$: We add vertex $y_{s+\tau_0}$. We then add $\frac{m}{8000(\alpha+1)^2}$ random edges incident with $y_{s+\tau_0}$ of the form $(y_{s+\tau_0}, w_i)$ for $i = 1, 2, \ldots, \frac{m}{8000(\alpha+1)^2}$. Here each w_i is chosen uniformly from W_s.

The idea is to couple the construction of G_n with the construction of H_N for $N \sim \text{Bi}(n-T, A_r/4)$ and $\tau_0 = \text{Bi}(T, A_r/4)$ such that **whp** H_N is a subgraph of G_n with vertex set $V_n(B)$. We are then going to show that **whp** $\text{diam}(H_N) \le 2(K_1+1)\ln n$, and therefore $\text{diam}(G_n(B)) \le 2(K_1+1)\ln n$.
To do the coupling we use two counters, t for the steps in G_n and s for the steps in H_N:

- Given G_{τ_0}, set $s = 0$. Let $W_0 = V_T(B)$. Notice that $\tau_0 = |W_0| \sim \text{Bi}(T, A_r/4)$ and that $\tau_0 \le K_1 \ln n$ **whp**.
- For every $t > T$.
  - If $x_t \notin B$, do nothing in H_s.
  - If $x_t \in B$, set $s := s+1$. Set $y_{s+\tau_0} = x_t$. As we want H_N to be a subgraph of G_n we must choose the neighbors of $y_{s+\tau_0}$ among the neighbors of x_t in G_n. Let A be the set of vertices chosen by x_t in $V_t(B)$. Notice that $|A|$ stochastically dominates $a_t \sim \text{Bi}\left(m, \frac{D_t(B)}{\max\{\alpha m A_r t, D_t(x_t)\}}\right)$. If $\frac{D_t(B)}{\max\{\alpha m A_r t, D_t(x_t)\}} \ge \frac{1}{50(\alpha+1)}$, then a_t stochastically dominates $b_t \sim \text{Bi}(m, \frac{1}{50\alpha})$ and so **whp** is at least $\frac{m}{100(\alpha+1)}$. If $\frac{D_t(B)}{\max\{\alpha m A_r t, D_t(x_t)\}} < \frac{1}{50(\alpha+1)}$ we declare failure, but as we see below this is unlikely to happen.
    For any $R > 0$,
    $$m|V_t(B_R(w))| \le D_t(B_R(w)) = m|V_t(B_R(w))| + D_t^-(B_R(w)) \le 2m|V_t(B_{R+r}(w))|. \quad (18)$$
    where $D_t^-(B_R(w))$ is the sum over vertices $x \in B_R(w)$ of the of the in-degree $\deg_t(x) - m$ of x.
    Now $|V_t(B_R(w)| \sim \text{Bi}(t, (R/r)^2 A_r)$ and so
    $$\mathbf{Pr}(D_t(x_t) \ge 8mA_r t \text{ OR } D_t(B) \notin [mA_r t/5, 3mA_r t] \text{ OR } |V_t(B)| < A_r t/5) \le n^{-K_1/100.} \quad (19)$$
    So we assume that G_t is such that the event described in (19) does not happen. Thus each vertex of B has probability at least $\frac{m}{8(\alpha+1)mA_r t} \ge \frac{1}{40(\alpha+1)|V_t(B)|}$ of being chosen under preferential attachment.
    Thus, as insightfully observed by Bollobás and Riordan [11] we can legitimately start the addition of x_t in G_t by choosing $\frac{m}{8000(\alpha+1)^2}$ random neighbours uniformly in B.

Notice that N, the number of times s is increased, is the number of steps for which $x_t \in B$, and so $N \sim \text{Bi}(n-T, A_r/4)$.
Now we are ready to show that H_N is connected **whp**.
By Chernoff's bound we have that
$$\mathbf{Pr}\left[\left|\tau_0 - \frac{K_1}{4}\ln n\right| \ge \frac{K_1}{8}\ln n\right] \le 2n^{-K_1/48}$$

and

$$\mathbf{Pr}\left[N \leq \frac{1}{3}(\ln n)^2\right] \leq e^{-c(\ln n)^2}$$

for some $c > 0$. Therefore, we can assume $\ln n \leq \tau_0 \leq K_1 \ln n$ and $N \geq \frac{1}{3}(\ln n)^2$.
Let X_s be the number of connected components of H_s. Then

$$X_{s+1} = X_s - Y_s, \qquad X_0 = \tau_0$$

where $Y_s \geq 0$ is the number of components (minus one) collapsed into one by $y_{s+\tau_0}$. So

$$\mathbf{Pr}\left[Y_s = 0 \mid H_s\right] \leq \sum_{i=1}^{X_s} \left(\frac{c_i}{s+\tau_0}\right)^{m/8000(\alpha+1)^2}$$

where the c_i are the component sizes of H_s. If $s < 2K_1 \ln n$ then because $m \geq K \ln n$, we have

$$\mathbf{Pr}\left[Y_s = 0 \mid X_s \geq 2\right] \leq 2\left(1 - \frac{1}{s+\tau_0}\right)^{m/8000(\alpha+1)^2} \leq 2e^{-m/(8000(\alpha+1)^2(s+\tau_0))} \leq 1/10.$$

So X_s is stochastically dominated by the random variable $\max\{1, \tau_0 - Z_s\}$ where $Z_s \sim \mathrm{Bi}(s, 9/10)$. We then have

$$\mathbf{Pr}\left[X_{2K_1 \ln n} > 1\right] \leq \mathbf{Pr}\left[Z_{2K_1 \ln n} < \tau_0\right] \leq \mathbf{Pr}\left[Z_{2K_1 \ln n} < K_1 \ln n\right] \leq n^{-3}.$$

And therefore

$$\mathbf{Pr}\left[H_{2K_1 \ln n} \text{ is not connected}\right] \leq n^{-3}.$$

Now, to obtain an upper bound on the diameter, we run the process of construction of H_N by rounds. The first round consists of $2K_1 \ln n$ steps and in each new round we double the size of the graph, i.e. it consists of as many steps as the total number of steps of all the previous rounds. Notice that we have less than $\ln n$ rounds in total. Let $\mathcal{A}$ be the event for all $i > 0$ every vertex created in the $(i+1)^{th}$ round is adjacent to a vertex in $H_{2^{i-1}K_1 \ln n}$, the graph at the end of the i^{th} round.
On the event $\mathcal{A}$, every vertex in H_N is at distance at most $\ln n$ of $H_{2K_1 \ln n}$ whose diameter is not greater than $2K_1 \ln n$. Thus, the diameter of H_N is smaller than $2(K_1 + 1)\ln n$.
Now, we have that if v is created in the $(i+1)^{th}$ round,

$$\mathbf{Pr}\left[v \text{ is not adjacent to } H_{2^{i-1}K_1 \ln n}\right] \leq \left(\frac{1}{2}\right)^m.$$

Therefore

$$\mathbf{Pr}\left[\neg\mathcal{A}\right] \leq \left(\frac{1}{2}\right)^m n(\ln n) \leq \frac{\ln n}{n^{K \ln 2 - 1}}.$$

□

To finish the proof of connectivity and the diameter, let u, v be two vertices of G_n. Let $C_1, C_2, \ldots, C_M$, $M = O(1/r)$ be a sequence of spherical caps of radius $r/4$ such that u is the center of C_1, v is the center of C_M and such that the centers of C_i, C_{i+1} are distance $\leq r/2$ apart. The intersections of C_i, C_{i+1} have area at least $A_r/40$ and so **whp** each

intersection contains a vertex. Using Lemma 6 we deduce that **whp** there is a path from u to v in G_n of size at most $O(\ln n/r)$.
Acknowledgement We thank Olivier Riordan for detailed comments which pointed to a major error in our proof in earlier version of this paper. We also thank Zeng Jianyang for his comments.

References

[1] W. Aiello, F. R. K. Chung, and L. Lu, A random graph model for massive graphs, *Proc. of the 32nd Annual ACM Symposium on the Theory of Computing*, (2000) 171–180.

[2] N. Alon and J. Spencer, The Probabilistic Method, Second Edition, Wiley-Interscience 2000.

[3] W. Aiello, F. R. K. Chung, and L. Lu, Random Evolution in Massive Graphs, *Proc. of IEEE Symposium on Foundations of Computer Science*, (2001) 510–519.

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

[5] A. Barabasi and R. Albert, Emergence of scaling in random networks, *Science* 286 (1999) 509–512.

[6] N. Berger, B. Bollobas, C. Borgs, J. Chayes, and O. Riordan, Degree distribution of the FKP network model, *Proc. of the 30th International Colloquium of Automata, Languages and Programming*, (2003) 725–738.

[7] N. Berger, C. Borgs, J. Chayes, R. D'Souza, and R. D. Kleinberg, Competition-induced preferential attachment, Proceedings of the 31st International Colloquium on Automata, Languages and Programming (ICALP), 208-221, Lecture Notes in Computer Science 3142 (2004).

[8] D. Blandford, G. E. Blelloch, and I. Kash, Compact Representations of Separable Graphs, *Proc. of ACM/SIAM Symposium on Discrete Algorithms* (2003) 679–688.

[9] B. Bollobás and O. Riordan, Mathematical Results on Scale-free Random Graphs, in *Handbook of Graphs and Networks*, Wiley-VCH, Berlin, 2002.

[10] B. Bollobás and O. Riordan, The diameter of a scale-free random graph, *Combinatorica*, 4 (2004) 5–34.

[11] B. Bollobás and O. Riordan, Coupling scale free and classical random graphs, *Internet Mathematics* 1 (2004), no. 2, 215–225.

[12] 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.

[13] A. Broder, R. Kumar, F. Maghoul, P. Raghavan, S. Rajagopalan, R. Stata, A. Tomkins, and J. Wiener, Graph structure in the web, *Proc. of the 9th Intl. World Wide Web Conference* (2002) 309–320.

[14] G. Buckley and D. Osthus, Popularity based random graph models leading to a scale-free degree distribution, *Discrete Mathematics* 282 (2004) 53–68.

[15] F.R.K. Chung, L. Lu, and V. Vu, Eigenvalues of random power law graphs, *Annals of Combinatorics* 7 (2003) 21–33.

[16] F.R.K. Chung, L. Lu, and V. Vu, The spectra of random graphs with expected degrees, *Proceedings of national Academy of Sciences* 100 (2003) 6313–6318.

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

[18] E. Drinea, M. Enachescu, and M. Mitzenmacher, *Variations on Random Graph Models for the Web*, Harvard Technical Report TR-06-01 (2001).

[19] P. Erdős and A. Rényi, On random graphs I, *Publicationes Mathematicae Debrecen* 6 (1959) 290–297.

[20] A. Fabrikant, E. Koutsoupias, and C. H. Papadimitriou, Heuristically Optimized Trade-Offs: A New Paradigm for Power Laws in the Internet, *Proc. of 29th International Colloquium of Automata, Languages and Programming* (2002) .

[21] M. Faloutsos, P. Faloutsos, and C. Faloutsos, On Power-law Relationships of the Internet Topology, *ACM SIGCOMM Computer Communication Review* 29 (1999) 251–262.

[22] J. Gómez-Gardeñes and Y. Moreno, Local versus global knowledge in the Barabási-Albert scale-free network model, *Physical Review E* 69 (2004) 037103.

[23] B. Hayes, Graph theory in practice: Part II, *American Scientist* 88 (2000) 104-109.

[24] J. M. Kleinberg, R. Kumar, P. Raghavan, S. Rajagopalan, and A. S. Tomkins, The Web as a Graph: Measurements, Models and Methods, *Proc. of the 5th Annual Intl. Conf. on Combinatorics and Computing (COCOON)* (1999).

[25] R. Kumar, P. Raghavan, S. Rajagopalan, D. Sivakumar, A. Tomkins, and E. Upfal, Stochastic Models for the Web Graph, *Proc. IEEE Symposium on Foundations of Computer Science* (2000) 57.

[26] R. Kumar, P. Raghavan, S. Rajagopalan, D. Sivakumar, A. Tomkins, and E. Upfal, The Web as a Graph, *Proc. 19th ACM SIGACT-SIGMOD-AIGART Symp. Principles of Database Systems (PODS)* (2000) 1–10.

[27] R. Kumar, P. Raghavan, S. Rajagopalan, and A. Tomkins, Trawling the Web for emerging cyber-communities, *Computer Networks* 31 (1999) 1481–1493.

[28] C. J. H. McDiarmid, Concentration, in *Probabilistic methods in algorithmic discrete mathematics*, (1998) 195-248.

[29] M. Mihail and C. H. Papadimitriou, On the Eigenvalue Power Law, *Proc. of the 6th International Workshop on Randomization and Approximation Techniques* (2002) 254–262.

[30] M. Mihail, C. H. Papadimitriou, and A. Saberi, On Certain Connectivity Properties of the Internet Topology, *Proc. IEEE Symposium on Foundations of Computer Science* (2003) 28.

[31] M. Mitzenmacher, A brief history of generative models for power law and lognormal distributions, *Internet Mathematics* 1 (2004), no. 2, 226–251.

[32] M. D. Penrose, *Random Geometric Graphs*, Oxford University Press (2003).

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

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

[35] 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.