Internet Mathematics Vol. 1, No. 1: 1-35

Robustness and Vulnerability of Scale-Free Random Graphs

Béla Bollobás and Oliver Riordan

Abstract. Recently many new "scale-free" random graph models have been introduced, motivated by the power-law degree sequences observed in many large-scale, real-world networks. Perhaps the best known, the Barabási-Albert model, has been extensively studied from heuristic and experimental points of view. Here we consider mathematically two basic characteristics of a precise version of this model, the LCD model, namely robustness to random damage, and vulnerability to malicious attack. We show that the LCD graph is much more robust than classical random graphs with the same number of edges, but also more vulnerable to attack. In particular, if vertices of the n-vertex LCD graph are deleted at random, then as long as any positive proportion remains, the graph induced on the remaining vertices has a component of order n. In contrast, if the deleted vertices are chosen maliciously, a constant fraction less then 1 can be deleted to destroy all large components. For the Barabási-Albert model, these questions have been studied experimentally and heuristically by several groups.

1. Introduction

Recently there has been much interest in developing and studying new random graph models that capture certain observed common features of many large-scale, real-world networks. Although the recent activity in this area perhaps started with the "small-world" model of Watts and Strogatz [Watts and Strogatz 98], the main focus now seems to be on so called "scale-free" random graphs, whose

1542-7951/03 $0.50 per page

degree distributions follow power laws. Their introduction was motivated by observations of Faloutsos, Faloutsos, and Faloutsos [Faloutsos et al. 99] and many other groups, who noticed power laws in the degree sequences and other properties of many real-world graphs, including the internet and web graphs. Such power laws had been described much earlier in several contexts (see Lotka [Lotka 26], Simon [Simon 55], and Zipf [Zipf 49], for example), without leading to the same kind of explosion in activity.

Two approaches to scale-free models have been proposed. The first, introduced by Aiello, Chung, and Lu [Aiello et al. 01], takes the degree sequence as a given; the study of such models then tells us what other properties of the network follow from its scale-free nature. The second approach seeks to explain the degree distribution in terms of simple underlying principles. One of the first, and perhaps the most studied, of the models in this vein is the Barabási-Albert model described in [Barabási and Albert 99].

The Barabási-Albert model has two key features. The first is that the graph grows one vertex at a time, each new vertex sending a fixed number m of edges to older vertices. The second is the principle of "preferential attachment"; at each step the probability that a certain old vertex is chosen is proportional to its degree at the time. Together these two features give a power-law degree sequence. (Other mechanisms, such as "copying" have also been considered; see [Kumar et al. 00], for example.)

Although the Barabási-Albert model has been much studied, most of the work on it is of a heuristic or experimental rather than mathematical nature. In particular, it is not often pointed out that the definition given by Barabási and Albert is incomplete, and, strictly speaking, does not make sense. (See [Bollobás and Riordan to appear] for further details.) A precisely defined model, the *linearized chord diagram* or LCD model, was introduced in [Bollobás and Riordan to appear], motivated by the vague description of Barabási and Albert, and incorporating its key features as well as other useful mathematical properties. (See Section 2.)

For heuristic and experimental studies of the Barabási-Albert model, we refer the reader to the extensive surveys [Albert and Barabási 02] and [Dorogovtsev and Mendes 02]; these references also describe many generalizations, and contain much background material on the by now rather large subject of scale-free random graphs. In contrast, so far there has been rather little rigorous mathematical work; what there is sometimes confirms and sometimes contradicts the heuristic results. See [Aiello et al. 01, Cooper and Frieze 02, Bollobás et al. 01, Buckley and Osthus to appear, Cooper and Frieze 01] and [Cooper and Frieze 02] for some examples, or the survey [Bollobás and Riordan 02].

The properties we consider here are among the most basic properties of real-

world networks, namely *robustness*, i.e., resistance to random damage, and *vulnerability*, i.e., vulnerability to malicious attack. These were considered experimentally in [Albert et al. 00] and heuristically in [Callaway et al. 00, Cohen et al. 00, Cohen et al. 01]. Writing n for the total number of vertices of the graph, we ask when the damaged graph contains a *giant component*, i.e., a component whose order is $\Theta(n)$ as $n \to \infty$. In particular, we measure robustness or vulnerability by asking what fraction p_c of vertices must remain in order to have a giant component, when vertices are deleted at random or so as to cause the most damage. This kind of measure is usual in random graphs, and corresponds to the critical probability in percolation.

We shall show that the LCD graph is much more robust but also more vulnerable than classical random graphs with the same number of edges. Under malicious attack, the critical proportion p_c of vertices needed for a giant component is roughly 4 times as high in the LCD graph as in classical random graphs. In the other direction, and much more strikingly, the LCD graph is in some sense "infinitely" robust, unlike classical models. More precisely, the critical probability is zero. In other words, for any constant $p > 0$, after randomly deleting all but pn vertices of an LCD graph with n vertices a giant component of order roughly $\lambda(p)n$ remains, where $\lambda(p)$ is a positive constant. However, we shall show that $\lambda(p)$ tends to zero very quickly as $p \to 0$, so in practice for p not very much smaller than the threshold for classical models, the "giant" component might as well not exist. We also investigate the reason for these results, showing that the main difference comes from the preferential attachment rather than the "growing" nature of the graph.

The threshold question considered by Aiello, Chung, and Lu [Aiello et al. 01] is related but rather different: Instead of deleting vertices, the power law is varied and the question is which power laws give (undamaged) graphs with giant components.

The rest of the paper is organized as follows. In the next section we give the definition of the LCD model $G_m^{(n)}$ itself, and describe exactly the graphs G_p and G_c derived by random and malicious deletion of vertices respectively. In Section 3 we state our main results, described approximately above. In Section 4 we describe an equivalent formulation of the LCD model also introduced in [Bollobás and Riordan to appear], which gives a large amount of conditional independence. In Section 5 we give the motivation for, and definition of, a certain continuous process that we shall use to approximate the growth of neighbourhoods in G_p and G_c; an exact approximation result for $G_m^{(n)}$ is stated and proved in Section 6. In the next section we modify the process to account for vertex deletion, while also simplifying its definition. Using the new process the proof of the robustness result is completed in Section 8, and that of the vulnerability result in Section 9.

The final section concerns comparisons with rigorous results for classical random graphs and growing graphs with uniform rather than preferential attachment, as well as with heuristic results for scale-free graphs.

2. Models

Throughout the paper, we consider the random graph process $(G_m^{(t)})_{t\geq 0}$ defined by the LCD model of [Bollobás and Riordan to appear], where m is a positive integer describing the number of edges sent out by each vertex. The definition from [Bollobás and Riordan to appear] is as follows:

Consider a fixed sequence of vertices $v_1, v_2, \ldots$. (Most of the time we shall take $v_i = i$ to simplify the notation.) We write $d_G(v)$ for the degree of the vertex v in the graph G. We shall inductively define a random graph process $(G_1^{(t)})_{t\geq 0}$ so that $G_1^{(t)}$ is a graph on $\{v_i : 1 \leq i \leq t\}$, as follows. Start with $G_1^{(0)}$, the empty "graph" with no vertices, or with $G_1^{(1)}$ the graph with one vertex and one loop. Given $G_1^{(t-1)}$, we form $G_1^{(t)}$ by adding the vertex v_t together with a single edge between v_t and v_i, where i is chosen randomly with

$$\Pr(i = s) = \begin{cases} d_{G_1^{(t-1)}}(v_s)/(2t-1) & 1 \leq s \leq t-1, \\ 1/(2t-1) & s = t. \end{cases} \tag{2.1}$$

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 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. Note that the definition allows loops and multiple edges, but there will not be many. The reason for choosing exactly this definition is that it has a very useful equivalent form, described in Section 4.

From now on, we shall take $v_i = i$, so $G_m^{(t)}$ is a graph on $[t] = \{1, 2, \ldots, t\}$. Most of the time we shall consider not the whole process, but the random graph $G_m^{(n)}$ on $[n]$ resulting at some particular time n, which will tend to infinity. Note that $G_m^{(n)}$ is an undirected graph, but has a very natural orientation: Direct each edge ij with $i > j$ from i to j.

Having described the model, we now turn to the measurement of robustness and vulnerability.

For robustness, we consider the graph G_p obtained from $G_m^{(n)}$ by deleting vertices independently of each other, and of the graph itself, keeping each vertex with probability p. Here p will be a constant between 0 and 1, and we consider the following question: For which p does the "damaged graph" G_p have with high probability a *giant component*, i.e., a component of order $\Theta(n)$, as $n \to \infty$ with m fixed?

For vulnerability to attack, one would perhaps like to consider deleting cn vertices from $G_m^{(n)}$ so as to cause the most damage, i.e., so as to minimize the size of the largest component in the graph that remains. However, such an attack seems very hard to analyze precisely. Also, such an attack can in practice not be carried out; the attacker would need to know the complete graph and perform an intractable calculation to decide which vertices to knock out. In $G_m^{(n)}$, an enemy who knows the general properties but not the details of the graph can still direct an attack in a sensible way; earlier vertices tend to have higher degrees, and are likely to be more important in "holding the graph together." The attack we shall consider is to delete the first cn vertices of $G_m^{(n)}$, obtaining a graph G_c, where $0 < c < 1$ is a constant; given the model but not the random realization, it is easy to see that this is the best attack. We are interested in the following question: For which c is the graph G_c likely to have a giant component as $n \to \infty$ with m fixed?

3. Results

Let us write $L_1(G)$ for the order of the largest component of a graph G, and $L_2(G)$ for the order of the second largest component (or 0 if G is connected). In the following results, the graphs $G_p = G_p(m,n)$ and $G_c = G_c(m,n)$ are defined by deleting vertices from $G_m^{(n)}$ as described in the previous section. Note that in line with standard random graphs notation, we *keep* vertices with probability p, whereas some authors delete vertices with probability p. On the other hand, c is a cutoff. Thus as p increases the graph G_p grows, while as c increases G_c shrinks.

Theorem 3.1. *Let $m \geq 2$ be fixed. For $0 < p \leq 1$, there is a function $\lambda(p) > 0$ such that with probability $1-o(1)$ we have $L_1(G_p) = (\lambda(p)+o(1))n$ and $L_2(G_p) = o(n)$.*

Although our main interest is the absence of a critical probability, i.e., the existence of a giant component for any $p > 0$, we shall also obtain bounds on the size of the giant component, showing that as $p \to 0$ we have

$$\exp(-\Theta(1/p^2)) \leq \lambda(p) \leq \exp(-\Theta(1/p)). \tag{3.1}$$

Here the lower bound is what we happen to obtain from our proof of Theorem 3.1; from a theoretical point of view, the main thing is that it is positive. From a more practical point of view, the upper bound is more important; this shows that while for any positive p a theoretical "giant" component exists, for even moderately small p this component might as well not be there for practical purposes. Numerical calculations suggest that the upper bound is the truth.

For G_c, the situation is very different: There is a threshold.

Theorem 3.2. *Let $m \geq 2$ and $0 < c < 1$ be constant. If $c \geq \frac{m-1}{m+1}$, then with probability $1 - o(1)$ we have $L_1(G_c) = o(n)$. If $c < \frac{m-1}{m+1}$, then there is a positive constant $\theta(c)$ such that with probability $1 - o(1)$ we have $L_1(G_c) = (\theta(c) + o(1))n$ and $L_2(G_c) = o(n)$.*

The restriction to $m \geq 2$ is important; let us note the following simple result for $m = 1$, for later comparison with heuristics.

Theorem 3.3. *Set $m = 1$. For any constant $p < 1$, we have $L_1(G_p) = o(n)$. Also, for any constant $c > 0$ we have $L_1(G_c) = o(n)$.*

Proof. When $m = 1$, the graph $G_m^{(n)}$ is a forest with loops; in this case, other variants of the model are more natural, leading to trees. The key observation is that there is at most one path between any pair of vertices. It is easy to check that the length of the path between a random pair of vertices is typically $\Theta(\log n)$; to be more precise, but much more crude, for any constant K the number of pairs of vertices within distance K is $o(n^2)$. If a constant fraction of the vertices is deleted (either at random or from the start), then almost all long paths are destroyed; since there are not many short paths to start with, $o(n^2)$ paths remain and there cannot be a giant component. □

The result above shows that the case $m = 1$ is very different, but also rather uninteresting. For the rest of the paper (except the conclusions), we consider only the case $m \geq 2$.

4. Pairings on [0,1]

One of the most important properties of the precise model introduced in [Bollobás and Riordan to appear] is that the graph $G_m^{(n)}$ has a static (nonrecursive) definition which allows the reintroduction of independence to a great extent. As shown in [Bollobás and Riordan to appear], we can define $G_m^{(n)}$ as a graph on

$[n] = \{1, 2, \ldots, n\}$ as described below. By an $M_2(0,1)$ random variable we mean a $[0,1]$-valued random variable with density function $2x$, $0 < x < 1$, for example, the maximum of two independent uniform $[0,1]$-valued random variables. We ignore, as we may, probability zero events; in particular, we shall not be careful or indeed consistent in how we treat the case of equality between continuous random variables.

Let r_k, $1 \le k \le mn$, be iid $M_2(0,1)$ random variables. Sort these into ascending order to obtain $0 < R_1 < \cdots < R_{mn}$, and let $W_i = R_{mi}$, $1 \le i \le n$. Let $W_0 = 0$ and write w_i for $W_i - W_{i-1}$. The graph $G_m^{(n)}$ may be formed as follows: Given the R_k, take mn independent random variables $L_{i,r}$, $1 \le i \le n$, $1 \le r \le m$, where $L_{i,r}$ is uniform on $R_{m(i-1)+r}$. Each vertex i sends out m edges to vertices $t_{i,r}$, $1 \le r \le m$, where each $t_{i,r}$ is the unique t such that $W_{t-1} < L_{i,r} \le W_t$. As shown in [Bollobás and Riordan to appear], the random graph produced this way has exactly the distribution of $G_m^{(n)}$.

We shall see below that for our purposes it makes no difference if we take $L_{i,r}$ uniform on $[0, W_i]$ rather than $[0, R_{m(i-1)+r}]$; the bounds $W_{i-1} \le R_{i,r} \le W_i$ (or indeed much cruder bounds) are sufficient for the approximations we make. Thus we shall consider the following almost equivalent but slightly simpler description: Starting from the W_i as above, the random variables $t_{i,r}$ describing to which vertices vertex i sends it edges are independent, with

$$\Pr(t_{i,r} = t) = \begin{cases} w_t/W_i & 1 \le t \le i, \\ 0 & \text{otherwise.} \end{cases} \tag{4.1}$$

Note that in the above descriptions, we think of the edges of $G_m^{(n)}$ as directed, but this orientation carries no information as it is always from higher numbered vertices to lower number vertices. Although throughout we study the properties of $G_m^{(n)}$ as an undirected graph, it is sometimes convenient to consider the edges being directed in this manner.

Let us note some key properties of the W_i and w_i we shall use. Here and later we say that an event holds *with very high probability*, or **wvhp**, if it holds with probability $1 - o(n^{-\epsilon})$ as $n \to \infty$, where $\epsilon = 1/1000$. Throughout, we keep m fixed. Given this, all constants in $O(.)$ notation are universal; similarly the implied functions in $o(.)$ notation depend on n only. Most of the time we shall write ϵ instead of $1/1000$, as we shall think of it as a small enough constant whose numerical value is not important.

Note first that $W_i \le x$ if and only if at least mi of the mn variables r_k fall in $[0, x]$. The number of r_k falling in $[0, x]$ has a Binomial $\mathrm{Bi}(mn, x^2)$ distribution, so one can check that **wvhp** we have

$$W_i = \sqrt{i/n}\left(1 + O(n^{-1/4}\log n)\right) \tag{4.2}$$

holding (uniformly) for all $n^{1/2} \le i \le n$.

Secondly, it is easy to describe the distribution of w_i given $W_1, \dots, W_{i-1}$. Indeed, given that $W_{i-1} = y$ (and given also any information about $W_1, \dots, W_{i-2}$), we see that exactly $m(n-i+1)$ of the r_k exceed y. Furthermore, the values $r'_1, \dots, r'_{m(n-i+1)}$ taken by these variables are independent with density $2x/(1-y^2)$, $y < x < 1$, so the conditional distribution of W_i is given by the m^{th} smallest of $m(n-i+1)$ iid random variables with this density. As long as i is not too close to 1 or n and y is such that (4.2) holds, one can see that near y the values of the r'_j resemble a Poisson process with density $m(n-i+1)2y/(1-y^2) \sim 2m\sqrt{in}$, and the distribution of w_i given W_{i-1} is similar to

$$\frac{Z_m}{2m\sqrt{in}}, \tag{4.3}$$

where Z_m is given by the sum of m independent exponential random variables, each with mean 1. We shall not make the meaning of "similar" precise here; a precise statement is made in Section 6. Note that the unconditional distribution of w_i is also similar to that given above.

5. Continuous Approximation for Neighbourhood Growth

In proving Theorems 3.1 and 3.2 the main idea is to consider a random vertex of $G_m^{(n)}$ and look at the early growth of its neighbourhoods, comparing this with a certain continuous random process. Although we are interested in G_p or G_c, where vertices of $G_m^{(n)}$ have been deleted randomly or from the beginning, we shall state the key approximation result for $G_m^{(n)}$ itself; the relevant modifications will be immediate.

Since the result is a little awkward to state, we start with some motivation, considering how the neighbourhoods of a vertex $i \in V(G_m^{(n)}) = [n]$ grow. From (4.1), we see that the distribution of the neighbours of i depends on two parameters. The first is W_i, which itself is essentially determined by i, or, rescaling, by $\alpha(i) = i/n$. The second is w_i, or, rescaling in a natural way, $x(i) = 2m\sqrt{in}w_i$. If we pick a vertex i uniformly at random, then $\alpha(i)$ is essentially uniform on $[0,1]$, while $x(i)$ has essentially the distribution Z_m described above, and is essentially independent of $\alpha(i)$. (Precise statements are postponed until the next section.) In fact we shall condition on the values of all the W_i, and hence the w_i, assuming that they have certain nice properties stated later, including (4.2) and an assumption about the distribution of the w_i; the comment about the distribution of $\alpha(i), x(i)$ still applies, using only the randomness in the choice of i.

The vertex i has two kinds of neighbours: *left*-neighbours $j < i$, to which i

sends an edge in the oriented version of $G_m^{(n)}$, and *right*-neighbours $j > i$. (We ignore loops for the moment.) Considering the r^{th} left-neighbour $j = t_{i,r}$ of i, we see that for each $t \le i$, we have $\Pr(j = t) = w_t/W_i$. Considering only i and j not too close to 1 or n (say in the range $[n^{1/2}, n - n^{1/2}]$) we have from (4.2) that

$$\Pr(j = t) = w_t/W_i \sim \frac{x(t)}{2m\sqrt{it}}. \tag{5.1}$$

We are interested in the chance that the parameters $\alpha(j) = j/n$ and $x(j)$ take certain values, or rather fall in certain ranges. Let us write α for $\alpha(i)$ and x for $x(i)$, and $f_Z(y) = y^{m-1}e^{-y}/(m-1)!$, $y > 0$, for the density function of Z_m. One can check that, roughly speaking, the chance that $\alpha(j)$ lies in $[\beta, \beta + \mathrm{d}\beta]$ and $x(j)$ in $[y, y + \mathrm{d}y]$ is given by

$$n\mathrm{d}\beta f_Z(y)\mathrm{d}y \frac{y}{2m\sqrt{\alpha n \beta n}} = \frac{y f_Z(y)}{2m\sqrt{\alpha\beta}}\mathrm{d}y\mathrm{d}\beta, \tag{5.2}$$

for $\beta < \alpha$ not too small and $y > 0$ not too large. Indeed, there are $n\mathrm{d}\beta$ vertices t with $\alpha(t)$ in the required range. From the distribution assumption on the w_i we shall make, based on (4.3), the number of these with $x(t)$ in the required range is roughly $n\mathrm{d}\beta f_Z(y)\mathrm{d}y$, and multiplying by the probability in (5.1) gives (5.2). Note that the value $x = x(i)$ does not appear in (5.2).

We now turn to the right-neighbours of i, noting that the number of these is itself a random variable. In suitable ranges, the probability that i has a right-neighbour j with $\alpha(j) \in [\beta, \beta + \mathrm{d}\beta]$ and $x(j) \in [y, y + \mathrm{d}y]$ is given by

$$mn\mathrm{d}\beta f_Z(y)\mathrm{d}y \frac{x}{2m\sqrt{\alpha n \beta n}} = \frac{x f_Z(y)}{2\sqrt{\alpha\beta}}\mathrm{d}y\mathrm{d}\beta. \tag{5.3}$$

To see this, note that there are again $n\mathrm{d}\beta f_Z(y)\mathrm{d}y$ candidate vertices j. Each of these sends out m edges. Relation (5.1) shows that each has probability

$$w_i/W_j \sim \frac{x(i)}{2m\sqrt{ij}}$$

of landing at i, giving (5.3). Provided we take sufficiently small intervals, the chance of i having two right-neighbours with these parameters will be negligible.

Note that the degree $d(i)$ of vertex i is just m plus the number of its right-neighbours. It follows that (after conditioning on the Ws) we have

$$\mathbb{E}(d(i)) \sim m + \frac{x(i)}{2\sqrt{i}} \sum_{j=i}^{n} j^{-1/2} \sim m + x(i)(\sqrt{n/i} - 1), \tag{5.4}$$

for $i \geq n^{1/2}$, say. Also, when this expectation is large it is the sum of many small terms corresponding to independent events. It is thus easy to check that, given the Ws, **wvhp** every vertex with $\mathbb{E}(d(i)) \geq 2n^{\epsilon}$ has $d(i) \geq n^{\epsilon}$.

So far we have discussed the neighbours of a single vertex. Of course, as we work outwards finding the neighbourhoods of a random initial vertex, we do not have complete independence. We do, however, have almost complete independence, in a sense we shall now describe.

We use the following standard notation: For a vertex v of a graph G and an integer $k \geq 0$, let $\Gamma_k(v)$ be the set of vertices of G at graph distance exactly k from v. Also, let $N_k(v)$ be the set of vertices at distance at most k. Let the initial (random) vertex be v_0, and set $\Gamma_k = \Gamma_k(v_0)$. Thus $\Gamma_0 = \{v_0\}$, while for $k = 1, 2, \ldots$ the set Γ_k consists of those vertices of $G_m^{(n)} \setminus (\Gamma_0 \cup \cdots \cup \Gamma_{k-1})$ adjacent to some vertex of Γ_{k-1}. We shall also write N_k for $N_k(v_0) = \Gamma_0 \cup \ldots \cup \Gamma_k$. When determining the distribution of Γ_{k+1} given $\Gamma_0, \ldots, \Gamma_k$, we must remember not only which vertices are in Γ_k, and their $x(.)$ values, but also how each vertex $j \in \Gamma_k$ was reached—*from the left*, i.e., as a right-neighbour of some $i \in \Gamma_{k-1}$, or *from the right*, as a left-neighbour. Of course, v_0 is reached in neither way. Also, a vertex might be reached both ways (or one way twice or more), but we shall see later that we can ignore this.

Given $\Gamma_0, \ldots, \Gamma_k$, the corresponding $x(.)$ values and that a vertex $j \in \Gamma_k$ was reached from the right, all we know about its left-neighbours $t_{j,r}$, $1 \leq r \leq m$, is that they are not in N_{k-1}. Provided this set is not too large, $t_{j,r}$ thus has roughly its unconditioned distribution. Since for each $i > j$ the events that i sends an edge to j are independent (given the Ws), except that no vertex of N_{k-1} can be a right-neighbour of j, the set of right-neighbours also has essentially its unconditioned distribution. If j was reached from the left, the situation is different for left-neighbours; this time, we know that some particular left-neighbour of j lies in N_{k-1}. However, the $m-1$ remaining left-neighbours will have essentially their unconditioned distribution.

Motivated by the above, we define (precisely) a statistical process $(\tilde{\Gamma}) = (\tilde{\Gamma}_0, \tilde{\Gamma}_1, \ldots)$ as follows. Each generation $\tilde{\Gamma}_k$, $k \geq 0$, will consist of a finite number of points v each of which has an integer $l(v) \in \{m-1, m\}$ and two real numbers $\alpha(v) \in (0,1)$ and $x(v) > 0$ associated with it. We start with $\tilde{\Gamma}_0$ consisting of a single point v_0 with $\alpha(v_0)$ chosen uniformly from $[0,1]$ and $x(v_0)$ having the distribution Z_m described above. Given $\tilde{\Gamma}_0, \ldots, \tilde{\Gamma}_k$, each point $v \in \tilde{\Gamma}_k$ gives rise independently to children in the next generation as described below. We write α for $\alpha(v)$ and x for $x(v)$. Firstly, v gives rise to exactly $l(v)$ "left-children." Each such left-child w has $l(w) = m$, and the values $\beta = \alpha(w)$, $y = x(w)$ are chosen according to the density (5.2), independently for each left-child. Secondly, v gives rise to a Poisson number of

"right-children" w, each with $l(w) = m - 1$, where the chance that v has such a right-child w with $\beta = \alpha(w)$ and $y = x(w)$ in a certain small interval is given by (5.3).

Putting together the remarks in this section suggests that the first few neighbourhoods Γ_k of a random vertex of $G_m^{(n)}$ should behave like the sets $\tilde{\Gamma}_k$, in the sense that the α and x values are similar. We shall make this precise in the next section.

6. Proof of the Continuous Approximation

In this section, we prove that the neighbourhoods of a random vertex of $G_m^{(n)}$ do grow roughly as given by the continuous process described above. By (Γ), we shall mean the process $(\Gamma_0, \Gamma_1, \ldots)$ defined by first constructing $G_m^{(n)}$, then choosing a random vertex v_0 of this graph, setting $\Gamma_0 = \{v_0\}$ and defining Γ_k as the k^{th} neighbourhood of v_0 as in the previous section. The process $(\tilde{\Gamma}) = (\tilde{\Gamma}_0, \tilde{\Gamma}_1, \ldots)$ is defined at the end of the previous section; note that this latter process does not depend on n.

Theorem 6.1. *For each n, the processes (Γ) and $(\tilde{\Gamma})$ can be coupled so that with probability $1 - o(n^{-1/1000})$ we have*

$$|\Gamma_k| = |\tilde{\Gamma}_k|$$

for $0 \le k \le K$, and either $|\Gamma_K| = |\tilde{\Gamma}_K| = 0$ or

$$\sum_{k=0}^{K+1} |\Gamma_k|, \quad \sum_{k=0}^{K+1} |\tilde{\Gamma}_k| \ge n^{1/1000}.$$

In other words, unless a certain event of small probability holds, as far as size is concerned the neighbourhoods behave the same way in the two models until their total size reaches at least $n^{1/1000}$ or they die out. We have made no attempt to optimize the constant $1/1000$ in the above result, and indeed will write ϵ for this constant throughout the proof, so as not to be distracted by its numerical value.

Proof. Throughout, we shall assume that n is larger than some very large constant n_0. Although the proof we present is complete, we shall not dot all the i's and cross all the t's, omitting straightforward but sometimes tedious details at various points, and giving a somewhat informal presentation where it is clear this can be made formal.

The basic method is to construct the coupling inductively. Essentially, we have to show that, off a certain bad set, the transition probabilities of the two processes are very close. To do this, we consider not only the size of Γ_k or $\tilde{\Gamma}_k$, but also the corresponding values $\alpha(v)$, $x(v)$ for the vertices (points) in these sets, showing that **wvhp** these are "the same" for all k in the indicated range. Of course, they cannot be exactly the same; for example the $\alpha(.)$ values are discrete in one case and continuous in the other.

Let us write ϵ for $1/1000$, which we shall think of just as a small constant. Let $\delta = \delta(n) = n^{-10\epsilon}$. Note that events that occur with probability $1 - O(n^{\epsilon}\delta^{1/2})$ hold **wvhp**. Since when "growing" the neighbourhoods Γ_k from v we do not need to consider more than n^{ϵ} vertices, we may ignore events that hold with probability $O(\delta)$ for each neighbour found. For both processes, we shall quantize the values of $\alpha(v)$ and $x(v)$ within a factor of $(1+\delta)$. We shall also only consider the ranges

$$\alpha \in [\delta^3, 1] \text{ and } x \in [\delta^{1/2}, 20\epsilon \log n]. \tag{6.1}$$

As before, let Z_m be the sum of m iid exponential random variables each with mean 1, so Z_m has probability density function $f_Z(x) = x^{m-1}e^{-x}/(m-1)!$, $x > 0$. Then we have

$$\Pr(Z_m \le \delta^{1/2}) \le (\delta^{1/2})^m/m! \le \delta, \tag{6.2}$$

and

$$\Pr(Z_m \ge 20\epsilon \log n) = n^{-20\epsilon + o(1)} \le \delta. \tag{6.3}$$

Let us write s_0 for the maximal integer with $(1+\delta)^{s_0} < \delta^{1/2}$, and s_1 for the minimal integer with $(1+\delta)^{s_1} > 20\epsilon \log n$. It is easy to check that for $s_0 \le s < s_1$ we have

$$\Pr\left(Z_m \in [(1+\delta)^s, (1+\delta)^{s+1})\right) = \delta x f_Z(x)(1 + O^*(\delta)) \ge n^{-1/4}, \tag{6.4}$$

where $x = (1+\delta)^s$ and the $O^*(.)$ notation hides bounded powers of $\log n$.

We consider first the process (Γ). For $0 \le r \le R$, let us write $c_r = \lfloor (1-\delta)^r n \rfloor$, where R is chosen so that $c_R = \delta^3 n(1 + O(\delta))$, so $R = \Theta(\delta^{-1} \log n)$. Let $I_r = [c_{r+1} + 1, c_r]$, for $0 \le r < R$. Note that from (4.2), we have that **wvhp**

$$W_{c_r} = (1-\delta)^{r/2}(1 + O(\delta^2)) \tag{6.5}$$

for each $r = 0, 1, \ldots, R$. It follows that

$$W_i = (1-\delta)^{r/2}(1 + O(\delta)) \tag{6.6}$$

for every $i \in I_r$, for $0 \le r < R$. It is at this point that we see that in the precise description of the model no harm is done by taking the $L_{i,j}$ uniform on $[0, W_i]$; from now on, the only bound on the range of each $L_{i,j}$ we shall use is (6.6).

For $0 \le r < R$ and $s_0 \le s < s_1$, let $C_{r,s}$ be the number of $i \in I_r$ with $x(i) \in [(1+\delta)^s, (1+\delta)^{s+1})$. We claim that, **wvhp**, for all such r and s we have

$$C_{r,s} = n\delta\alpha f_Z(x)\delta x(1+O^*(\delta)), \tag{6.7}$$

where $\alpha = (1-\delta)^r$ and $x = (1+\delta)^s$. Also, let B_r be the set of $i \in I_r$ for which $x(i) \le \delta^{1/2}$ or $x(i) \ge 20\epsilon \log n$. We also claim that, **wvhp**, for every $0 \le r < R$ we have

$$|B_r| = O(\delta |I_r|) \tag{6.8}$$

and

$$\sum_{i \in B_r} w_i = O^*(\delta(W_{c_r} - W_{c_{r+1}})). \tag{6.9}$$

Note that $W_{c_r} - W_{c_{r+1}}$ is exactly the sum of all w_i for $i \in I_r$.

The claims above are straightforward but tedious to check using (6.2), (6.3), and (6.4); omitting the details, to check (6.7), (6.8), and (6.9) we condition on $W_{c_{r+1}}$ taking some fixed value consistent with (6.5) and consider examining each w_i, $i = c_{r+1}+1, \dots, c_r$, in turn. As indicated in the previous section, until i is very close to n, unless some highly unlikely bad event occurs, given all the previous w_j, each w_i is very close in distribution to $Z_m/(2m\sqrt{in})$, so off a certain bad event the relevant $x(i)$ are well approximated by independent copies of Z_m. (Here the approximation is that the two quantities are within a factor of $(1+\delta)$ when in the range $(\delta^{1/2}, 20\epsilon \log n)$, and that if one is above/below this range, so is the other.) Since the last interval has length about δn which is much larger than $n^{1/2}$, the breakdown of the approximation when i is close to n does not affect (6.7), (6.8), or (6.9).

So far, we have only considered the W_i. From now on, we shall condition on the W_i, assuming only that (6.6)–(6.9) hold. The key point is that, within the ranges indicated, the number of vertices with $\alpha(v)$ and $x(v)$ taking values in certain small ranges is close to what it "should" be.

We now turn to the range limitations, showing that we do not need to consider vertices with α or x values outside these ranges.

Recall that we are considering the neighbourhood process (Γ) in the graph $G_m^{(n)}$. Let us say a vertex $i \in \Gamma_k$ is *bad* if one of the following four conditions holds: (a) $\alpha(i) < \delta^3$, (b) $x(i) \ge 20\epsilon \log n$, (c) $x(i) \le \delta^{1/2}$, or (d) i has more than one neighbour in Γ_{k-1}.

Claim 6.2. *The probability that there is a $k \ge 0$ such that $|N_k| < n^\epsilon$ and Γ_k contains a bad vertex is $O(n^\epsilon \delta^{1/2})$.*

To establish the claim, consider such a bad vertex $i \in \Gamma_k$ with k minimal. Note that we may assume $k > 0$ as it is immediate from (6.8) that the probability that the randomly chosen initial vertex v_0 is bad is $O(\delta)$.

Suppose first that $\alpha(i) < \delta^3$. Let j be a neighbour of i in Γ_{k-1}, and note that $i < j$, so i is a left-neighbour of j. We may assume that $j \le \delta^2 n$ because each left-neighbour of a vertex j with $j > \delta^2 n$ has probability

$$\sum_{i \le \delta^3 n} w_i / W_j = W_{\lfloor \delta^3 n \rfloor} / W_j \le \sqrt{\delta^3/\delta^2}(1 + o(1)) \sim \delta^{1/2}$$

of sending an edge to some $i \le \delta^3 n$. Thus, conditioning on Γ_{k-1}, the probability that some such $j \in \Gamma_{k-1}$ sends an edge to some $i < \delta^3 n$ is $O(|\Gamma_{k-1}|\delta^{1/2})$. Since we only consider k for which $|N_k| < n^\epsilon$, this probability is negligible. We may thus assume that our bad vertex i with $\alpha(i) < \delta^3$ was reached from a $j \in \Gamma_{k-1}$ with $j < \delta^2 n$. But we know that $x(j) \ge \delta^{1/2}$. From (5.4), any vertex j with these properties has expected degree $\Theta(\sqrt{n/j}x(j)) = \Omega(\delta^{-1/2})$. As noted earlier, the probability that there is such a vertex with fewer than n^ϵ neighbours is very small. Hence we may assume that j has at least n^ϵ neighbours, which contradicts our assumption on the size of N_k. This establishes that the probability that a "first" bad vertex i has $\alpha(i) < \delta^3$ is very small.

We may now assume that $\alpha(i) > \delta^3 n$, so $i \in I_r$ for some $0 \le r < R$. It is now immediate from (6.8) and (6.9) that the probability that there is such a first bad vertex satisfying (b) or (c) is very small: Consider choosing a left- or right-neighbour i of a vertex $j \in \Gamma_{k-1}$. We may first decide which of the I_r the vertex i lies in (noting that with probability $1 - O(\delta)$ this is not the interval containing j), and then exactly where in I_r the vertex i is. For a right-neighbour i, each vertex of I_r is roughly equally likely to be chosen (the denominator in (5.1) does not vary much over I_r), and so from (6.8) it is very unlikely that $i \in B_r$, i.e., that i satisfies (b), or (c). For a left-neighbour we use (6.9) instead, because of the dependence of the probability of t being chosen on $x(t)$.

Finally it is very easy to check that the chance of Γ_{k-1} sending at least two edges to a vertex i not satisfying (a), (b), or (c) is very small, establishing Claim 6.2.

We have now done the groundwork necessary to prove Theorem 6.1. We shall actually prove a much stronger statement, that one can couple the processes so that, **wvhp**, until too many vertices/points are involved they are "almost identical," in that corresponding α and x values lie in the same small intervals. Claim 6.2 shows that we can ignore α or x values outside the ranges given by (6.1). (Actually, we have shown this for (Γ). A similar but much simpler argument shows this for $(\tilde{\Gamma})$.)

Note that (Γ) consists of sets of vertices, while $(\tilde{\Gamma})$ consists of sets of "points." We have chosen different words to avoid confusion between the two processes.

For $0 \le r < R$, let $\tilde{I}_r$ be the interval $((1-\delta)^{r+1}, (1-\delta)^r]$, so that for $i \in [n]$ we have $i \in I_r$ if and only if $\alpha(i) = i/n \in \tilde{I}_r$. For $s_0 \le s < s_1$, let J_s be the interval $[(1+\delta)^s, (1+\delta)^{s+1})$. We shall inductively couple (Γ) and $(\tilde{\Gamma})$ so that **wvhp**, until both processes involve more than n^ϵ vertices/points, we have bijections between Γ_k and $\tilde{\Gamma}_k$ preserving the following information: from which vertex/point of the previous generation each vertex/point is reached, whether from the left or the right, which $\tilde{I}_r$ its α value lies in, and in which J_s its x value lies. (We avoid a cumbersome formal statement.) All we have to do is show that the transition probabilities going from generation k to $k+1$ are similar for the two processes. We do this by considering the vertices $v \in \Gamma_k$ one at a time. Throughout, we only consider values within the ranges (6.1).

Let us suppose first that v was reached from the right (or was the initial vertex). Conditioning on the exact values of $\Gamma_0, \dots, \Gamma_k$, as well as the neighbourhoods of the vertices in Γ_k we have already examined, all we know about each $t_{v,a}$ is that it does not lie in $N_{k-1} = \Gamma_0 \cup \dots \cup \Gamma_{k-1}$, a set of at most n^ϵ vertices. From (5.1), it follows that for each $t \notin N_{k-1}$, the conditional probability that $t_{v,a} = t$ is given by

$$\Pr(t_{v,a} = t \mid \dots) = w_t/(W_v - \sum_{j \le v,\ j \in N_{k-1}} w_j).$$

Now from the bound (6.1), for each $j \in N_{k-1}$, we have

$$w_j = x(j)/(2m\sqrt{jn}) \le 10\epsilon \log n/(\delta^{3/2} n) < n^{-2/3}.$$

As $|N_{k-1}| \le n^\epsilon$, while $v \ge \delta^3 n$, so $W_v = \sqrt{v/n}(1 + O(\delta)) \ge n^{-1/3}$, say, we see that the sum above is negligible compared to W_v, so

$$\Pr(t_{v,a} = t \mid \dots) = w_t \sqrt{n/v}(1 + O(\delta)) = \frac{x(t)}{2m\sqrt{vt}}(1 + O(\delta)).$$

Let us suppose that $v \in I_r$ (or equivalently that $\alpha(v) \in \tilde{I}_r$), and that $x(v) \in J_s$. Set $\alpha = (1-\delta)^r$ and $x = (1+\delta)^s$, so $\alpha(v) = \alpha(1+O(\delta))$ and $x(v) = x(1+O(\delta))$. Let us consider $t \in I_{r'}$ with $x(t) \in J_{s'}$, setting $\beta = (1-\delta)^{r'}$ and $y = (1+\delta)^{s'}$. Then we see that

$$\Pr(t_{v,a} = t \mid \dots) = \frac{y}{2mn\sqrt{\alpha\beta}}(1 + O(\delta)).$$

Finally, summing over the $C_{r',s'}$ such points t, we see from (6.7) that

$$\Pr(\alpha(t_{v,a}) \in \tilde{I}_{r'}, x(t_{v,a}) \in J_{s'} \mid \dots) = n\delta\beta f_Z(y)\delta y \frac{y}{2mn\sqrt{\alpha\beta}}(1 + O^*(\delta)).$$

Noting that $\delta\beta$ is the width of $\tilde{I}_{r'}$ and δy that of $J_{s'}$, we see that this is within a factor $1 + O^*(\delta)$ of the corresponding probability for $(\tilde{\Gamma})$, given essentially by (5.2). (The density (5.2) does not change by more than a factor $1 + O^*(\delta)$ over the relevant intervals for α, β, x and y.) Thus we may couple the left-neighbours of v with those of the corresponding point in $\tilde{\Gamma}_k$ as required, with probability $1 - O^*(\delta)$. Note that we do not mind accumulating such error probabilities as we stop after finding n^ϵ neighbours.

The case where v was reached from the left is almost identical, except that this time we already know one of the left-neighbours of v, and consider only the $m - 1$ others. The definition of $\tilde{\Gamma}$ ensures that the corresponding point has only $m - 1$ left-neighbours in the next generation.

For the right-neighbours, there is a slight complication as their number is not fixed, but this is easy to deal with. Let us enumerate as $(r_i, s_i)_{i=1}^N$ the possible pairs (r', s') for which v may have a right-neighbour w with $\alpha(w) \in I_{r'}$, $x(w) \in J_{s'}$. We shall ignore the event that v has two or more neighbours with the same r', s' values; if this has non-negligible probability then the expected number of neighbours of v is very large, and as noted below we are done. Let p_i be the probability that v has a right-neighbour w with $\alpha(w) \in I_{r_i}$, $x(w) \in J_{s_i}$. Then since the events that two vertices send edges to v are independent, for (Γ) the transition probabilities we wish to consider are of the form

$$p_S = \prod_{i \in S} p_i \prod_{i \in [N] \setminus S} (1 - p_i)$$

for $S \subset [N]$. Let p'_i be the probabilities corresponding to p_i but for $(\tilde{\Gamma})$, given by (5.3), up to a factor $1 + O^*(\delta)$, which is how much the density can vary over the relevant intervals. Then by the independence in the definition of $(\tilde{\Gamma})$, the corresponding transition probability is of the form

$$p'_S = \prod_{i \in S} p'_i \prod_{i \in [N] \setminus S} (1 - p'_i).$$

Arguing as for left-neighbours, we see that for each i, we have $p_i = p'_i(1 + O^*(\delta))$. For small p, we have $(1 - p)/(1 - p(1 + \eta)) = 1 + O(p\eta)$, so

$$p_S / p'_S = (1 + O^*(\delta))^{|S|} \exp(O^*(\delta) \sum_{i \in [N]} p_i).$$

Now the first term above may be neglected, as we accumulate a factor $1 + O^*(\delta)$ for each right-neighbour we find, but we may stop once we find n^ϵ. We may neglect the second factor unless $\sum_{i \in [N]} p_i$ is at least $n^{2\epsilon} < \delta^{-1/3}$, say. However, this sum is just the expected number of right-neighbours of v. If this expectation

is at least $n^{2\epsilon}$ then, as noted after (5.4), **wvhp** the actual number of right-neighbours is at least n^{ϵ}. Also, in this case the expected number $\sum_{i\in[N]} p'_i$ of neighbours in $(\tilde{\Gamma})$ is also very large (as $p'_i \sim p_i$), and **wvhp** we have many neighbours in $(\tilde{\Gamma})$ also. This means that we can abandon the construction of the coupling at this point. Put together with the argument for left-neighbours, this completes the inductive proof of the existence of the required coupling, proving Theorem 6.1. □

7. General Analysis of the Continuous Process

In order to deduce results about $G_m^{(n)}$ from Theorem 6.1, we need first to analyze the process $(\tilde{\Gamma})$. In fact, we are interested in the graphs G_p and G_c obtained from $G_m^{(n)}$, respectively, by deleting vertices independently with probability $1-p$ (keeping them with probability p) and by deleting all vertices i with $i \le cn$. To study these, we shall need a corresponding generalization of $(\tilde{\Gamma})$. We shall also note that $(\tilde{\Gamma})$ is essentially "one-dimensional" rather than two-dimensional, in that the variables $x(.)$ can be eliminated. This will be convenient when it comes to analyzing $(\tilde{\Gamma})$. As we shall modify $(\tilde{\Gamma})$ in several ways, we describe the modifications informally separately, and then give a formal definition of the new process.

Throughout this section, we fix $0 \le c < 1$ and $0 < p \le 1$. We shall always take either $c = 0$, $p < 1$, corresponding to G_p, or $c > 0$, $p = 1$, corresponding to G_c. Corresponding to deleting all vertices i of $G_m^{(n)}$ with $i \le cn$, we shall modify the density (5.2) by replacing it with zero when $\beta \le c$. We no longer have a probability density—in the modified process each point now has at most $l(v)$ left-children rather than exactly $l(v)$. Also, we replace $\tilde{\Gamma}_0$ by $\emptyset$ if $\alpha(v_0) \le c$.

Corresponding to deleting vertices of $G_m^{(n)}$ at random with probability $1-p$, we change the process $(\tilde{\Gamma})$ as follows: For each $v \in \tilde{\Gamma}_k$, use the "old" rules to construct a set of *potential children* of v. The actual children of v in $\tilde{\Gamma}_{k+1}$ are obtained by selecting the potential children independently, each with probability p. Before starting, we also replace $\tilde{\Gamma}_0$ by $\emptyset$ with probability $1-p$.

Turning to the final modification, let us note that in $(\tilde{\Gamma})$ the distribution of $x(v)$ for some $v \in \tilde{\Gamma}_k$ is very simple. We condition on $\tilde{\Gamma}_{k-1}$, and on how v was reached (as a left- or right-child). If v was a right-child, then from (5.3) we see that $x(v)$ has the distribution Z_m, with density $f_Z(x)$, independently of $\alpha(v)$. Similarly, if v was a left-child, then from (5.2) $x(v)$ is distributed as the size-biased version of Z_m, with density $xf_Z(x)/m$. We may thus construct a process equivalent to $(\tilde{\Gamma})$ as follows: For each point, we record only $l(v)$, and

hence whether v was reached from the left or the right, except that v_0 must be treated differently, and $\alpha(v)$. The rule for obtaining children is to first pick a value x according to the distribution Z_m or its size-biased version as appropriate, and then construct children according to versions of (5.2) and (5.3) where the y part has been integrated out.

Putting the above together, we define a process $(\tilde{\Gamma})^{p,c} = (\tilde{\Gamma}_0, \tilde{\Gamma}_1, \ldots)$ as follows: Each $\tilde{\Gamma}_k$ will consist of a finite number of points v each of which has integers $l(v) \in \{m-1, m\}$ and $s(v) \in \{0, 1\}$ and a real number $\alpha(v) \in (0, 1)$ associated with it. ($s(v)$ records whether or not to used size-biased selection as described above.) Let $\tilde{\Gamma}'_0 = \{v_0\}$, where $l(v_0) = m$, $s(v_0) = 0$ and $\alpha(v_0)$ is chosen uniformly from $[0, 1]$. (We ignore probability zero events, so we can assume all α values actually lie in $(0, 1)$.) If $\alpha(v_0) \le c$, set $\tilde{\Gamma}_0 = \emptyset$, otherwise with probability p set $\tilde{\Gamma}_0 = \tilde{\Gamma}'_0$, and with probability $1 - p$ set $\tilde{\Gamma}_0 = \emptyset$.

Given $\tilde{\Gamma}_k$, construct $\tilde{\Gamma}_{k+1}$ as follows: For each $v \in \tilde{\Gamma}_k$, we construct independently $l(v)$ *potential* left-children of v. Each potential left-child is an actual left-child $w \in \tilde{\Gamma}_{k+1}$ with probability the integral of (7.1) below, and has $l(w) = m$, $s(w) = 1$, and $\beta = \alpha(w)$ distributed according to the (normalized version of) the density

$$\frac{p}{2\sqrt{\alpha\beta}}\mathrm{d}\beta, \quad c < \beta < \alpha, \tag{7.1}$$

where $\alpha = \alpha(v)$. For the right-children of v, first construct a random variable $x = x(v)$ with the distribution Z_m, if $s(v) = 0$, or the size-biased version of Z_m, if $s(v) = 1$. Then construct a Poisson number of right-children $w \in \tilde{\Gamma}_{k+1}$ all with $l(w) = m - 1$ and $s(w) = 0$, so that the probability density that such a w is created with $\beta = \alpha(w)$ in a small interval is given by

$$\frac{px}{2\sqrt{\alpha\beta}}\mathrm{d}\beta. \tag{7.2}$$

We state without proof the equivalent of Theorem 6.1 for $(\tilde{\Gamma})^{p,c}$; the proof is exactly the same. The process $(\Gamma)^{p,c}$ is defined exactly as (Γ), except that we delete from $G_m^{(n)}$ all vertices i with $i \le cn$, and all other vertices independently with probability $1 - p$. In particular, in $(\Gamma)^{p,c}$ we have $\Gamma_0 = \emptyset$ if the uniformly chosen initial vertex was deleted. Recall that $\epsilon = 1/1000$.

Theorem 7.1. *Let $0 < p \le 1$ and $0 \le c < 1$ be fixed. For each n, the processes $(\Gamma)^{p,c} = (\Gamma_0, \Gamma_1, \ldots)$ and $(\tilde{\Gamma})^{p,c} = (\tilde{\Gamma}_0, \tilde{\Gamma}_1, \ldots)$ can be coupled so that with probability $1 - o(n^{-\epsilon})$ we have*

$$|\Gamma_k| = |\tilde{\Gamma}_k|$$

for $0 \le k \le K$, *and either* $|\Gamma_K| = |\tilde{\Gamma}_K| = 0$ *or*

$$\sum_{k=0}^{K+1} |\Gamma_k|, \quad \sum_{k=0}^{K+1} |\tilde{\Gamma}_k| \ge n^\epsilon.$$

We shall be interested in the probability that the process $(\tilde{\Gamma})^{p,c}$ never dies out, i.e., that $\tilde{\Gamma}_k \neq \emptyset$ for all $k \ge 0$. Let us say that $v \in \tilde{\Gamma}_k$ *propagates* if v has descendants in all later generations. With p and c fixed, the probability that v propagates is a function $P(\alpha, l, s)$ of $\alpha(v)$, $l(v)$ and $s(v)$. Since v propagates if and only if at least one of its left- or right-children does so, it makes sense to consider separately survival in each of these ways.

For $c < \alpha \le 1$, let $L(\alpha)$ be the chance that a particular potential left-child of some v with $\alpha(v) = \alpha$ is actual, and propagates. From the form of (7.1), we see that this probability depends only on α. The chance that no left-child of v propagates is then

$$(1 - L(\alpha))^{l(v)}. \tag{7.3}$$

For $c < \alpha \le 1$ and $x > 0$, let $r(\alpha, x)$ be the conditional probability that some right-child of v propagates, where $\alpha(v) = \alpha$ and we condition on $x(v) = x$. Then the chance that no right-child of v propagates is

$$\int_0^\infty (1 - r(\alpha, x)) f(x) \mathrm{d}x,$$

where $f(x)$ is the probability density function of Z_m or its size-biased version, according to $s(v)$, so

$$f(x) = \frac{x^{m-1+s(x)} e^{-x}}{(m-1+s(x))!}.$$

From the Poisson nature of the process and the fact that (7.2) is proportional to x, it is easy to see that

$$1 - r(\alpha, x) = \exp(-xR(\alpha))$$

for some function $R(\alpha)$. Hence the probability that no right-child of v propagates is given by

$$\int_0^\infty e^{-xR(\alpha)} \frac{x^{m-1+s(x)} e^{-x}}{(m-1+s(x))!} \mathrm{d}x = (1 + R(\alpha))^{-m-s(x)}, \tag{7.4}$$

using $\int_0^\infty x^a e^{-bx} \mathrm{d}x = a!/b^{a+1}$.

Combining (7.3) and (7.4), recalling that v fails to propagate if and only if all its left- and right-children so fail, we have

$$P(\alpha, l, s) = 1 - \frac{(1 - L(\alpha))^{l(v)}}{(1 + R(\alpha))^{m+s(x)}}. \tag{7.5}$$

Note that the overall probability $\sigma(p, c)$ that $(\tilde{\Gamma})^{p,c}$ does not die out is given by

$$\sigma(p, c) = p \int_{\alpha=c}^{1} \left(1 - \frac{(1 - L(\alpha))^m}{(1 + R(\alpha))^m}\right) \mathrm{d}\alpha, \tag{7.6}$$

since the initial point has $l(v_0) = m$, $s(v_0) = 0$ and $\alpha(v_0)$ chosen uniformly from $[0, 1]$, but this point is kept only if $\alpha(v_0) > c$, and then only with probability p.

Returning to the definition of $L(\alpha)$, we see that $L(\alpha)$ is given by integrating $P(\beta, m, 1)$ times the density (7.1), so

$$L(\alpha) = \frac{p}{2\sqrt{\alpha}} \int_{\beta=c}^{\alpha} \frac{1}{\sqrt{\beta}} \left(1 - \frac{(1 - L(\beta))^m}{(1 + R(\beta))^{m+1}}\right) \mathrm{d}\beta. \tag{7.7}$$

Similarly, after a moment's thought we see from the definition of $R(\alpha)$, (7.2) and (7.5) that

$$R(\alpha) = \frac{p}{2\sqrt{\alpha}} \int_{\beta=\alpha}^{1} \frac{1}{\sqrt{\beta}} \left(1 - \frac{(1 - L(\beta))^{m-1}}{(1 + R(\beta))^m}\right) \mathrm{d}\beta. \tag{7.8}$$

Let us write the two equations above as

$$(L, R) = \mathbf{F}((L, R)) \tag{7.9}$$

where $\mathbf{F}$ is the functional (on functions from $(c, 1]$ to $\mathbb{R}^2$) given by the right hand sides of (7.7) and (7.8). Note that $\mathbf{F}$ is monotonic with respect to pointwise comparison: If $L'(\alpha) \geq L(\alpha)$ and $R'(\alpha) \geq R(\alpha)$ for all α, then the same inequalities hold for $\mathbf{F}(L', R')$ and $\mathbf{F}(L, R)$. Standard probability theory tells us that the propagation probabilities $L(\alpha)$ and $R(\alpha)$ are given by the maximum solution to (7.9). Note that this makes sense as the supremum of all solutions gives a solution (there are easy pointwise upper bounds on any solution to (7.9)).

In the next two sections, we apply Theorem 7.1 to G_p and G_c, for which we need to solve, or bound solutions to, (7.9) in two different cases.

8. Robustness

In this section, we fix $p > 0$ and use Theorem 7.1 with $c = 0$ to analyze G_p, the graph obtained from $G_m^{(n)}$ by deleting vertices independently with probability $1-p$, keeping them with probability p. Our aim is to prove Theorem 3.1, showing that, whatever the value of p, the size of the giant component is $\Theta(n)$.

The first step is to show that in this case the process $(\tilde{\Gamma})^{p,0}$ has positive probability of never dying out. From (7.6), this is equivalent to showing that $L(\alpha)$ and $R(\alpha)$ are not (almost everywhere) zero. As long as we do not want the best possible bound, this is easy.

Let us consider "trial functions" $L_0(\alpha) = 0$,

$$R_0(\alpha) = \begin{cases} 1, & 0 < \alpha < \alpha_0, \\ 0, & \alpha_0 < \alpha \leq 1, \end{cases}$$

for some small α_0 to be chosen later. Let $(L_1, R_1) = \mathbf{F}((L_0, R_0))$. Then, very crudely, for $\alpha > \alpha_0$ we see from (7.7) that $L_1(\alpha) \geq 7p/8\sqrt{\alpha_0/\alpha}$: The part of the integrand in brackets is at least $7/8$ for $\alpha \leq \alpha_0$, and $\int_0^{\alpha_0} \beta^{-1/2} \mathrm{d}\beta = 2\sqrt{\alpha_0}$. For the rest of the range, we use only $L_1(\alpha) \geq 0$. Now set $(L_2, R_2) = \mathbf{F}((L_1, R_1))$. From monotonicity of (7.8), our lower bound on L_1, and $R_1 \geq 0$ we have, for $\alpha \leq \alpha_0$,

$$R_2(\alpha) \geq \frac{p}{2\sqrt{\alpha}} \int_{\beta=\alpha_0}^{1} \frac{1}{\sqrt{\beta}} \left(1 - \left(1 - \frac{7p\sqrt{\alpha_0}}{8\sqrt{\beta}}\right)^{m-1}\right) \mathrm{d}\beta.$$

Now as $m \geq 2$, the quantity in the large brackets is at least $7p\sqrt{\alpha_0}/(8\sqrt{\beta})$, so, for $\alpha < \alpha_0$,

$$R_2(\alpha) \geq \frac{p}{2\sqrt{\alpha_0}} \int_{\beta=\alpha_0}^{1} \frac{7p\sqrt{\alpha_0}}{8\beta} \mathrm{d}\beta = \frac{7p^2}{16} \log(1/\alpha_0).$$

Setting $\alpha_0 = \exp(-16p^{-2}/7)$, we see that, for $0 < \alpha \leq \alpha_0$, we have $R_2(\alpha) \geq 1 = R_0(\alpha)$, and hence we have $(L_2, R_2) \geq (L_0, R_0)$ pointwise. It follows immediately that $(L, R) \geq (L_0, R_0)$: We can keep iterating the monotone functional $\mathbf{F}$, obtaining a sequence (L_i, R_i) with $(L_i, R_i) \geq (L_{i-2}, R_{i-2})$ for each i. Since we have local upper bounds on the functions L, R, the even and odd terms of this sequence each converge pointwise. Using monotonicity again, we see that these limits are the same, and are a solution of (7.9) at least as large as (L_0, R_0). But (L, R) is the maximum solution to (7.9). Plugging back in to (7.6), we see that the probability $\sigma(p, 0)$ that $(\tilde{\Gamma})^{p,0}$ does not die out is always positive, and in fact that

$$\sigma(p, 0) \geq \exp(-\Theta(p^{-2})) \tag{8.1}$$

as $p \to 0$.

We can now complete the proof of Theorem 3.1.

Proof of Theorem 3.1. Fix $0 < p \leq 1$ and set $c = 0$. For $0 \leq k \leq n$, let us write N_k for the number of vertices of G_p in components of order k, counting deleted vertices as being in components of order 0, so N_0 has a Binomial $\mathrm{Bi}(n, (1-p))$ distribution.

For $0 \leq k < \infty$, let us write μ_k for the probability that in the process $(\tilde{\Gamma})^{p,0}$ we have $|\cup_{t=0}^{\infty} \tilde{\Gamma}_t| = k$, so $\sum_{k=0}^{\infty} \mu_k = 1 - \sigma(p, 0)$. From the approximation result, Theorem 7.1, we have

$$\mathbb{E}(N_k) = n\mu_k + o(n^{1-\epsilon})$$

for $0 \le k < n^\epsilon$. It is easy to check that N_k is concentrated about its mean. An outline proof is as follows: Recall that in the proof of Theorem 7.1 we fixed $\mathbf{W} = (W_1, W_2, \ldots, W_n)$, assuming only certain conditions about the distribution of w_i values for i in certain intervals, which hold **wvhp**. Thus, the proof actually shows that **wvhp** we have $\mathbb{E}(N_k \mid \mathbf{W}) = n\mu_k + o(n^{1-\epsilon})$. Let us write $X \subset [n]$ for the vertex set of G_p. Then it is easy to see that **wvhp** we have

$$\mathbb{E}(N_k \mid \mathbf{W}, X) = n\mu_k + o(n^{1-\epsilon}).$$

The proof proceeds as before assuming that among vertices i in certain intervals with $w(i)$ in certain intervals, roughly the right proportion (p) lie in X. Now, having fixed $\mathbf{W}$ and X, the graph $G_m^{(n)}$ and thus G_p is determined by the mn choices of the variables $t_{i,r}$. Changing one such choice affects only one edge of $G_m^{(n)}$. The effect on G_p is to delete at most one edge and then add at most one, changing N_k by at most 4. Applying a suitable martingale inequality, **wvhp** N_k is within $O(n^{2/3})$ of its expectation given $\mathbf{W}$ and X, which is **wvhp** close to $n\mu_k$.

Thus, **wvhp** we have

$$N_k = \mathbb{E}(N_k) + O(n^{2/3}) = n\mu_k + o(n^{1-\epsilon}) \tag{8.2}$$

for all k in the range $0 \le k < n^\epsilon$. Noting that μ_k does not depend on n, and that since $\sum_k \mu_k$ converges we have $\mu_k \to 0$, it follows that **wvhp** in G_p there are $\sigma(p,0)n + o(n)$ vertices in large components. It only remains to show that nearly all such vertices are in a unique giant component.

The uniqueness of the giant component can be established using the methods in the next section, or, more directly, as follows. Let us note that once the neighbourhood expansion considered in [Bollobás and Riordan to appear] takes off, it keeps going. In particular, denoting by i_0 the vertex of G_p with smallest index, the argument there easily shows that with probability $1 - o(1)$ the graph G_p is such that every vertex v with $|\Gamma_k(v)| \ge (\log n)^{10}$ for some k is connected to i_0. To complete the proof, it only remains to show that it is unlikely that for a fixed vertex v we have $|\cup_{k=0}^{\infty} \Gamma_k(v)|$ large (at least $(\log n)^{20}$, say) but $|\Gamma_k(v)|$ small for every k. We can do this without further calculation using Theorem 7.1 and basic properties of the process $(\tilde{\Gamma})^{p,0}$.

It is easy to check that $P(\alpha, l, s)$ decreases with α but is nonzero at $\alpha = 1$, and hence is uniformly bounded away from zero. Also, for $m \ge 3$, the probability given $\alpha(v)$, $l(v)$, and $s(v)$ that v has at least two children is bounded below (by p^2), so the probability that v has at least two children which propagate is bounded below. (For $m = 2$ we must look two steps ahead—we shall ignore this complication.) In particular, the (larger) probability that v has at least two

children which propagate, *given that v propagates*, is bounded below by some p_0, say. It follows that, conditioning on v_0 propagating, the process formed by taking only those descendants of v_0 which themselves propagate dominates a branching process in which each point has either one or two children, having two with probability p_0. Such a process diverges exponentially, so given that v_0 propagates, with probability $1-o(1)$ we have $|\tilde{\Gamma}_{\log n}| \geq (\log n)^{10}$, say. It follows that the probability that $|\tilde{\Gamma}_k| < (\log n)^{10}$ for every k for which $|\cup_{j=0}^{k} \tilde{\Gamma}_j| < n^{\epsilon}$ is $o(1)$. Thus, using Theorem 7.1 and the remarks above, the expected number of vertices of G_p in large components but not the giant component is $o(n)$, completing the proof. □

With a little work, the $o(1)$ error probability above may be replaced by $o(n^{-\epsilon})$. Note also that together with (8.1), the argument above proves the lower bound given in (3.1).

We have now proved our main result, Theorem 7.1, showing that for any $p > 0$ the graph G_p has a giant component, and given a (very small) lower bound on its size. We now turn to the upper bound on the size of this component.

Note that from (8.2), since $\sum_k \mu_k = 1 - \sigma(p,0)$, to get an upper bound on $L_1(G_p)$ it suffices to obtain an upper bound for $\sigma(p,0)$. For this, it turns out to be convenient to go back to the definition of $(\tilde{\Gamma})^{p,0}$ itself, rather than using (7.7), (7.8). The basic method is as follows: We consider a small α_0, and call a point v with $\alpha \leq \alpha_0$ or $\alpha > \alpha_0$ *good* or *bad*, respectively. We note that if v_0 survives, then with probability 1 it has a good descendant; loosely speaking, infinitely many of its descendants each has positive probability of having a good left-child. We shall show that v_0 is unlikely to have a good descendant by showing that when iterated in the right way the neighbourhoods of bad vertices tend to shrink.

For a set V of points of $(\tilde{\Gamma})^{p,c}$, let us write $\mathcal{R}(V)$ for the set of right-children of points in V. Let $\mathcal{R}^k(V) = \mathcal{R}(\mathcal{R}^{k-1}(V))$, let $\mathcal{R}_\infty(V) = \cup_{k=0}^{\infty} \mathcal{R}^k(V)$, and let $\mathcal{R}_\infty^+(V) = \mathcal{R}_\infty(\mathcal{R}(V)) = \cup_{k=1}^{\infty} \mathcal{R}^k(V)$. Suppose that $\alpha(v) = \alpha$. Then from (7.2) and the fact that Z_m has mean m and its size-biased version mean $m+1$, the probability density $\rho_1(\alpha,\beta)\mathrm{d}\beta$ that $\mathcal{R}(\{v\})$ contains a point w with $\beta = \alpha(w)$ in a small interval is given by

$$\rho_1(\alpha,\beta) = \frac{p(m+s(v))}{2\sqrt{\alpha\beta}}\mathrm{d}\beta,$$

for $\alpha < \beta < 1$. Since each point w of $\mathcal{R}(\{v\})$ has $s(w) = 0$, it follows by induction that for $k \geq 1$, the corresponding probability density $\rho_k(\alpha,\beta)$ for $\mathcal{R}^k(\{v\})$ is given by

$$\rho_k(\alpha,\beta) = \int_{\alpha<\beta_1<\cdots<\beta_{k-1}<\beta} \frac{p^k(m+s(v))m^{k-1}}{2^k\sqrt{\alpha\beta}\beta_1\cdots\beta_{k-1}}\mathrm{d}\beta_1\cdots\mathrm{d}\beta_{k-1},$$

for $\alpha < \beta < 1$. The integral above is just $1/(k-1)!$ times a product of independent integrals, so we obtain

$$\rho_k(\alpha,\beta) = \frac{p^k(m+s(v))m^{k-1}(\log(\beta/\alpha))^{k-1}}{2^k(k-1)!\sqrt{\alpha\beta}}.$$

Summing over k, the corresponding density for $\mathcal{R}^+_\infty(V)$ is

$$\sum_{k\geq 1}\rho_k(\alpha,\beta) = \frac{p(m+s(v))}{2\sqrt{\alpha\beta}}\exp\left(mp\log(\beta/\alpha)/2\right).$$

Let $\alpha_0 = \exp(-1/(2mp))$. Then we have

$$\sum_{k\geq 1}\rho_k(\alpha,\beta) \leq \frac{pm}{\sqrt{\alpha\beta}} \tag{8.3}$$

whenever $\alpha > \alpha_0$, using only $\beta \leq 1$ to bound the exponential term.

Corresponding to the definition for vertices of $G_m^{(n)}$ in [Bollobás and Riordan to appear], let us define the *weight* $w(v)$ of a point v of $(\tilde{\Gamma})^{p,c}$ as $1/\sqrt{\alpha(v)}$, and the *weight* of a set as the sum of the weights of its points. Then, provided v is bad, i.e., $\alpha = \alpha(v) > \alpha_0$, from (8.3) we have

$$\begin{aligned}\mathbb{E}(w(\mathcal{R}^+_\infty(\{v\}))) &= \int_{\beta=\alpha}^1 \frac{1}{\sqrt{\beta}}\sum_{k\geq 1}\rho_k(\alpha,\beta)\mathrm{d}\beta \\ &\leq \int_{\beta=\alpha}^1 \frac{pm}{\beta\sqrt{\alpha}} \leq pm\log(1/\alpha)w(v) \leq w(v)/2.\end{aligned}$$

Since $\mathcal{R}_\infty(\{v\}) = \{v\}\cup\mathcal{R}^+_\infty(\{v\})$ it follows that $\mathbb{E}(w(\mathcal{R}_\infty(\{v\}))) \leq 3w(v)/2$. Thus for any set V of bad vertices we have $\mathbb{E}(w(\mathcal{R}_\infty(V))) \leq 3w(V)/2$. In particular, writing $B(V)$ for the event that a random set V consists only of bad vertices, and $\mathbb{I}_{B(V)}$ for its indicator function, we have

$$\mathbb{E}(\mathbb{I}_{B(\mathcal{R}_\infty(V))}w(\mathcal{R}_\infty(V)) \mid V) \leq 3\mathbb{I}_{B(V)}w(V)/2. \tag{8.4}$$

Let us write $\mathcal{L}(V)$ for the set of left-children of points of V. Note first that, integrating (7.1) from 0 to α_0, the probability that a particular left-child of a point v is good is exactly $p\sqrt{\alpha_0/\alpha(v)} = p\sqrt{\alpha_0}w(v)$. From this, it follows that

$$\Pr(\mathcal{L}(v) \text{ contains a good point}) \leq mp\sqrt{\alpha_0}w(V). \tag{8.5}$$

Considering now bad points in $\mathcal{L}(V)$, integrating $1/\sqrt{\beta}$ times (7.1) over the relevant range for each of the $l(v)$ left-children of a point v with $\alpha(v) = \alpha > \alpha_0$,

we see that

$$\begin{aligned}\mathbb{E}(\mathbb{I}_{B(\mathcal{L}(\{v\}))}w(\mathcal{L}(\{v\}))) &\leq l(v)\int_{\beta=\alpha_0}^{\alpha}\frac{1}{\sqrt{\beta}}\frac{p}{2\sqrt{\alpha\beta}}\mathrm{d}\beta \\ &= l(v)\frac{p\log(\alpha/\alpha_0)}{2\sqrt{\alpha}} \\ &\leq mp\log(1/\alpha_0)w(v)/2 = w(v)/4.\end{aligned} \tag{8.6}$$

All the pieces are now in place. Let us note that the probability that v_0 is good is exactly α_0. Let $S_0 = \mathcal{R}_\infty(\{v_0\})$ and for $k \geq 1$ set $S_k = \mathcal{R}_\infty(\mathcal{L}(S_{k-1}))$, so the union of the S_k is all the descendants of v_0. If v_0 is bad but one of its descendants is good, then for some k there is a good point in $\mathcal{L}(S_k)$, as to reach a good point from a bad point we must take a left-child. Considering the minimum such k we may assume that $S_0, \ldots, S_k$ consist only of bad points. Using (8.5),

$$\Pr(S_k \text{ is bad but } \mathcal{L}(S_k) \text{ is not} \mid S_k) \leq mp\sqrt{\alpha_0}\mathbb{I}_{B(S_k)}w(S_k),$$

so

$$\Pr(S_0 \cup \cdots \cup S_k \text{ is bad but } \mathcal{L}(S_k) \text{ is not}) \leq mp\sqrt{\alpha_0}\,\mathbb{E}(\mathbb{I}_{B(S_0\cup\cdots\cup S_k)}w(S_k)).$$

Hence $\sigma(p,0)$, which, as noted above is the same as the probability that v_0 has a good descendant, can be bounded as follows:

$$\sigma(p,0) \leq \alpha_0 + \sum_{k=0}^{\infty} mp\sqrt{\alpha_0}\,\mathbb{E}(\mathbb{I}_{B(S_0\cup\cdots\cup S_k)}w(S_k)). \tag{8.7}$$

But from (8.4) we have

$$\mathbb{E}(\mathbb{I}_{B(S_0)}w(S_0)) \leq \frac{3}{2}\,\mathbb{E}(\mathbb{I}_{B(\{v_0\})}w(v_0)) = \frac{3}{2}\int_{\beta=\alpha_0}^{1}\frac{1}{\sqrt{\beta}}\mathrm{d}\beta = 3\sqrt{\alpha_0}.$$

From (8.4) and (8.6) and the inductive definition of S_k, we have for $k \geq 1$ that

$$\mathbb{E}(\mathbb{I}_{B(S_0\cup\cdots\cup S_k)}w(S_k)) \leq \frac{3}{8}\,\mathbb{E}(\mathbb{I}_{B(S_0\cup\cdots\cup S_{k-1})}w(S_{k-1})).$$

Returning to (8.7), it follows that

$$\begin{aligned}\sigma(p,0) &\leq \alpha_0 + mp\sqrt{\alpha_0}3\sum_{k=0}^{\infty}\sqrt{\alpha_0}(3/8)^k \\ &\leq (1+5mp)\alpha_0 = (1+5mp)\exp(-1/(2mp)).\end{aligned}$$

As noted earlier, $(\sigma(p,0)+o(1))n$ is an upper bound on the size $L_1(G_p)$ of the giant component in G_p with high probability, proving the upper bound in (3.1).

9. Vulnerability

In this section, we turn to the question of the vulnerability of $G_m^{(n)}$ to malicious attack, considering the graph G_c obtained from $\mathcal{G}_m^{(n)}$ by deleting all vertices i with $i \le cn$. It turns out that using the approximation result Theorem 7.1 we can prove Theorem 3.2 with very little actual calculation, for the most part using only very general properties of the process $(\tilde{\Gamma})^{p,c}$.

Throughout this section, we take $p = 1$, and consider the process $(\tilde{\Gamma})^{1,c}$ for $0 \le c \le 1$. At the very end, we translate our results back to G_c.

Let us write $\theta(c) = \sigma(1, c)$ for the probability that the initial vertex v_0 of $(\tilde{\Gamma})^{1,c}$ propagates, i.e., the process $(\tilde{\Gamma})^{1,c}$ consists of infinitely many points. Let us write $\mu_{k,c}$ for the probability that $(\tilde{\Gamma})^{1,c}$ consists of exactly k points, emphasizing the dependence on c. Note that, by definition,

$$\sum_{k=0}^{\infty} \mu_{k,c} = 1 - \theta(c). \tag{9.1}$$

Now, for $c_1 < c_2$ we can couple the processes $(\tilde{\Gamma})^{1,c_1}$ and $(\tilde{\Gamma})^{1,c_2}$ so that they are nested in the obvious sense, with $(\tilde{\Gamma})^{1,c_1}$ containing $(\tilde{\Gamma})^{1,c_2}$. Indeed, we can think of $(\tilde{\Gamma})^{1,c_2}$ as obtained from $(\tilde{\Gamma})^{1,c_1}$ by deleting all points v with $c_1 < \alpha(v) \le c_2$. It follows that, unsurprisingly, $\theta(c)$ decreases with c.

We claim that $\theta(c)$ is continuous. Left-continuity is fairly easy to check: fix $c > 0$ and consider an arbitrary $\eta > 0$. From (9.1), we see that there is a K with

$$\sum_{k=0}^{K} \mu_{k,c} \ge 1 - \theta(c) - \eta/2. \tag{9.2}$$

For $0 < \delta < c$, let us couple $(\tilde{\Gamma})^{1,0}$, $(\tilde{\Gamma})^{1,c-\delta}$, and $(\tilde{\Gamma})^{1,c}$ so that they are nested, thinking of $(\tilde{\Gamma})^{1,c}$ and $(\tilde{\Gamma})^{1,c-\delta}$ as obtained from $(\tilde{\Gamma})^{1,0}$ by deleting points as above. We condition on the entire process $(\tilde{\Gamma})^{1,c}$, showing that adding a few more points by passing to $(\tilde{\Gamma})^{1,c-\delta}$ does not increase the chance of v_0 propagating by much.

Let us condition on $(\tilde{\Gamma})^{1,c}$, supposing that the process consists of exactly k points $v_0, \dots, v_{k-1}$, $k \le K$. Each v_i has at most $l(v_i)$ left-children in $(\tilde{\Gamma})^{1,c}$; the "missing" left-children w have $\alpha(w)$ distributed on $[0, c]$ in proportion to the density (7.1). Thus each of the, at most, mk missing left-children has probability

$$\left(\int_{c-\delta}^{c} \beta^{-1/2} \mathrm{d}\beta\right) / \left(\int_{0}^{c} \beta^{-1/2} \mathrm{d}\beta\right) = O(\delta)$$

of being added back when we pass to $(\tilde{\Gamma})^{1,c-\delta}$. In particular, for each k the probability that $(\tilde{\Gamma})^{1,c-\delta} = (\tilde{\Gamma})^{1,c}$, given that $(\tilde{\Gamma})^{1,c}$ consists of exactly k points,

tends to 1. Hence, for each k, we have $\mu_{k,c-\delta} \geq \mu_{k,c} - \eta/(2K+2)$ for δ sufficiently small. Since the sum involved has finitely many terms, it follows that if δ is small enough, then

$$1 - \theta(c-\delta) \geq \sum_{k=0}^{K} \mu_{k,c-\delta} \geq \sum_{k=0}^{K} \mu_{k,c} - \eta/2 \geq 1 - \theta(c) - \eta,$$

where the last step is from (9.2). In other words, if $\delta > 0$ is small enough, then $\theta(c-\delta) < \theta(c) - \eta$. As $\theta(.)$ is decreasing and η was arbitrary, this establishes continuity of $\theta(c)$ from the left.

The argument for right-continuity is slightly more involved and we give only an outline. Note first that if $\theta(c) = 0$, then θ is automatically right-continuous at c, as θ is decreasing and non-negative. Thus we may suppose $\theta(c) > 0$. Now it is easy to see from the fact that (7.1) and (7.2) are uniformly bounded below that the same holds for the propagation probability $P(\alpha, l, s)$. Arguing along the same lines as in Section 8, we see that conditioning on v_0 propagating, with high probability the process $(\tilde{\Gamma})^{1,c}$ grows exponentially at some positive rate. From this, it follows that when passing to $(\tilde{\Gamma})^{1,c+\delta}$, for v_0 to now fail to propagate, with high probability we must delete at least a certain positive fraction of points in some Γ_k. But the probability that we do this tends to zero as $\delta \to 0$, implying right-continuity of $\theta(c)$.

We have now shown that $\theta(c)$ is decreasing and continuous. As $\theta(0) = 1$ and $\theta(1) = 0$ the critical value $c_0 = \inf\{c : \theta(c) = 0\}$ satisfies $0 < c_0 < 1$, and we have $\theta(c_0) = 0$.

For for $0 < c < c_0$, we have $\theta(c) > 0$, so the left and right propagation probabilities $L_c(\alpha)$ and $R_c(\alpha)$ given by the largest solution to equations (7.7) and (7.8) are nonzero. Since as $c \to c_0$ we have $\theta(c) \to 0$, it follows that L_c and R_c tend to zero uniformly as c increases to c_0. Here we again use the fact that the densities (7.1) and (7.2) are bounded below. It follows that suitable "normalized limits" $\tilde{L}$, $\tilde{R}$ of L_c and R_c satisfy the linearized versions of (7.7), (7.8):

$$\begin{aligned} \tilde{L}(\alpha) &= \frac{1}{2\sqrt{\alpha}} \int_{\beta=c}^{\alpha} \frac{1}{\sqrt{\beta}} \Big(m\tilde{L}(\beta) + (m+1)\tilde{R}(\beta) \Big) \, \mathrm{d}\beta, \\ \tilde{R}(\alpha) &= \frac{1}{2\sqrt{\alpha}} \int_{\beta=\alpha}^{1} \frac{1}{\sqrt{\beta}} \Big((m-1)\tilde{L}(\beta) + m\tilde{R}(\beta) \Big) \, \mathrm{d}\beta. \end{aligned}$$

Converting these equations to differential form after multiplying through by $\sqrt{\alpha}$, it is easy to solve the equations explicitly (we omit the tedious details), and we find that they have a nonzero solution if and only if $c = \frac{m-1}{m+1}$. As shown above, these equations have a nonzero solution if $c = c_0$, so we have $c_0 = \frac{m-1}{m+1}$.

At this point, we have shown that $\theta(c) = \sigma(1, c)$ is zero if $c \geq \frac{m-1}{m+1}$ and positive if $c < \frac{m-1}{m+1}$. Let us say that a component is *small* if it has fewer than $n^{\epsilon/2}$ vertices, and *large* otherwise. Arguing as in the previous section, from Theorem 7.1 it is easy to deduce the analogue of (8.2), namely that **wvhp** we have

$$N_{k,c} = \mu_{k,c} n + o(n^{1-\epsilon}) \tag{9.3}$$

for $0 \leq k < n^{\epsilon}$, where $N_{k,c} = N_{k,c}(n)$ is the number of vertices of G_c in components of order k. In particular, we see that **wvhp** the graph G_c has $(1 - \theta(c) + o(1))n$ vertices in small components and hence $\theta(c) + o(1)$ vertices in large components. If $c \geq c_0 = \frac{m-1}{m+1}$, so $\theta(c) = 0$, it follows that $L_1(G_c) = o(n)$ as required. For $c < c_0$, it remains to show that most of the vertices in large components are, in fact, in a single component. We can do this with very little extra calculation by considering the graphs $G_m^{(n)}$ for different values of n.

Proof of Theorem 3.2. Let us now suppose that $c < c_0 = \frac{m-1}{m+1}$ is constant; as noted above we have already dealt with the case $c \geq c_0$. Let $\gamma = \gamma(n)$ be some function tending to zero sufficiently slowly; for definiteness we shall take $\gamma = n^{-\epsilon/3}$. Set $n' = \lfloor (1-\gamma)n \rfloor$. Returning to the original process defining $G_m^{(n)}$, let us consider the nested sequence of graphs $G_m^{(t)}$, $t = n', n'+1, \ldots, n$, and their truncated versions H_t, where H_t is obtained from $G_m^{(t)}$ by deleting all vertices i with $i \leq cn$, so in particular, $H_n = G_c$, the graph we are interested in.

Since $\gamma \to 0$, for n sufficiently large we have $c' = cn/n' < c_0$. Applying (9.3) with n, c replaced by n', c' we see that **wvhp** the graph $H_{n'}$ has $(\theta(c') + o(1))n'$ vertices in large components. (We use n rather than n' in the definition of large.) As θ is continuous, this quantity is $(\theta(c) + o(1))n$. Let us call a vertex $1 \leq i \leq n$ *important* if it is in a large component of $H_{n'}$; note that there are $(\theta(c) + o(1))n$ important vertices. Also, let us call a component of H_t, $n' \leq t \leq n$, *important* if it contains an important vertex. Note that the set of important vertices is fixed as t varies, while passing from H_t to H_{t+1} no new important components are created, and one or more pairs (or larger groups) of important components may unite to form important components in H_{t+1}. In particular, the number of important components decreases with t. Our aim is to show that $H_n = G_c$ has a single important component; this suffices to complete the proof as this component then contains all $(\theta(c) + o(1))n$ important vertices, which is within $o(n)$ of the total number of vertices of G_c in large components.

Let A and B be distinct components of some H_t with a and b vertices, respectively. Let $d(A)$ be the sum over $i \in A$ of the degree of i *in the graph* $G_m^{(t)}$. Then when we pass to $G_m^{(t+1)}$, from the general m version of (2.1) we see that the probability that the first edge from the new vertex lands in A is exactly $d(A)/(2mt + 1)$. Given this, the probability that the second edge lands in B

is $d(B)/(2mt+3)$. Hence, the probability that in H_{t+1} the components A and B have been united is at least $d(A)d(B)/(2mt+1)/(2mt+3)$. Now $d(A) \geq a$, $d(B) \geq b$, while $t \sim n$, so this probability is at least $\kappa ab/n^2$ for some positive constant κ. For $n' \leq t \leq n$, let $S_2(t)$ be the sum of $|A||B|$ over all unordered pairs $\{A, B\}$ of distinct important components of H_t. If $S_2(t) \geq (\theta(c)n)^2/5$, then with probability $\kappa' = \theta(c)^2\kappa/5$ two important components of H_t become united in H_{t+1}, so H_{t+1} has fewer important components than H_t. As $H_{n'}$ has at most $n^{1-\epsilon/2}$ important components, this can happen for at most $n^{1-\epsilon/2}$ values of t. It follows from basic properties of the binomial distribution that **wvhp** we have $S_2(t_0) < (\theta(c)n)^2/5$ for some t_0 with $t_0 \leq n' + 2n^{1-\epsilon/2}/\kappa'$. As S_2 is at least the number of important vertices choose 2 minus the sum over each important component of its size choose 2, it follows that H_{t_0} has an important component C of order at least $\theta(c)n/2$. For a fixed other important component A of H_{t_0}, we see that at each remaining step $t_0 < t \leq n$, the chance that A (or the component containing it) becomes united with C (or the component containing it) is at least $\kappa|A||C|/n^2 \geq \kappa''|A|/n \geq \kappa'' n^{\epsilon/2-1}$ for some constant $\kappa'' > 0$. As there are $n - t \sim n^{1-\epsilon/3}$ steps remaining, the probability that A and C eventually unite is at least $1 - o(n^{-2})$, say. Hence, **wvhp** all important components of H_t are united in H_n. As noted above, this completes the proof of Theorem 3.2. □

10. Comparisons and Conclusions

We finish by comparing the results presented here on the robustness and vulnerability of the LCD model $G_m^{(n)}$ with those for other random graph models. We start with some classical examples.

The simplest classical random graph models are $\mathcal{G}(n, M)$, introduced by Erdős and Rényi [Erdős and Rényi 59], where the graph is chosen uniformly at random from all graphs on n vertices with exactly M edges, and the closely related model $\mathcal{G}(n, \pi)$, introduced by Gilbert [Gilbert 59], where pairs of vertices are joined independently with probability π. Since $G_m^{(n)}$ has mn edges and average degree $2m$, it is natural to take $M = mn$ and $\pi = 2m/(n-1)$; from now on, these are the values we shall take for M and π. For the graphs $\mathcal{G}(n, M)$ and $\mathcal{G}(n, \pi)$ the notion of malicious attack used here does not make sense, so we consider deleting vertices independently at random, keeping them with some constant probability $0 < p < 1$, and ask for the threshold probability p_c above which a giant component appears. We shall refer to graphs obtained in this way as *damaged* graphs. Standard methods tell us that the threshold probability is the same for $\mathcal{G}(n, M)$ and $\mathcal{G}(n, \pi)$, and that there is a giant component if and only if the neighbourhood growth from a vertex "takes off" with positive probability.

In $\mathcal{G}(n,\pi)$, his happens if and only if the expected degree in the damaged graph is greater than one, showing that $p_c = 1/(n-1)\pi = 1/(2m)$. For $G_m^{(n)}$ the corresponding threshold given by Theorem 3.1 is zero, so we see that $G_m^{(n)}$ is much less vulnerable to random attack than $\mathcal{G}(n,M)$ or $\mathcal{G}(n,\pi)$. On the other hand, when $G_m^{(n)}$ is attacked, we see from Theorem 3.2 that we need to keep a fraction $2/(m+1)$ of the vertices to have a giant component, so by this measure $G_m^{(n)}$ is roughly four times more vulnerable to attack.

Although we are treating $G_m^{(n)}$ as an undirected graph, in the definition each vertex sends out m edges, so, for example, the minimum degree of $G_m^{(n)}$ is (at least) m. Thus a more natural classical model to use for comparison is $\mathcal{G}_{m\text{-out}}$, where each vertex sends out m edges independently to other vertices chosen uniformly at random, again considered as an undirected graph. (The usual definition of $\mathcal{G}_{m\text{-out}}$ disallows multiple edges; this makes no difference here.) In this case, to find the critical probability is very slightly more complicated: Again, standard methods tell us that what matters is whether neighbourhoods initially expand, but this time there are two ways we can follow an edge, with or against its orientation in the original directed graph. Not counting the edge we have just followed, in the first case the vertex reached has m outgoing edges in $\mathcal{G}_{m\text{-out}}$, and in the second only $m-1$ other outgoing edges. In either case, the expected number of (other) incoming edges is m. It follows that p_c is given by the reciprocal of the largest eigenvalue of the matrix

$$\begin{pmatrix} m & m \\ m-1 & m \end{pmatrix},$$

so $p_c = 1/(m+\sqrt{m(m-1)}) = 1/(2m-1/2) + O(m^{-3})$, which is close enough to the value for $\mathcal{G}(n,M)$ not to affect the comparison with $G_m^{(n)}$ significantly. We could also compare with a random $2m$-regular graph, although this is less natural. Again, it makes little difference; this time $p_c = 1/(2m-1)$.

The conclusion that scale-free random graphs are more robust than classical random graphs but also more vulnerable was reached on the basis of experimental results by Albert, Jeong, and Barabási [Albert et al. 00]. Detailed comparison with their results is rather difficult, as [Albert et al. 00] is mainly concerned with smaller deletion probabilities, and the results concerning thresholds are rather unclear. However, it is stated that for the scale-free model the data "are consistent with an extremely delayed percolation transition" (at $p = 0.25$ in our notation, for $m = 2$) and "indicate the existence of a critical point." Heuristic results of Cohen, Erez, ben-Avraham, and Havlin [Cohen et al. 00], and Callaway, Newman, Strogatz, and Watts [Callaway et al. 00] (see also [Albert and Barabási 02]) suggest that, as shown here, no such transition occurs (for $p > 0$ constant); the apparent transition noted in [Albert et al. 00] is presumably due

to the small network sizes considered (n up to 20,000). Note, however, that these heuristics only take account of the degree distribution, not the structure of the graph. A modest example of their failure is for $\mathcal{G}_{m\text{-out}}$, where for fixed m the threshold predicted is off by a small constant factor. A more striking example is $G_m^{(n)}$ with $m = 1$, for which the heuristics again predict $p_c = 0$. As shown in Theorem 3.3, the truth is as far from this as possible; in fact $p_c = 1$. While such heuristics are informative, in general a lot of work is needed to tell whether they apply to any given model.

Having noted that $G_m^{(n)}$ is more robust but also more vulnerable than classical random graphs, the natural question is "Why"? There are two main differences in the model; that the graph grows one vertex at a time, with edges sent only to earlier vertices, and "preferential attachment," that new edges are more likely to connect to vertices whose degree is already high. Both of these seem likely to account for some part of the observed difference. To see how important each factor is, let us consider a model $H_m^{(n)}$ defined exactly as $G_m^{(n)}$ is, except with uniform rather than preferential attachment. For definiteness, we may take each vertex i to send out independently m edges, each of which lands at a vertex chosen uniformly from 1 to i, although the details turn out not to matter. The graph $H_m^{(n)}$ is much easier to analyze that $G_m^{(n)}$, due to the independence built into the definition.

One can consider a continuous model for neighbourhood growth in $H_m^{(n)}$ analogous to $(\tilde{\Gamma})^{p,c}$; now instead of (7.1), (7.2), we have the simpler densities

$$\frac{p}{\alpha}\mathrm{d}\beta, \quad c < \beta < \alpha,$$

for each of the $l(v)$ outgoing edges from a vertex v with $\alpha(v) = \alpha$, and

$$\frac{mp}{\beta}\mathrm{d}\beta, \quad \alpha < \beta \leq 1,$$

for incoming edges. From these, arguing as in Section 7 , we obtain simultaneous equations

$$\begin{aligned} L(\alpha) &= \frac{p}{\alpha}\int_{\beta=c}^{\alpha}\left(1-(1-L(\beta))^m e^{-R(\beta)}\right)\mathrm{d}\beta, \\ R(\alpha) &= mp\int_{\beta=\alpha}^{1}\frac{1}{\beta}\left(1-(1-L(\beta))^{m-1}e^{-R(\beta)}\right)\mathrm{d}\beta, \end{aligned}$$

for suitably defined survival probabilities $L(\alpha)$, $R(\alpha)$, and the question is when these equations have a positive solution. Again, we consider the linearized form

of the equations:

$$\begin{aligned} \tilde{L}(\alpha) &= \frac{p}{\alpha}\int_{\beta=c}^{\alpha}\left(m\tilde{L}(\beta)+\tilde{R}(\beta)\right)\mathrm{d}\beta, \\ \tilde{R}(\alpha) &= mp\int_{\beta=\alpha}^{1}\frac{1}{\beta}\left((m-1)\tilde{L}(\beta)+\tilde{R}(\beta)\right)\mathrm{d}\beta. \end{aligned} \tag{10.1}$$

These are easy to solve by setting $\tilde{S}=\alpha\tilde{L}$, multiplying through by α and differentiating.

For vulnerability, we can argue as in Section 9; taking $p=1$, the equations above have a solution if and only if $c=c_0=\frac{m-1}{m}$. Hence, under malicious attack we must keep a fraction greater than $1/m$ of the vertices of $H_m^{(n)}$ to have a giant component. Note that this threshold is (geometrically) roughly half way between the value for $G_m^{(n)}$ and those for classical random graphs.

For robustness, when $c=0$, the situation is rather different, as there is a problem with the continuous model near $\alpha=0$. Although $H_m^{(n)}$ is not our main focus, we include a brief discussion of this as a cautionary tale about the limitations of heuristic arguments when used without rigorous justification. Allowing $p<1$ but, for the moment, keeping $c>0$, the arguments used here give an equivalent form of Theorem 7.1. Setting

$$p_u=\frac{1}{2}\left(1-\sqrt{\frac{m-1}{m}}\right)=\frac{1}{4m}+O(m^{-2})$$

and arguing as in Section 9, we see that for $p>p_u$, there is a giant component if and only if $c<c_0(p)$ for a certain function $c_0(p)$ that can be obtained explicitly by solving the linearized equations. In particular, with $c=0$ and $p>p_u$, there is a giant component. Also, we see that $c_0(p)\to 0$ as p tends to p_u from above, strongly suggesting that the critical probability p_c is equal to p_u, which we believe to be the case. Unfortunately, it is not so easy to analyze the continuous model when $c=0$.

Returning to the original graph $H_m^{(n)}$, the expected number of edges between vertices i and j in $H_m^{(n)}$ is $m/\max\{i,j\}$. (We ignore complications due to loops, which make very little difference.) Let $\mathbf{A}=\mathbf{A}(n)$ be an n by n matrix with entries $a_{ij}=1/\max\{i,j\}$. If we had complete independence between different steps in the walk, then the entries of $(m\mathbf{A})^r$ would give the expected numbers of walks of length r between each pair of vertices $H_m^{(n)}$. In fact, if we only consider paths, the only dependence is negative, so the entries of $(m\mathbf{A})^r$ are upper bounds on the expected number of paths. It follows (we omit the details) that in the damaged graph, where we keep vertices independently with constant probability

p, there can be a giant component only if $mp\lambda_{\max} > 1$, where $\lambda_{\max}$ is the limit as $n \to \infty$ of $\lambda_{\max}(n)$, the maximum eigenvalue of $\mathbf{A}(n)$. Now it is easy to see from standard results (e.g., comparison with a Hilbert matrix with entries $1/(i+j)$) that $\lambda_{\max}(n)$ is bounded by 2π. In fact, with some work one can check that $\lambda_{\max} = 4$, and it follows that for $p \le p_l = \frac{1}{4m}$, the damaged version of $H_m^{(n)}$ has no giant component. Together with the result for $p > p_u$ in the preceding paragraph, this shows that $p_c = \frac{1}{4m} + O(m^{-2})$.

One would expect that the continuous version of the eigenvalue problem above has maximum eigenvalue $\lambda_{\max}$, but this turns out not to be the case; the continuous problem has arbitrarily large eigenvalues. This does not contradict the corresponding continuous Hilbert inequality as the eigenfunctions are not square-integrable. The linearized equations (10.1), although obtained in a different way, correspond to this eigenvalue problem when more of the structure of $H_m^{(n)}$ is taken into account, where $c = 0$ and the eigenvalue is $1/p$. In particular, if these equations have a solution for a certain p, then for all larger p there is an eigenvalue greater than 1, and one would expect a giant component. In fact the equations (10.1) have solutions for all $p > 0$, but, as shown above, the damaged graph can only have a giant component if $p > 1/(4m)$.

Returning to our topic, recall that, in terms of vulnerability, $H_m^{(n)}$ lies roughly half way between $G_m^{(n)}$ and the classical random graphs. Although we have not found the precise threshold for a giant component to remain after random deletion of vertices from $H_m^{(n)}$, this threshold is positive, in fact around $1/(4m)$. Thus $H_m^{(n)}$ is more robust than $\mathcal{G}(n, M)$, by roughly a factor of two in the critical probability. This contrasts sharply with $G_m^{(n)}$, which is much more robust, with critical probability zero. Heuristically, this last difference is not surprising: There is a close connection between robustness and diameter, as suggested by considering expanding neighbourhoods. Ignoring complications, one can say that the reason $G_m^{(n)}$ is so robust is that its "neighbourhood expansion factor" tends to infinity as $n \to \infty$. Again ignoring complications, this is generally speaking the condition giving sublogarithmic diameter. In contrast, $H_m^{(n)}$ has a bounded neighbourhood expansion factor, and, as one can check by counting paths, logarithmic diameter.

Acknowledgments. The work of the first author was supported in part by NSF grant ITR 0225610 and DARPA grant F33615-01-C-1900.

References

[Aiello et al. 01] W. Aiello, F. Chung, and L. Lu. "A Random Graph Model for Power Law Graphs." *Experiment. Math.* 10 (2001), 53–66.

[Albert and Barabási 02] R. Albert and A.-L. Barabási. "Statistical Mechanics of Complex Networks." *Rev. Mod. Phys.* 74 (2002), 47–97.

[Albert et al. 00] R. Albert, H. Jeong, and A.-L. Barabási. "Error and Attack Tolerance of Complex Networks." *Nature* 406 (2000), 378–382.

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

[Bollobás and Riordan to appear] B. Bollobás and O. Riordan. "The Diameter of a Scale-Free Random Graph." To appear in *Combinatorica*. Preprint available from World Wide Web (http://www.dpmms.cam.ac.uk/∼omr10/).

[Bollobás and Riordan 02] B. Bollobás and O. Riordan. "Mathematical Results on Scale-Free Random Graphs." In *Handbook of Graphs and Networks*, edited by Stefan Bornholdt and Heinz Georg Schuster, pp. 1–34. Weinheim: Wiley-VCH, 2002.

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

[Buckley and Osthus to appear] P. G. Buckley and D. Osthus. "Popularity Based Random Graph Models Leading to a Scale-Free Degree Sequence." Preprint available from World Wide Web (http://www.informatik.hu-berlin.de/∼osthus/).

[Callaway et al. 00] D. S. Callaway, M. E. J. Newman, S. H. Strogatz, and D. J. Watts. "Network Robustness and Fragility: Percolation on Random Graphs." *Phys. Rev. Lett.* 85 (2000), 5468–5471.

[Cohen et al. 00] R. Cohen, K. Erez, D. Ben-Avraham, and S. Havlin. "Resilience of the Internet to Random Breakdowns." *Phys. Rev. Lett.* 85 (2000), 4626–4628.

[Cohen et al. 01] R. Cohen, K. Erez, D. Ben-Avraham, and S. Havlin, "Breakdown of the Internet under Intentional Attack." *Phys. Rev. Lett.* 86 (2001), 3682–3685.

[Cooper and Frieze 01] C. Cooper and A. Frieze. "On a General Model of Web Graphs." In *Proceedings of ESA, Lecture Notes in Computer Science 2161*, pp. 500–511. Heidelberg: Springer-Verlag, 2001.

[Cooper and Frieze 02] C. Cooper and A. Frieze. "Crawling on Web Graphs." *STOC* (2002).

[Dorogovtsev and Mendes 02] S. N. Dorogovtsev and J. F. F. Mendes. "Evolution of Networks." *Adv. Phys.* 51 (2002), 1079.

[Erdős and Rényi 59] P. Erdős and A. Rényi. "On Random Graphs I." *Publicationes Mathematicae Debrecen* 5 (1959), 290–297.

[Faloutsos et al. 99] M. Faloutsos, P. Faloutsos, and C. Faloutsos. "On Power-Law Relationships of the Internet Topology." *Comput. Commun. Rev.* 29 (1999), 251.

[Gilbert 59] E. N. Gilbert. "Random Graphs." *Annals of Mathematical Statistics* 30 (1959), 1141–1144.

[Kumar et al. 00] R. Kumar, P. Raghavan, S. Rajagopalan, D. Sivakumar, A. Tomkins, and E. Upfal. "Stochastic Models for the Web Graph." *FOCS* (2000).

[Lotka 26] A. J. Lotka. "The Frequency Distribution of Scientific Productivity." *J. of the Washington Acad. of Sci.* 16 (1926), 317.

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

[Watts and Strogatz 98] D. J. Watts and S. H. Strogatz. "Collective Dynamics of 'Small-World' Networks." *Nature* 393 (1998), 440–442.

[Zipf 49] G. K. Zipf. "Human Behavior and the Principle of Least Effort." New York: Hafner, 1949.

Béla Bollobás, Department of Mathematical Sciences, University of Memphis, Memphis, TN 38152 (bollobas@msci.memphis.edu)

Oliver Riordan, Royal Society Research Fellow, Department of Pure Mathematics and Mathematical Statistics, University of Cambridge, UK (o.m.riordan@dpmms.cam.ac.uk)

Received November 5, 2002; accepted December 11, 2002