Internet Mathematics Vol. 7, No. 1: 67–80

Graphs with Asymptotically Invariant Degree Sequences under Restriction

Joshua Cooper and Linyuan Lu

Abstract. Scaling-free graphs are often used to describe a class of graphs that have the self-similarity property. The degree sequences of many scaling-free graphs follow the power-law distribution. In this paper, we study the distributions of graphical degree sequences that are invariant under "scaling." We show that the invariant degree sequence must be a power-law distribution for sparse graphs if we ignore isolated vertices, or more generally, the vertices of degree less than a fixed constant k. We obtain a concentration result on the degree sequence of a random induced subgraph. The case of hypergraphs (or set systems) is also examined.

1. Introduction

What distribution of a graphical degree sequence is invariant under "scaling"? Are these graphs always power-law graphs? Quite a few recent papers use the term "scale-free networks" to refer to large sparse graphs formed from real-world data. Such graphs often exhibit power-law degree distributions. Namely, the number of vertices with degree d is roughly proportional to $d^{-\beta}$, for some positive β. However, the term "scale-free" is rarely defined in the literature, at

ISSN: 1542-7951 print

least in the rigorous mathematical sense. Furthermore, accounts in the literature of how power laws arise have been largely model-dependent. That is, a number of models of random-graph growth have been proposed that give rise, under circumstances of varying generality, to power-law degree distributions. The most popular growth model of this kind is the "preferential attachment" scheme, exemplified by [Aiello et al. 02, Barabási and Albert 99, Barabási et al. 00, Kleinberg et al. 99].

It is easy to show that power-law graphs are "scale-free." Here "scaling the graph down" means "taking an induced subgraph." Of course, subgraphs may look quite different from one another. Hence, we consider only the average behavior.

1.1. The Random Induced Subgraph G_p

For any $0 < p < 1$, let G_p be the induced subgraph of G on a random subset of vertices S. For each vertex v of G, v is in $V(G_p)$ with independent probability p.

There are some simple cases in which the graph G_p is similar to G. For example:

- Let G be a complete graph on n vertices. Then G_p is also a complete graph on around pn vertices.
- Let G be an empty graph on n vertices. Then G_p is also an empty graph on around pn vertices.
- For any constant $q \in (0, 1)$, let G be the random graph $G(n, q)$. Then G_p is also a random graph $G(m, q)$ over a randomly chosen set of size $m \sim pn$.

Crucially, these examples are not "real-world graphs," in the sense that graphs appearing "in nature" tend to be quite sparse. Most vertices have small degrees. To characterize this property, we use the following definition:

Definition 1.1. For a given sequence $\{\lambda_d\}_{d=0}^{\infty}$ satisfying $\sum_{d=0}^{\infty} \lambda_d = 1$, with $\lambda_d \geq 0$ for all $d \geq 0$, a sequence of graphs $\{G^n\}$ on n vertices is said to have *degree sequence with limit distribution* $\{\lambda_d\}_{d=0}^{\infty}$ if the number of vertices with degree d in G^n is $\lambda_d n + o(n)$ for each $d \geq 0$. We also say that $\{G^n\}$ has *limit distribution* $\{\lambda_d\}_{d=k}^{\infty}$ *for* $\sum_{d \geq k} \lambda_d \leq 1$ *if* G^n *has* $\lambda_d n + o(n)$ *vertices of degree* d *for each* $d \geq k$.

We consider two questions.

1. If the degree sequence of G in $\{G_n\}$ has a limit distribution, then for any fixed p, does the degree sequence of the random induced subgraph G_p also have a limit distribution?

2. For what distribution $\{\lambda_k\}_{k=0}^{\infty}$ is the limit distribution of the degree sequence of G_p essentially the same as the limit distribution of the degree sequence of G?

To answer the first question, we observe that a vertex of degree cn in G would badly affect the concentration of the degree sequence of G_p. On the other hand, using the vertex-exposure martingale, we can show that the degree sequence of G_p will have a limit distribution if

$$\sum_v \deg^2(v) = O(n^{2-\epsilon}).$$

This condition is satisfied, for example, if G has maximum degree bounded by $n^{1/2-\epsilon}$.

Suppose $a_0, a_1, a_2, \ldots,$ is the degree frequency sequence of a graph G, with a_d representing the number of vertices in G with degree d. What is the degree frequency sequence of G_p? If a vertex v survives in G_p, its degree has binomial distribution $B(d_G(v), p)$. There is no simple way to describe the joint distribution because of edge correlations. Nonetheless, the expected degree frequency sequence for G_p is easy to compute. Let $b_0, b_1, b_2, \ldots$ be the expected degree frequency sequence of the random induced subgraph G_p. We have

$$b_d = p \sum_{k \geq i} a_k \binom{k}{d} p^d (1-p)^{k-d}$$

for all $d = 0, 1, 2, \ldots$. Note that $\{b_d\}_{d\geq 0}$ depends linearly on $\{a_d\}_{d\geq 0}$. We can therefore normalize both sequences by dividing by n.

Therefore, from now on, we assume that each a_i is the fraction of the number of vertices with degree i in the graph G. More precisely, we consider a sequence of graphs G_n such that the number of vertices with degree d in G_n is $a_d n + o(n)$. We consider only sparse graphs such that

$$\sum_{i\geq 0} a_i = 1.$$

We have the following theorem.

Theorem 1.2. *For any integer $k > \beta > 1$, the degree frequency sequence starting at k defined by $a_d = C_\beta \binom{d-\beta}{d} + o_n(1)$ is scale-free. Moreover, if a graph G on n*

vertices such that

$$\sum_{v \in G} \deg(v)^2 = O(n^{2-\epsilon}) \tag{1.1}$$

for some $\epsilon > 0$ *has a scale-free degree sequence starting at* k*, then there is a* $\beta \in (1, k)$ *such that* $a_d = C_\beta \binom{d-\beta}{d} + o_n(1)$*. As a consequence, sparse graphs with scale-free degree sequences are power-law graphs.*

It is worth remarking that this manuscript can be read, in effect, as a response to the well-known paper [Stumpf et al. 05] and its authors' related publications. Although the present authors became aware of this work only after discovering the results below, it is clear that there is a very strong resemblance to their work. However, we offer a counterassertion to the authors' "subnets of scale-free networks are not scale-free," namely, "subnets of scale-free networks *are* scale-free, as long as one ignores vertices of suitably small-degree." We also take a somewhat different tack by studying, in particular, the asymptotic conditions under which scale-freeness holds.

This paper is organized as follows. In Section 2, we will derive scale-free degree sequences starting at $k = 0, 1$. The concentration result is proved in Section 3. The proof of our main theorem is given in Section 4. Scale-free set systems and remarks are given in Sections 5 and 6, respectively.

2. Scale-Free Degree Sequences

Let $A(x) = \sum_{i=0}^{\infty} a_i x^i$ be the generating function of $\{a_i\}_{i \geq 0}$, and let $B(x) = \sum_{i=0}^{\infty} b_i x^i$ be the generating function of $\{b_i\}_{i \geq 0}$. Both $A(x)$ and $B(x)$ converge on the interval $[-1, 1]$. We have

$$\begin{aligned}
B(x) &= \sum_{i=0}^{n} b_i x^i \\
&= \sum_{i=0}^{\infty} p \sum_{k \geq i} (a_k + o(1)) \binom{k}{i} p^i (1-p)^{k-i} x^i \\
&= p \sum_{k=0}^{\infty} a_k \sum_{i=0}^{k} \binom{k}{i} p^i (1-p)^{k-i} x^i + o(1) \cdot \sum_{k=0}^{\infty} \sum_{i=0}^{k} \binom{k}{i} p^i (1-p)^{k-i} x^i \\
&= p \sum_{k=0}^{\infty} a_k (1 - p + px)^k + o(1) \sum_{k=0}^{\infty} (1 - p + px)^k \\
&= pA(1 - p + px) + \frac{o(1)}{1 - x}.
\end{aligned}$$

2.1. Scale-Free Degree Sequences Starting at 0

A naive way to define scale-freeness is to require

$$b_i = f(p)a_i + o(1) \quad \text{for all } i \geq 0,$$

where $f(p)$ is a quantity depending only on p.

Equivalently, for any $x \in [-1, 1]$ and $p \in (0, 1)$, we have

$$pA(1 - p + px) = f(p)A(x). \tag{2.1}$$

To solve (2.1), let $x = 1$. We get $pA(1) = f(p)A(1)$. Thus $f(p) = p$. We have

$$A(1 - p + px) = A(x).$$

Let $x = 0$. We have $A(0) = A(1 - p)$. Therefore,

$$A'(0) = \lim_{x \to 0} \frac{A(x) - A(0)}{x} = \lim_{x \to 0} \frac{A(1 - p + px) - A(1 - p)}{x} = pA'(1 - p).$$

Since this holds for any $p \in (0, 1)$, we have

$$\begin{aligned} A(p) &= A(0) + \int_{1-p}^{1} A'(1 - p)\, dp = A(0) + \int_{1-p}^{1} \frac{A'(0)}{p}\, dp \\ &= A(0) - A'(0) \ln(1 - p). \end{aligned}$$

Thus,

$$A(x) = A(0) - A'(0) \ln(1 - x).$$

We have

$$\begin{aligned} A(1 - p + px) &= A(0) - A'(0) \ln(p - px) = A(0) - A'(0)(\ln p + \ln(1 - x)) \\ &= A(x) - A'(0) \ln p. \end{aligned}$$

This forces $A'(0) = 0$. The only solution for (2.1) is $A(x) \equiv A(0)$ (the constant function, corresponding to a graph with no edges). This solution is not interesting.

2.2. Scale-Free Degree Sequences Starting at 1

In many cases, we do not care about the number of isolated vertices. We require only that

$$b_d = f(p)a_d + o(1) \quad \text{for all } d \geq 1,$$

where $f(p)$ is a quantity depending only on p.

Equivalently, for any $p \in (0, 1)$ and $x \in [-1, 1]$, we have

$$f(p)(A(x) - A(0)) = p(A(1 - p + px) - A(1 - p)).$$

Take the derivative with respect to x on both sides. We then have, for any $p \in (0,1)$ and $x \in (-1,1)$,

$$f(p)A'(x) = p^2 A'(1-p+px). \tag{2.2}$$

Let $\alpha = \int_0^1 \frac{f(p)}{p^2}dp$ be a positive constant. Divide both sides of (2.2) by p^2 and integrate with respect to p from 0 to 1. We obtain

$$\alpha A'(x) = \int_0^1 A'(1-p+px)dp = \frac{A(1)-A(x)}{1-x} = \frac{1-A(x)}{1-x}.$$

Rewriting this expression yields

$$\frac{A'(x)}{1-A(x)} = \frac{1}{\alpha(1-x)}.$$

Now integrate with respect to x from 0 to x. We get

$$\ln\frac{1-A(0)}{1-A(x)} = -\frac{1}{\alpha}\ln(1-x).$$

Therefore, we have

$$A(x) = 1-(1-A(0))(1-x)^{1/\alpha}. \tag{2.3}$$

It is easy to verify that (2.3) satisfies (2.2) with $f(p) = p^{1+1/\alpha}$.

We do not care about $A(0) = a_0$, the number of isolated vertices. Hence the solution is uniquely determined by the parameter α up to a constant factor. For $d \geq 1$, we have

$$a_d = (1-a_0)\binom{1/\alpha}{d}(-1)^{d+1} = -(1-a_0)\binom{d-1/\alpha-1}{d} = O\left(d^{-(1+1/\alpha)}\right).$$

In other words, the degree frequency sequence follows a power-law distribution with exponent $\beta = 1+1/\alpha$. However, not all a_d are positive. Particularly, if $\beta > 2$, then there are negative terms a_d, $d \geq 1$.

3. Concentration

Since we know that the only degree sequences that are scale-free *in expectation* have power-law limit distributions, it is crucial to show that such graphs have degree sequences that are close to their means with high probability.

Theorem 3.1. *Suppose that $\{G^n\}_{n=1}^{\infty}$ is a sequence of graphs on $n \to \infty$ vertices with degree sequence of limit distribution $\{\lambda_d\}_{d=k}^{\infty}$. Further suppose that*

$$\sum_{v\in G}\deg(v)^2 = O(n^{2-\epsilon})$$

for some $\epsilon > 0$*. Then the degree sequence of* G_p^n *also has a limit distribution* $\{\lambda'_d\}_{d=k}^{\infty}$*.*

Proof. Let $a_d = a_d(n)$ be the fraction of vertices of degree d in G^n and let $b_d = b_d(n)$ be the fraction of vertices of degree d in G_p^n. Let $\lambda'_d = \mathbf{E}(b_d)$. Clearly it suffices to show that b_d is concentrated about its expectation.

To that end, we apply the Azuma–Hoeffding inequality to the "vertex exposure" martingale. In particular, consider the following process. Fix $d \geq k$, order the vertices of G^n as $v_1, \dots, v_n$, and let A_m denote the event that $v_m \in G_p^n$. Let $X_0 = \mathbf{E}[b_d n]$, and let $X_{m+1} = \mathbf{E}[X_m \mid A_{m+1}]$. That is, at stage m, we "expose" vertex m and recalculate the expected number of vertices of degree d based on the new information concerning whether $v_m \in G_p^n$. It is easy to see that this is a martingale, and furthermore, that $|X_{m+1} - X_m| \leq \deg(v_{m+1}) + 1$, where $\deg(\cdot)$ denotes degree in G. Since $b_d n = X_n$, we may apply the Azuma–Hoeffding inequality to get

$$\mathbf{P}\left[|b_d - \lambda'_d| \geq t/n\right] \leq \exp\left(\frac{-t^2}{2\sum_{m=1}^{n}(\deg(v_m)+1)^2}\right)$$

for $t \geq 0$. Since $\sum_{m=1}^{n} \deg(v_m)^2 = O(n^{2-\epsilon})$ and

$$\sum_{m=1}^{n} \deg(v_m) \leq \sqrt{n}\left(\sum_{m=1}^{n} \deg(v_m)^2\right)^{1/2} = O(n^{3/2-\epsilon/2})$$

by Cauchy–Schwarz, we can set $t = n^{1-\epsilon/4}$, thereby obtaining

$$\mathbf{P}\left[|b_d - \lambda'_d| \geq t/n\right] \leq e^{-\Omega(n^{\epsilon/2})}.$$

Let $t' = t/n = n^{-\epsilon/4}$. Then, since

$$\sum_{n=1}^{\infty} \mathbf{P}\left[\bigwedge_{d=k}^{n} (|b_d - \lambda'_d| \geq t')\right] \leq \sum_{n=1}^{\infty} n e^{-n^{\epsilon/2}} < \infty,$$

the Borel–Cantelli lemma implies that asymptotically almost surely, $|b_d - \lambda'_d| \leq t' = o(n)$ for all $d \geq k$. $\square$

4. Proof of the Main Theorem

Proof of Theorem 1.2. Suppose $a_d = C_\beta\binom{d-\beta}{d} + o_n(1)$ for all $d \geq k > \beta > 1$. We have

$$\begin{aligned}
b_d &= (1+o_n(1))\sum_{i=d}^{\infty} a_i \binom{i}{d} p^d (1-p)^{i-d} \\
&= (1+o_n(1)C_\beta \sum_{i=d}^{\infty} \binom{i-\beta}{i}\binom{i}{d} p^d (1-p)^{i-d} \\
&= (1+o_n(1))C_\beta p^d \sum_{j=0}^{\infty} \binom{d+j-\beta}{d+j}\binom{d+j}{d} (1-p)^j \\
&= (1+o_n(1))C_\beta p^d \binom{d-\beta}{d} \sum_{j=0}^{\infty} \binom{j-(\beta-d)}{j} (1-p)^j \\
&= (1+o_n(1))C_\beta p^d \binom{d-\beta}{d} \sum_{j=0}^{\infty} \binom{\beta-d}{j} (p-1)^j \\
&= (1+o_n(1))C_\beta p^d \binom{d-\beta}{d} p^{\beta-d} \\
&= p^\beta a_d + o_n(1).
\end{aligned}$$

Thus, the degree sequence is scale-free.

Now we assume that inequality (1.1) holds. By Theorem 3.1, the degree sequence of G_p^n has a limit distribution. Now we assume that the degree sequence distribution, considering only degrees at least k, is scale-free. That is,

$$b_d = f(p)a_d + o_n(1) \quad \text{for all } d \geq k,$$

where $f(p)$ is a quantity depending only on p. Or equivalently, for any $p \in (0,1)$ and $x \in [-1,1]$, we have

$$f(p)\left(A(x) - \sum_{d=0}^{k-1} a_d x^d\right) = p\left(A(1-p+px) - \sum_{d=0}^{k-1} a_d x^d \binom{k}{d} p^d (1-p)^{k-d}\right). \tag{4.1}$$

Take the kth derivative with respect to x on both sides to get rid of all terms of degree up to $k-1$. We have, for any $p \in (0,1)$ and $x \in (-1,1)$,

$$f(p)A^{(k)}(x) = p^{k+1}A^{(k)}(1-p+px). \tag{4.2}$$

Let

$$\alpha_k = \int_0^1 \frac{f(p)}{p^{k+1}}\, dp.$$

Similar arguments to those used in the cases $k = 0, 1$ (in Section 2) show that the solution of (4.2) is of the form

$$A^{k-1}(x) = C_1 - C_2(1-x)^{1/\alpha}.$$

If we then integrate $k-1$ times with respect to x, the result is

$$A(x) = P_k(x) - C(1-x)^{k+1/\alpha_k}. \tag{4.3}$$

Here $P_k(x)$ is a polynomial in x of degree $k-1$. It is easy to verify that (4.3) is the solution of (4.1) with $f(p) = p^{\alpha_k + k}$. Let $\beta = k + 1/\alpha_k$. For any $d \geq k$, we have

$$a_d = C\binom{d-\beta}{d}.$$

If we set

$$C = C_\beta = \left(\sum_{d \geq \lceil \beta \rceil} \binom{d-\beta}{d}\right)^{-1},$$

then the a_d are positive for $d > \beta$. Note that $\operatorname{sgn}(C_\beta) = (-1)^{\lfloor \beta \rfloor}$. □

5. Scale-Free Set Systems

Many power-law graphs such as the collaboration graph and the Hollywood graph are actually better modeled by set systems (or hypergraphs) than by graphs. For example, in the *Math Reviews* database, each published item has one or more authors. The family of all papers considered as collections of authors forms a set system. The collaboration graph captures only part of the information in this set system. Here we quote from the Erdős number project [Grossman et al. 11]:

> There are about 1.9 million authored items in the *Math Reviews* database, by a total of about 401,000 different authors.... Approximately 62.4% of these items are by a single author, 27.4% by two authors, 8.0% by three authors, 1.7% by four authors, 0.4% by five authors, and 0.1% by six or more authors.

In this example, the distribution of set sizes follows a power-law distribution (see Figure 1). Is this just a coincidence? Is a "scale-free" distribution of a set system always a power-law distribution?

Motivated by this example and "scale-free" graphs, we consider the following problem. For a set system $\mathcal{F}$ and any probability $p \in (0,1)$, the random sub–set

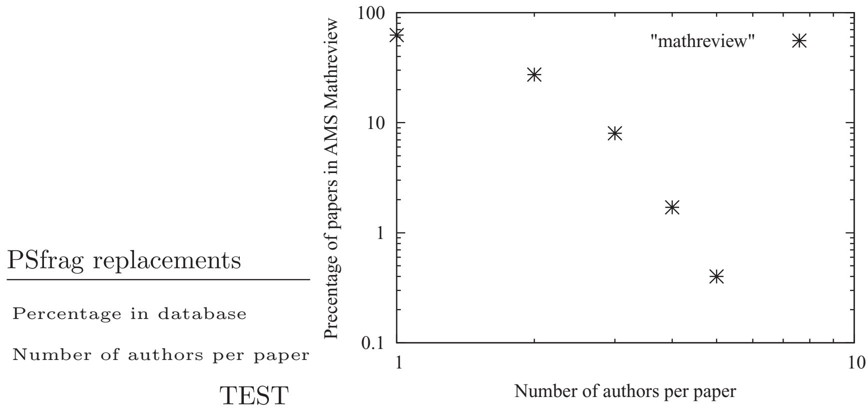

Figure 1. The percentages of articles by number of authors in the *Math Reviews* database.

system $\mathcal{F}_p$ is chosen by independently removing vertices with probability $1-p$ and reducing the sets to their remaining elements.

Problem 5.1. For what sequence of set sizes in a set system $\mathcal{F}$ is the sequence of the set sizes in random sub–set system $\mathcal{F}_p$ essentially the same as the original sequence up to a scaling factor?

For $i \geq 1$, let a_i be the number of i-sets in $\mathcal{F}$ and let b_i be the number of i-sets in $\mathcal{F}_p$. We are asking whether there is a function $f(p)$ such that

$$b_i = f(p)a_i + o(n)$$

for all $i \geq k$. Here k is a small positive integer.

Since the expected value $\mathbf{E}(b_i)$ satisfies

$$\mathbf{E}(b_i) = \sum_{j \geq i} a_j \binom{j}{i} p^i (1-p)^{j-i},$$

it is necessary to have

$$\sum_{j \geq i} a_j \binom{j}{i} p^i (1-p)^{j-i} = f(p)a_i$$

for all $i \geq k$.

Let $A(x) = \sum_i a_i x^i$ be the generating function. For any $p \in (0,1)$ and $x \in [-1,1]$, we have

$$f(p)\left(A(x) - \sum_{d=0}^{k-1} a_d x^d\right) = \left(A(1-p+px) - \sum_{d=0}^{k-1} a_d x^d \binom{k}{d} p^d (1-p)^{k-d}\right).$$

This is essentially the same equation as (4.1). Thus we have the following theorem.

Theorem 5.2. *If the sequence of set sizes in a set system starting at $k > 1$ is scale-free, then there are constants $\beta \in (1,k)$ and C such that the number of i-sets in this set system is $C_\beta \binom{i-\beta}{i} n + o(n)$ for all $i \geq k$.*

6. Remarks and Questions

Note that the results of the preceding sections have a probabilistic interpretation. Suppose that for each n, we have a probability distribution $\mathcal{G}$ over graphs on n vertices with the property that the expected number of vertices of degree d is a_d. Then what must $\mathbf{E}[a_d]$ be if when G is sampled from $\mathcal{G}$ and a random subgraph G_p is taken, the expected number b_d of vertices of degree d is the same as a_d after scaling so that $\sum_d a_d = \sum_d b_d$? The above analysis provides the answer: the expectation of a_d must be a power law in d.

It is natural to ask the following: if the variance of the b_d is scaled as the square of the scaling factor for the expectations, then what must $\sigma^2(a_d)$ be? In fact, one can ask the same question of all moments, leading to the following open problem:

Problem 6.1. Fix $p \in (0,1)$. Let G be drawn from a probability distribution $\mathcal{G}$ on graphs with n vertices. Suppose that a_d, $d \geq 0$, is the number of vertices of degree d in G, and b_d, $d \geq 0$, is the number of vertices of degree d in G_p. For which distributions $\mathcal{G}$ is it true that there exists some $c(p) \in \mathbb{R}$ such that $\{a_d\}_{d \geq k}$ and $\{c(p) b_d\}_{d \geq k}$ have approximately the same distribution for large n? Is it possible to find such $\mathcal{G}$ for all $p \in (0,1)$ simultaneously?

Currently, the exponents of power laws of "real-world" scale-free networks is estimated in a rather ad hoc fashion, usually using a regression on the log-log plot of frequency vs. degree after removing the extremes of the data. If it were possible to describe scale-free distributions exactly, then it would make sense to ask the following very practical question:

Problem 6.2. Find an unbiased estimator for the exponent of a power-law degree distribution.

For the matter of the variance of the a_d, we note that at least for $\beta \in (1, 2)$, the following must be true:

$$p^{2\beta}\sigma^2(a_d) = \sum_k \binom{k}{d}^2 p^{2d}(1-p)^{2k-2d}\left(\sigma^2(a_k) + \binom{k-\beta}{k}\right) - p^{\beta-1}\binom{d-\beta}{d}.$$

This statement can be proven by applying the formula

$$\sigma^2\left(\sum_{i=1}^{N} X_i\right) = \mathbf{E}[X_1]^2\sigma^2(N) + \mathbf{E}[N]\sigma^2(X_1)$$

for i.i.d. variables X_i and an independent variable N taking on nonnegative integer values.

We also ask, what can be proved by extending the definition of scale-freeness to hypergraphs? We believe that the situation is very similar to that of graphs when the hypergraphs being considered are uniform (with edges removed whenever at least one of their vertices is removed). Perhaps the answer lies in a more refined description of scale-freeness. For example, consider the quantity $a_H(G)$, the number of occurrences of H as an induced subgraph of G. Suppose that $a_H(G)/n \to \alpha_H$ for each H and some $\alpha_H \in \mathbb{R}^+$, and that this sequence is scale-free, i.e.,

$$a_H(G_p) \propto a_H(G)$$

for any fixed p with $0 < p < 1$ and H varying over all graphs on at least k vertices. Then what must G look like?

Acknowledgments. Linyuan Lu's research was supported in part by NSF grant DMS 0701111.

References

[Aiello et al. 00] W. Aiello, F. Chung, and L. Lu. "A Random Graph Model for Massive Graphs." In *Proceedings of the Thirty-Second Annual ACM Symposium on Theory of Computing*, pp. 171–180, 2000.

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

[Aiello et al. 02] W. Aiello, F. Chung, and L. Lu. "Random Evolution in Massive Graphs." In *Handbook of Massive Data Sets*, volume 2, edited by J. Abello et al., pp. 97–122. Dordrecht: Kluwer Academic Publishers, 2002.

[Alon and Spencer 00] N. Alon and J. H. Spencer. *The Probabilistic Method*, second edition. New York: Wiley-Interscience, 2000.

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

[Barabási et al. 00] A. Barabási, R. Albert, and H. Jeong. "Scale-Free Characteristics of Random Networks: The Topology of the World Wide Web." *Physica a* 281 (2000), 69–77.

[Bollabás and Riordan 03] B. Bollabás and O. Riordan. "Robustness and Vulnerability of Scale-free Random Graphs." *Internet Mathematics* 1:1 (2003), 1–35.

[Broder et al. 00] A. Broder, R. Kumar, F. Maghoul, P. Raghavan, S. Rajagopalan, R. Stata, A. Tompkins, and J. Wiener. "Graph Structure in the Web." *Proceedings of the WWW9 Conference*, Amsterdam, May 2000.

[Chung 97] F. Chung. *Spectral Graph Theory*. Providence: AMS Publications, 1997.

[Chung and Lu 01] F. Chung and L. Lu. "The Diameter of Random Sparse Graphs." *Advances in Applied Math.* 26 (2001), 257–279.

[Chung and Lu 02a] F. Chung and L. Lu. "Connected Components in a Random Graph with Given Degree Sequences." *Annals of Combinatorics* 6 (2002), 125–145.

[Chung and Lu 02b] F. Chung and L. Lu. "The Average Distance in Random Graphs with Given Expected Degrees." *Proceedings of the National Academy of Sciences* 99, (2002), 15879–15882.

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

[Chung et al. 03a] F. Chung, L. Lu, and V. Vu. "The Spectra of Random Graphs with Given Expected Degrees." *Proceedings of National Academy of Sciences* 100:11 (2003), 6313–6318.

[Chung et al. 03b] F. Chung, L. Lu, T. G. Dewey, and D. J. Galas. "Duplication Models for Biological Networks." *Journal of Computational Biology* 10:5 (2003), 677–688.

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

[Cooper et al. 03] C. Cooper, A. Frieze, and J. Vera. "Random Vertex Deletion in a Scale Free Random Graph." *Internet Mathematics* 1:4 (2003), 463–483.

[Erdős and Gallai 61] P. Erdős and T. Gallai. "Gráfok előírt fokú pontokkal" (Graphs with Points of Prescribed Degrees, in Hungarian). *Mat. Lapok* **11** (1961), 264–274.

[Erdős and Rényi 59] P. Erdős and A. Rényi. "On Random Graphs. I." *Publ. Math. Debrecen* 6 (1959), 290–291.

[Grossman et al. 11] J. Grossman, P. Ion, and R. De Castro. "The Erdős Number Project." Available online (http://www.oakland.edu/enp/), Accessed February 25, 2011.

[Janson et al. 00] S. Janson, T. Łuczak, and A. Rucinski. *Random Graphs*. New York: Wiley-Interscience, 2000.

[Jeong et al. 00] H. Jeong, B. Tomber, R. Albert, Z. Oltvai, and A. L. Babárasi. "The Large-Scale Organization of Metabolic Networks." *Nature* 407 (2000), 378–382.

[Kim and Vu 00] J. H. Kim and V. Vu. "Concentration of Multi-variate Polynomials and Its Applications." *Combinatorica* 20:3 (2000), 417–434.

[Kleinberg et al. 99] J. Kleinberg, S. R. Kumar, P. Raphavan, S. Rajagopalan, and A. Tomkins. "The Web as a Graph: Measurements, Models and Methods." *Proceedings of the International Conference on Combinatorics and Computing*, 1999.

[Lotka 26] A. J. Lotka. "The Frequency Distribution of Scientific Productivity." *Journal of the Washington Academy of Sciences* 16 (1926), 317–323.

[Lu 01] L. Lu. "The Diameter of Random Massive Graphs." In *Proceedings of the Twelfth ACM-SIAM Symposium on Discrete Algorithms*, pp. 912–921, 2001.

[McDiarmid 98] C. McDiarmid. "Concentration." In *Probabilistic Methods for Algorithmic Discrete Mathematics*, Algorithms Combin. 16, pp. 195–248. Berlin: Springer, 1998.

[Milgram 67] S. Milgram. "The Small World Problem." *Psychology Today* 2 (1967), 60–67.

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

[Stumpf et al. 05] M. P. H. Stumpf, C. Wiuf, and R. M. May. "Subnets of Scale-Free Networks Are Not Scale-Free: Sampling Properties of Networks." *Proceedings of the National Academy of Sciences* 102 (2005), 4221–4224.

Joshua Cooper, University of South Carolina, Columbia, SC 29208 (cooper@math.sc.edu)

Linyuan Lu, University of South Carolina, Columbia, SC 29208 (lu@math.sc.edu)

Received March 14, 2007; accepted November 22, 2010.