The maximum degree of the rth power of a sparse random graph

Alan Frieze* and Aditya Raut
Department of Mathematical Sciences
Carnegie Mellon University
Pittsburgh PA-15213

Abstract

Let $G_{n,p}^r$ denote the rth power of the random graph $G_{n,p}$, where $p = c/n$ for a positive constant c. We prove that w.h.p. the maximum degree $\Delta\left(G_{n,p}^r\right) \sim \frac{\log n}{\log_{(r+1)} n}$. Here $\log_{(k)} n$ indicates the repeated application of the log-function k times. So, for example, $\log_{(3)} n = \log\log\log n$.

1 Introduction

The rth power G^r of a graph G is obtained from G by adding edges for all pairs of vertices at distance r or less from each other. Powers of graphs arise naturally in various contexts, e.g. in the study of Shannon capacity. In the context of random graphs, there has been little research.

Let $\log_{(k)} n$ indicate the repeated application of the log-function k times. So, for example, $\log_{(3)} n = \log\log\log n$.

One basic property of a class of graphs is their degree sequence. In particular, the maximum degree is of particular interest. Garapaty, Lokshtanov, Maji and Pothen [4] proved that if $p = c/n$ where $c > 0$ is constant, then w.h.p. the maximum degree $\Delta(G_{n,p}^r) = \Theta_r\left(\frac{\log n}{\log_{(r+1)} n}\right)$. The hidden multiplicative factor being in the range $[0.05 \cdot 2^{-r}, 6]$. We strengthen this and prove

Theorem 1. *Let $p = c/n$, where $c > 0$ is a constant and let $r \geq 2$ be a fixed positive integer. Then, w.h.p. $\Delta(G_{n,p}^r) \sim \frac{\log n}{\log_{(r+1)} n}$.*

(The case $r = 1$ is well-known, see e.g. Theorem 3.4 of [3].)

Remark 1. *The value of c does not contribute to the main term in the claim of Theorem 1. Thus we would expect that we could replace $p = c/n$ by $p \leq \omega/n$ for some slowly growing function $\omega = \omega(n) \to \infty$.*

*Research supported in part by NSF grant DMS1952285

2 Proof of Theorem 1

Given $v \in [n] = V(G), G = G_{n,p}$, let $N_t(v)$ denote the set of vertices of G at distance t from v. Let $d_t(v) := |N_t(v)|$. The probability that $|N_t(v)| = \ell_t, t = 1, 2, \ldots, r$ can be calculated as follows. Let the ℓ_i neighbors of v at distance i have $k_1^{(i)}, k_2^{(i)}, ..., k_{\ell_i}^{(i)}$ neighbors at distance $i+1$ respectively. Then,

$$\mathbb{P}\left[d_i(v) = \ell_i \ \forall i = 1, 2, ..., r\right]$$

$$= \binom{n-1}{\ell_1} p^{\ell_1}(1-p)^{n-1-\ell_1} \left(\sum_{\substack{k_1^{(1)}, k_2^{(1)}, ..., k_{\ell_1}^{(1)} \\ k_1^{(1)} + ... + k_{\ell_1}^{(1)} = \ell_2}} \prod_{i=1}^{\ell_1} \binom{n-(1+\ell_1)}{k_i^{(1)}} p^{k_i^{(1)}} (1-p)^{n-(1+\ell_1)-k_i^{(1)}} \times \cdots \right.$$

$$\left. \cdots \left(\sum_{\substack{k_1^{(r-1)}, k_2^{(r-1)}, ..., k_{\ell_{r-1}}^{(r-1)} \\ k_1^{(r-1)} + ... + k_{\ell_{r-1}}^{(r-1)} = \ell_r}} \prod_{i=1}^{\ell_{r-1}} \binom{n-(1+\ell_1+...+\ell_{r-1})}{k_i^{(r-1)}} p^{k_i^{(r-1)}} (1-p)^{n-(1+\ell_1+...+\ell_{r-1})-k_i^{(r-1)}} \right) \right)$$

Now it is well-known that $\Delta(G_{n,p}) = o(\log n)$ w.h.p., see for example [3]. In consequence $\ell_i = o(\log^i n)$ for $i \leq r$. We use the approximation $\binom{n}{\ell} = \frac{n^\ell}{\ell!} \cdot \left(1 + \mathcal{O}\left(\frac{\ell^2}{n}\right)\right)$ and simplify

$$\binom{n}{\ell} p^\ell (1-p)^{n-\ell} = \frac{n^\ell}{\ell!} \cdot \frac{c^\ell}{n^\ell} \cdot \left(1 - \frac{c}{n}\right)^{n-\ell} \cdot \left(1 + \mathcal{O}\left(\frac{\ell^2}{n}\right)\right) = \frac{c^\ell e^{-c}}{\ell!} \cdot \left(1 + \mathcal{O}\left(\frac{\ell^2}{n}\right)\right)$$

Simplifying all terms this way, we have

$$\mathbb{P}\left[d_i(v) = \ell_i \ \forall i = 1, 2, ..., r\right]$$

$$= \frac{c^{\ell_1} e^{-c}}{\ell_1!} \left(\sum_{\substack{k_1^{(1)}, k_2^{(1)}, ..., k_{\ell_1}^{(1)} \\ k_1^{(1)} + ... + k_{\ell_1}^{(1)} = \ell_2}} \prod_{i=1}^{\ell_1} \frac{c^{k_i^{(1)}} e^{-c}}{k_i^{(1)}!} \times \cdots \left(\sum_{\substack{k_1^{(r-1)}, k_2^{(r-1)}, ..., k_{\ell_{r-1}}^{(r-1)} \\ k_1^{(r-1)} + ... + k_{\ell_{r-1}}^{(r-1)} = \ell_r}} \prod_{i=1}^{\ell_{r-1}} \frac{c^{k_i^{(r-1)}} e^{-c}}{k_i^{(r-1)}!} \right) \right) \cdot \left(1 + \mathcal{O}\left(\frac{\sum_{i=1}^{r} \ell_i^2}{n}\right)\right)$$

$$= \frac{c^{\ell_1} e^{-c}}{\ell_1!} \left(\frac{c^{\ell_2} e^{-c\ell_1}}{\ell_2!} \sum_{\substack{k_1^{(1)}, k_2^{(1)}, ..., k_{\ell_1}^{(1)} \\ k_1^{(1)} + ... + k_{\ell_1}^{(1)} = \ell_2}} \frac{\ell_2!}{\prod_{i=1}^{\ell_1} k_i^{(1)}!} \times \cdots \left(\frac{c^{\ell_r} e^{-c\ell_{r-1}}}{\ell_r!} \sum_{\substack{k_1^{(r-1)}, k_2^{(r-1)}, ..., k_{\ell_{r-1}}^{(r-1)} \\ k_1^{(r-1)} + ... + k_{\ell_{r-1}}^{(r-1)} = \ell_r}} \frac{\ell_r!}{\prod_{i=1}^{\ell_{r-1}} k_i^{(r-1)}!} \right) \right) \cdot \left(1 + \mathcal{O}\left(\frac{\sum_{i=1}^{r} \ell_i^2}{n}\right)\right)$$

We now use $\sum_{\substack{k_1, ..., k_t \\ k_1 + ... + k_t = m}} \binom{m}{k_1, ..., k_t} = t^m$ to obtain

$$\mathbb{P}\left[d_i(v) = \ell_i \ \forall i = 1, 2, ..., r\right] = \frac{c^{\ell_1} e^{-c}}{\ell_1!} \cdot \left(\frac{c^{\ell_2} e^{-c\ell_1}}{\ell_2!} \cdot \ell_1^{\ell_2} \times \cdots \left(\frac{c^{\ell_r} e^{-c\ell_{r-1}}}{\ell_r!} \cdot \ell_{r-1}^{\ell_r}\right)\right) \cdot \left(1 + \mathcal{O}\left(\frac{\sum_{i=1}^{r} \ell_i^2}{n}\right)\right)$$

$$= \frac{c^{(\ell_1 + ... + \ell_r)} e^{-c(1+\ell_1+...+\ell_{r-1})}}{\ell_1!\, \ell_2! \, ...\ell_r!} \cdot \ell_1^{\ell_2} \cdot \ell_2^{\ell_3} \cdots \ell_{r-1}^{\ell_r} \cdot \left(1 + \mathcal{O}\left(\frac{\sum_{i=1}^{r} \ell_i^2}{n}\right)\right)$$

The exact probability that we are interested in is the degree of v being d in G^r, i.e. $\sum_{i=1}^{r} d_i(v) = d$.
We aim to show that $d \sim \frac{\log n}{\log_{(r+1)} n}$ happens with high probability.

$$\mathbb{P}\left[\sum_{i=1}^{r} d_i(v) = d\right] = \sum_{\substack{\ell_1,...,\ell_r \\ \ell_1+...+\ell_r = d}} \frac{c^d e^{-c(1+d-\ell_r)}}{\ell_1!\, \ell_2! \, ...\ell_r!} \cdot \ell_1^{\ell_2} \cdot \ell_2^{\ell_3} \cdots \ell_{r-1}^{\ell_r} \cdot \left(1 + \mathcal{O}\left(\frac{\sum_{i=1}^{r} \ell_i^2}{n}\right)\right)$$

$$= \left(\sum_{\substack{\ell_1,...,\ell_r \\ \ell_1+...+\ell_r = d}} u_{\ell_1,...,\ell_r}\right)\left(1 + \mathcal{O}\left(\frac{d^2}{n}\right)\right),$$

where

$$u_{\ell_1,...,\ell_r} = \frac{c^d e^{-c(1+d-\ell_r)}}{\ell_1!\, \ell_2! \, ...\ell_r!} \cdot \ell_1^{\ell_2} \cdots \ell_{r-1}^{\ell_r}. \tag{1}$$

For completeness, define $\ell_0 = 1$. Using Stirling's approximation $\log(n!) = n \log n - n + \mathcal{O}(\log n)$, we have

$$\log u_{\ell_1,...,\ell_r} = d \log c - c(1 + d - \ell_r) - \sum_{i=1}^{r} \ell_i \log \frac{\ell_i}{\ell_{i-1}} + \mathcal{O}(d) \tag{2}$$

The following lemma bounds the sum in (2):

Lemma 2. *For $\ell_0 = 1$ and $\ell_1, ..., \ell_r \in \mathbb{N}$ such that $\sum_{i=1}^{r} \ell_i = d$, we have $\min \sum_{i=1}^{r} \ell_i \log \frac{\ell_i}{\ell_{i-1}} \geq d \log_{(r)} d + \mathcal{O}(d)$, for sufficiently large d.*

Proof. We proceed by induction on r. For $r = 1$, the result holds since we have $\ell_1 = d$, implying that $\ell_1 \log \frac{\ell_1}{\ell_0} = d \log_{(1)} d$. Assume that the result holds for $r - 1$.

Case 1: $(d - \ell_r) \log_{(r-1)}(d - \ell_r) \geq d \log_{(r)} d$**:**
Because $\sum_{i=1}^{r-1} \ell_i = d - \ell_r$, from the induction hypothesis we have $\sum_{i=1}^{r-1} \ell_i \log \frac{\ell_i}{\ell_{i-1}} \geq (d - \ell_r) \log_{(r-1)}(d - \ell_r)$. So this case is done.

Case 2: $(d - \ell_r) \log_{(r-1)}(d - \ell_r) < d \log_{(r)} d$**:**
For d sufficiently large, we have $\frac{d}{10} \log_{(r-1)} \frac{d}{10} > d \log_{(r)} d$. Hence $\ell_r \geq \frac{9}{10} \cdot d$, implying $\ell_{r-1} \leq \frac{d}{10}$ and $\frac{\ell_r}{\ell_{r-1}} \geq 9$.

We now use the method of Lagrange multipliers. But first we deal with the constraints $\ell_i \geq 0$ for $i = 1, 2, \ldots, r$. If $\ell_i = 0$ and $\ell_{i+1} \neq 0$ then $\sum_{i=1}^{r} \ell_i \log \frac{\ell_i}{\ell_{i-1}} = \infty$. If $\ell_i = \ell_{i+1} = \cdots = \ell_r = 0$ then $\sum_{i=1}^{r} \ell_i \log \frac{\ell_i}{\ell_{i-1}} = \sum_{i=1}^{i-1} \ell_i \log \frac{\ell_i}{\ell_{i-1}}$ and the result follows by induction. So, in effect there is one constraint: $\sum_{i=1}^{r} \ell_i = d$.

Define $\mathcal{L}(\ell_1, ..., \ell_r, \lambda) = \sum_{i=1}^{r} \ell_i \log \frac{\ell_i}{\ell_{i-1}} + \lambda \left(\sum_{i=1}^{r} \ell_i - d \right)$. By the Lagrange multiplier theorem, the minima must satisfy $\frac{\partial \mathcal{L}}{\partial \ell_i} = 0$ for all i. Notice that $\frac{\partial \mathcal{L}}{\partial \ell_i} = 1 + \log \frac{\ell_i}{\ell_{i-1}} - \frac{\ell_{i+1}}{\ell_i} + \lambda$ for $1 \leq i < r$, and $\frac{\partial \mathcal{L}}{\partial \ell_r} = 1 + \log \frac{\ell_r}{\ell_{r-1}} + \lambda$. Let $p_i := \frac{\ell_i}{\ell_{i-1}}$. From $\frac{\partial \mathcal{L}}{\partial \ell_r} = 0$, we have $1 + \log p_r + \lambda = 0$ which implies that $p_r = e^{-(\lambda+1)}$. For $1 \leq i < r$, from $\frac{\partial \mathcal{L}}{\partial \ell_i} = 0$ we have $1 + \log p_i - p_{i+1} + \lambda = 0$ which implies that $p_i = e^{p_{i+1}-(1+\lambda)} = p_r \cdot e^{p_{i+1}} \geq p_r p_{i+1}$. Thus, we can iteratively obtain the exact expressions for $p_1, p_2, ..., p_{r-1}$ in terms of p_r. Now recall that $p_r \geq 9$ from induction hypothesis, hence $p_i \geq 9$ for all i.

Since $\ell_0 = 1$, $\ell_1 = p_1$. Now $\ell_i = p_i \cdot p_{i-1} \cdots p_1$, for all $i = 1, 2, ..., r$ and then $\frac{\ell_r}{\ell_i} = p_r p_{r-1} ... p_{i+1} \geq 9^{r-i}$. Now $\sum_{i=1}^{r} \ell_i = d$ and so clearly $d > \ell_r$. Moreover, $d = \ell_r \left(\sum_{i=1}^{r} \frac{\ell_i}{\ell_r} \right) \leq \ell_r \left(\sum_{i=1}^{r} \frac{1}{9^{r-i}} \right) < \ell_r \left(\sum_{j=0}^{\infty} \frac{1}{9^j} \right) = \frac{9}{8} \cdot \ell_r$. We can thus assume that $d = c_1 \cdot \ell_r$ for some $c_1 \in (1, 9/8)$. So,

$$\begin{aligned} d &= c_1 \cdot p_r \cdot p_{r-1} \cdots p_1 \\ &= c_1 \cdot p_r \cdot \left(p_r e^{p_r}\right) \cdot \left(p_r e^{p_r e^{p_r}}\right) \cdots \left(p_r \underbrace{e^{p_r e^{\cdots e^{p_r}}}}_{\substack{\text{exponential tower} \\ \text{of height } r-1}} \right) \\ &= c_1 \cdot p_r^r \cdot \left(e^{p_r} e^{p_r e^{p_r}} \cdots \underbrace{e^{p_r e^{\cdots e^{p_r}}}}_{\substack{\text{exponential tower} \\ \text{of height } r-1}} \right) \end{aligned}$$

Applying the logarithmic function to both sides $(r-1)$ times, we have $p_r \leq \log_{(r-1)} d = p_r + \mathcal{O}(\log p_r)$, implying that $\log_{(r-1)} d - c \log_{(r)} d \leq p_r \leq \log_{(r-1)} d$ for some constant $c > 0$. We see from the above that $p_{r-1} = p_r e^{p_r}$ and that $p_i \geq p_r p_{i+1}$. It follows that $\ell_i \leq \ell_r p_r^{i-r}$ and so

$$\ell_r = d(1 - \eta) \text{ for } \eta \leq \frac{2}{\log_{(r-1)} d} \to 0.$$

Let $T_i := \ell_i \log \frac{\ell_i}{\ell_{i-1}}$. Then we have

$$\frac{T_i}{T_{i-1}} = p_i \frac{\log p_i}{\log p_{i-1}} = \frac{p_i \log p_i}{p_i + \log p_r} \geq \frac{1}{2} \log p_i \gg 1.$$

It follows that for all $i < r - 1$, we have $T_i < T_{r-1}$. Now $\log p_{r-1} = p_r + \log p_r$ implies that

$$T_{r-1} = \ell_{r-1} \log p_{r-1} = \ell_{r-1}(p_r + \log p_r) \leq \frac{\ell_r}{p_r}(p_r + \log p_r) = \mathcal{O}(\ell_r) = \mathcal{O}(d).$$

Thus, the objective is dominated by last summand $T_r = \ell_r \log \frac{\ell_r}{\ell_{r-1}}$, resulting in minimum value of at least $d \log_{(r)} d + \mathcal{O}(d)$. □

Let

$$d_* = \frac{\log n}{\log_{(r+1)} n}.$$

2.1 Upper bound on $\Delta(G^r)$

We prove that $\Delta(G^r) \le d_*(1+\epsilon)$ w.h.p., where $\epsilon = \frac{1}{\log_{(r+1)} n}$. The following inequality will be useful.

Lemma 3. *Suppose that* $a \gg b$. *Then*

$$\log_{(s)}(a-b) \ge \log_{(s)} a - \frac{2sb}{a \log a \log\log a \cdots \log_{s-1} a}.$$

Proof. We prove this by induction on s. For $s=1$ we have

$$\log(a-b) = \log a + \log\left(1 - \frac{b}{a}\right) \ge \log a - \frac{2b}{a}.$$

Then for $s > 1$ we have

$$\begin{aligned}\log_{(s)}(a-b) &= \log\left(\log_{(s-1)}(a-b)\right) \\ &\ge \log\left(\log_{(s-1)} a - \frac{2(s-1)b}{a \log a \log\log a \cdots \log_{s-2} a}\right) \\ &\ge \log_{(s)} a - \frac{2sb}{a \log a \log\log a \cdots \log_{s-1} a}.\end{aligned}$$

□

Corollary 4. *Suppose that* $\log a \gg \log b$. *Then*

$$\log_{(s)}(a/b) \ge \log_{(s)} a - \frac{2(s-1)\log b}{\log a \log\log a \cdots \log_{s-1} a}.$$

□

Using $u_{\ell_1,\ldots,\ell_r}$ as in (2) and Lemma 2,

$$\begin{aligned}\log u_{\ell_1,\ldots,\ell_r} &= d \log c - c(1 + d - \ell_r) - \sum_{i=1}^{r} \ell_i \log \frac{\ell_i}{\ell_{i-1}} + \mathcal{O}(d) \\ &\le -d_*(1+\epsilon)\log_{(r)}(d_*(1+\epsilon)) + \mathcal{O}(d_*) \\ &\le -d_*(1+\epsilon)\log_{(r)} d_* + \mathcal{O}(d_*) \\ &\le -(1+\epsilon)\log n \left(1 - O\left(\frac{\log_{(r)} n}{\log n}\right)\right) \\ &\le -\left(1 + \frac{\epsilon}{2}\right)\log n\end{aligned}$$

Hence $u_{\ell_1,\ldots,\ell_r} \le \frac{1}{n^{1+\epsilon/2}}$. Since there are at most d^r terms in the summation, we have

$$\mathbb{P}\left[\sum_{i=1}^{r} d_i(v) = d\right] = \left(\sum_{\substack{\ell_1,\ldots,\ell_r \\ \ell_1+\ldots+\ell_r = d}} u_{\ell_1,\ldots,\ell_r}\right)\left(1 + \mathcal{O}\left(\frac{d^2}{n}\right)\right) \lesssim \frac{d^r}{n^{1+\epsilon/2}}.$$

Finally by taking the union bound over all n vertices,

$$\mathbb{P}[\Delta_r(G) \ge d] \le \sum_{i=1}^{n} \mathbb{P}\left[\sum_{j=1}^{r} d_j(v_i) = d\right] \lesssim n \cdot \frac{d^r}{n^{1+\epsilon/2}} = \frac{d^r}{n^{\epsilon/2}} \to 0.$$

2.2 Lower bound on $\Delta(G^r)$

We now use the second moment method to show that $\Delta_r(G) \geq d_*(1-\epsilon)$ w.h.p. For $i = 1, 2, ..., n$, let X_i be the indicator random variable for $\sum_{j=1}^{r} d_j(v_i) > d_*(1-\epsilon)$, i.e. the event that $v_i \in V(G)$ has degree greater than $d_*(1-\epsilon)$ in G^r. Let $X_* = \sum_{i=1}^{n} X_i$ and ℓ_* be values of ℓ_i which achieve the lower bound in Lemma 2.

$$\begin{aligned}\mathbb{P}[X_i = 1] \geq u_{\ell_*} &= \exp(-d_*(1-\epsilon)\log_{(r)} d_*(1-\epsilon) + \mathcal{O}(d_*)) \\ &\geq \exp(-(1-\epsilon)\log n + \mathcal{O}(d_*)) \\ &\geq \frac{1}{n^{1-\epsilon/2}}.\end{aligned}$$

Then we have

$$\mathbb{E}[X_*] \geq n^{\epsilon/2}.$$

On the other hand,

$$\begin{aligned}\mathbb{P}[X_* > 0] &\geq \frac{\mathbb{E}[X_*]^2}{\mathbb{E}[X_*^2]} \\ &= \frac{\mathbb{E}[X_*]^2}{\sum_{i=1}^{n} \mathbb{E}\left[X_i^2\right] + \sum_{i \neq j} \mathbb{E}\left[X_i X_j\right]} \\ &\geq \frac{\mathbb{E}[X_*]^2}{\mathbb{E}\left[X_*\right] + \mathbb{E}\left[X_*\right] \sum_{j: d(v_1, v_j) > r} \mathbb{E}\left[X_j\right] + \mathbb{E}\left[\left|\{j : d(v_1, v_j) \leq r\}\right|\right]}.\end{aligned}$$

Here we use the fact $X_j \leq 1$ for all j and that the only variables X_j that are dependent on X_1 are for the vertices v_j within a distance r of v_1.

Now,

$$\mathbb{E}\left[\left|\{j : d(v_1, v_j) \leq r\}\right|\right] \leq O(\log^{r+1} n) + n\mathbb{P}(\Delta(G_{n,p} \geq 10c\log n) \leq 2\log^{r+1} n.$$

So,

$$\mathbb{P}[X_* > 0] \geq \frac{\mathbb{E}[X_*]^2}{\mathbb{E}[X^*] + \mathbb{E}[X_*]^2 + 2\log^{r+1} n} \geq \frac{1}{n^{-\epsilon/2} + 1 + 2n^{-\epsilon}\log^{r+1} n} \to 1.$$

3 Conclusions

We have established the likely value of one of the key parameters related to powers of $G_{n,p}, p = c/n$. It would be interesting to explore other parameters. The chromatic number of $G_{n,p}^2$ was asymptotically determined w.h.p. in Frieze and Raut [2] and the independence number w.h.p. (for large c) in Atkinson and Frieze [1].

References

[1] G. Atkinson and A.M. Frieze, On the b-independence number of sparse random graphs, *Combinatorics, Probability and Computing* 13 (2004) 295-310.

[2] A.M. Frieze and A. Raut, A note on the chromatic number of the square of a sparse random graph.

[3] A.M. Frieze and M. Karoński, Introduction to Random Graphs, Cambridge University Press, 2015.

[4] K. Garapaty, D. Lokshtanov, H. Maji and A. Pothen, The Chromatic Number of Squares Of Random Graphs, *Journal of Combinatorics* 14 (2023) 507-537.