Random Graphs

BY WHZECOMJM

Spring 2020
Bar-Ilan University

Email: whzecomjm@gmail.com

Contents

1 Introduction and Examples

1.1 Lower bound of Ramsey numbers

Ramsey theorem claims that for any 2-colorings, and k a positive integer, there always exists a chromatic sub-complete graph K_s on some complete graph K_n. We denote the least number of n as $R(s)$ or $R(s,s)$.

We already know $R(1)=1$, $R(2)=2$, $R(3)=6$, and $R(4)=18$. However, $R(4)$ is the last one that's known exactly! There's a table of what's known about small Ramsey numbers on Wikipedia. Finding Ramsey numbers exactly is hard. It would still be interesting to have bounds (upper and lower) for these numbers. It turns out that the proof of Ramsey's theorem gives some sort of upper bound, although not the best possible. But here I'd like to talk about a proof of a lower bound. It's not the best known lower bound, because it's quite old and has since been improved, but it's not bad (compared with what's now known) and the proof is significant because it illustrates a technique called the probabilistic method. This proof was first given by Erdős in 1947 in this paper, and I think it's very elegant.

Theorem 1.1. **([4])** *The Ramsey number $R(s)$ is bigger than $2^{(s-1)/2}$.*

Let's remind ourselves what we're doing. We want a lower bound for the Ramsey number $R(s)$. That is, we want some n (ideally as large as possible) so that we can somehow color the edges of K_n using red and blue in such a way that we get no red K_s and no blue K_s (just as above we saw an example of a coloring of K_5 with no red triangle and no blue triangle). We don't actually need to see an example of such a coloring, we just need to know that one exists. (This might seem like a slightly odd idea the first time you come across it!) The method we'll use will prove that such a coloring exists, but will give us no information about how to find it. It is non-constructive.

Here's the idea (with details to follow afterwards). We'll pick a random coloring of the edges of K_n. Then we'll get an upper bound for the probability of a monochromatic K_s. We'll show that if we choose n carefully, we can make that probability less than 1. If the probability is less than 1, then there must be some coloring with no monochromatic K_s (because if every possible coloring of K_n had at least one monochromatic K_s, then the probability of a monochromatic K_s would be one). And that will do nicely.

How do we pick a random coloring? We color each edge of K_n red or blue, using each color with probability $1/2$. We color the edges independently.

What's the probability that we get a monochromatic K_s? It's certainly at most the sum of the probabilities that each K_s is monochromatic (using quite a crude bound). We can find this by counting the total number of K_ss in K_n and multiplying by the probability that an individual K_s is monochromatic.

How many K_ss are there? If we pick any collection of s vertices from our K_n, there's exactly one K_s with those vertices. So the number of K_ss is the number of ways to choose s vertices from n, which is the binomial coefficient $\binom{n}{s}$.

What's the probability that a single K_s is monochromatic? The probability that a single edge is red is $\frac{1}{2}$. The graph K_s has $\binom{s}{2}$ edges, and since the probabilities for different edges are independent we can multiply to see that the probability that $\mathrm{K_s}$ is red is $2^{-\binom{s}{2}}$. Also, the probability that K_s is blue is $2^{-\binom{s}{2}}$, so the probability that K_s is monochromatic is the sum of those, which is $2^{1-\binom{s}{2}}$.

So the probability of a monochromatic K_s is at most $\binom{n}{s}2^{1-\binom{s}{2}}$, and we want to choose n as large as possible so that this is less than 1. We've basically done all the thinking now — the rest is down to careful estimation! I'm not going to be very careful here, so this isn't the best possible bound that one can extract from this argument.

If we take $n = \lfloor 2^{(s-1)/2} \rfloor$ (where the funny brackets mean we round down to the nearest whole number — we take the integer part), what do we get? We use a crude bound for the binomial coefficient; I'll let you convince yourself that it's true. The probability of a monochromatic $\mathrm{K_s}$ is at most

$$2^{1-\binom{s}{2}}\binom{n}{s} \leq 2^{1-\binom{s}{2}}\frac{n^s}{s!} < 2^{-s(s-1)/2}2^{s(s-1)/2} = 1,$$

so the probability is strictly less than 1, as we wanted. If you look at Erdős's paper, you'll see that he phrases this argument using counting rather probability. The reason that it tends to be phrased probabilistically nowadays is that this language allows one to use lots of other ideas from probability (such as expectation and variance and so on).

1.2 Clique and coclique

Ramsey Theorem states that one will find monochromatic cliques in any edge coloring of a sufficiently large complete graph. To demonstrate the theorem for two colours, there exists a least positive integer $R(k,l)$ for which every blue-red edge coloring of the complete graph on $R(k,l)$ vertices contains a blue clique on k vertices or a red clique on l vertices.

Ramsey Numbers $R(k,l)$.

We will use

- $\alpha(G)$ - maximal size of independent set (stable set, **coclique**, 任意两点均不邻接的集合);
- $\omega(G)$ - size of the maximal clique (团, 任意两点邻接的顶点的集合).

更多内容参见 [1 in which 1.5&1.6.2 pp. 7, 9–10] 或者 Theorem of the week [5] 以及 Wikipedia 的 Clique problem 词条.

1.3 Hadwiger Conjecture

如果无向图H可以通过图G删除边和顶点或收缩边得到，称H为G的子式或次图。

Definition 1.2. **(Graph Minor)** *A graph $H=([k],F)$ is a minor(子式) of G if $V(G)$ contains k nonempty disjoint subsets $V_1,\dots,V_k$ such that:*

1. *the induced subgraph $G[V_i]$ is connected,*
2. *for every $f=\{i,j\}\in F$, there is an edge between V_i and V_j in G.*

我们有如下等价定义:

Definition 1.3. **(Minor)** *An undirected graph* H *is called a* ***minor*** *of the graph* G *if* H *can be obtained from* G *by a sequence of the following operations:*

1. *deleting a vertex,*
2. *deleting an edge,*
3. *contracting an edge. (边收缩)*

Definition 1.4. *The contraction clique number (Hadwiger number) of G, denoted ccl(G), is the maximum k for which K_k is minor of G.*

Conjecture 1.5. **(Hadwiger Conjecture, 1943)** *For each graph G, the chromatic number $\chi(G)\le \mathrm{ccl}(G)$.*

Notes.

- k=2 is true.
- k=3 is true.
- k=4 was proven by Hadwiger.
- k=5 equivalent to the four-color theorem. (Wagner, 1937).
- k=6 (Robertson, Seymour, Thomas, [7])
- $k\ge 7$ still open.

What if we give up on answering the question for all graphs, and weaken it to a probabilistic statement: does Hadwiger's conjecture hold for almost every graph?

Theorem 1.6. (Bollobas, Catlin, Erdos 1980, [6] or 1.6.3 in [1]) *Hadwiger conjecture is almost all true for all graphs.*

More precisely, Hadwiger conjecture holds a.a.s. for $G \sim G(n, 1/2)$.

Proof.

1. $\mathrm{ccl}(G) \geq \frac{n}{6\sqrt{\log_2 n}}$
2. $\chi(G) \leq \frac{6n}{\log_2 n}$

$G = G_1 \cup G_2$ in every G_i every edge is included with probability p, $1/2 = (1-p)^2$, $p \geq 1/4$.

We look for a path by choosing v1 (arb.) and as long as $i \leq n/2$, find a neighbour of v_i outside of $\{v_1, \ldots v_{i-1}\}$.

Details omitted. (cf. pp. 12-13, [7])

2 Basic Models of Random Graphs

2.1 Basic models

We denote $[n] = \{1, 2, \cdots, n\}$.

Erdos-Renyi Model is model which given fixed n vertices, and independent probabilities p for all edges. There are two ER models, namely, G(n,p) and G(n,m).

Definition 2.1. (binomial random graph) *(G(n,p)) probabilistic space over graph, $V = [n]$, for each pair $1 \leq i < j \leq n$,*

$$Pr\left[\{i, j\} \in E\right] = p = p(n)$$

independently of all other edges.

Equivalent to $Pr[G] = p^{e(G)} (1-p)^{\binom{n}{2} - e(G)}$.

Example 2.2. Kidney exchange graph

2个点看作是一个 vertex, donor and recipient. (受捐者和捐献者视为一个整体.)

Remark 2.3. If p=0.5, all graphs are equiprobable.

Definition 2.4. (uniform random graph, G(n,m)) *The sample space: $\Omega = \{G = (V, E) | V = [n], |E| = m\}$, p is the uniform distribution.*

$Pr[G] = \frac{1}{\Omega} = \frac{1}{\binom{\binom{n}{2}}{m}}$.

$d_{G(n,p)}(v) \sim B i n(n-1, p)$ *degree of vertices.*

Intuitively, we expect $m = E G(n, p) = \binom{n}{2} \cdot p$, then somehow $G(n, p)$ and $G(m, n)$ will be alike.

内积模型是对每一个顶点指定一个实系数的向量，两个顶点之间连接的概率是向量内积的函数。

Definition 2.5. (Random Graph process) *Denote $\binom{n}{2} =: N$. Pick a permutation $\sigma = (\sigma_1, \cdots \sigma_N)$ over the edges of K_n. Defined $G_t = ([n], \{e_1, \ldots, e_t\})$, $0 \leqslant t \leqslant N$ and $G_0 = ([n], \emptyset)$. Then $\tilde{G} = \tilde{G}(\sigma) = (G_i)_{t=0}^{N}$ is a random graph process.*

Exercise 2.1. Let $\tilde{G}=(G_t)$ be a r.g.p.. Then $G_t \sim G(n, m=t)$.

Proof cf. Prop. 12 in [1, 1.2&1.7].

2.2 Staged exposure

Proposition 2.6. (Staged exposure in G(n,p)) *Suppose* $0 \leq p \leq 1$, *and* $0 \leq p_1, \ldots, p_k \leq 1$, *such that* $1-p=\prod_{i=1}^k (1-p_i)$. *Then,* $G(n,p)=\bigcup_{i=1}^k G(n,p_i)$.

Proof. $G \sim G(n,p)$. Every edge appears independently in both models.

$Pr_{\cup G(n,p_i)}[e \notin G]=\prod_{i=1}^k (1-p_i)=1-p$.

2.3 Monotonicity

Definition 2.7. *A* graph property *is a class of graphs closed under isomorphism.*

Definition 2.8. *A graph property P is called* ***monotone*** *(increasing) if* $V[G]=V[H]$ *and*

$$G \in P, G \subset H \Longrightarrow H \in P.$$

Remark 2.9. Connectivity and Hamiltonicity are monotone increasing.

Proposition 2.10. ([1], 1.15) *Let P be a monotone increasing property,* $0 \leq p_1 \leq p_2 \leq 1$, $0 \leq m_1 \leq m_2 \leq \binom{n}{2}$. *then*

(1) $\Pr[G(n,p_1) \in P] \leq \Pr[G(n,p_2) \in P]$

(2) $\Pr[G(n,m_1) \in P] \leq \Pr[G(n,m_2) \in P]$

Proof. Easy to start with the second case.

(II) Consider R.G.P $\tilde{G}=(G_t)$ and $G_{m_1} \sim G(n,m_1), G_{m_2} \sim G(n,m_2)$, then event "$G(n, m_1)$ has P" is actually a subset of the event "$G(n,m_2)$ has P". (Using exercise) $\Longrightarrow \Pr[G(n,m_1) \in P] \leq \Pr[G(n,m_2) \in P]$.

(A coupling argument)

(I) Take $G_1 \sim G(n,p_1)$, $G_2 \sim G(n,p_2)$ and $G_0 \sim G(n,p_0)$ where $(1-p_1)(1-p_0)=1-p_2$ $(0 \leq p_0 \leq 1)$.

Then we have $G(n,p_2)=G(n,p_1) \cup G(n,p_0)$.

Then $\Pr[G_2 \,\text{has}\, P]=\Pr[G_1 \,\text{has}\, P]+\Pr[G_0 \cup G_1 \,\text{has}\, P | G_1 \notin p]$.

2.4 Asymptotic relation between $G(n,p)$ and $G(n,m)$

$\Rightarrow$ we expect G(n,p) to be closed to G(n,m) when $p\binom{n}{2}=m$.

$\Rightarrow$ set $m=p\binom{n}{2}$, intuitively we have

$$\Pr[G(n,m)\text{has m edges}]=1$$

$$\Pr[G(n,p)\text{has m edges}]=\frac{1}{\Omega\Big(\sqrt{\binom{n}{2}p(1-p)}\Big)}$$

since the number of edges in $G(n,p)\sim \mathrm{Bin}(\binom{n}{2},p)$.

Lemma 2.11. (Prop 1.29, [1]) *Let A be a graph property, let $p{=}p(n)$ be a sequence of real number with $0\le p(n)\le 1$ Finally, $0\le a\le 1$. If for all $m=\binom{n}{2}p+O(\sqrt{\binom{n}{2}p(1-p)})$,*

$\lim_{n\to\infty} Pr[G(n,m)\in A]=a$. *Then also*

$$\lim Pr[G(n,p)\in A]=a.$$

Example. A connectivity, $p(n)=\frac{\log n}{n}$.

Proof. By the law of total probability. choose large enough $c>0$.

Let $M=\{0\le m\le \binom{n}{2}||m-\binom{n}{2}p|\le c\sqrt{\binom{n}{2}p(1-p)}\}$,

Then by Chebyshev's inequality, we have

$$\Pr[\#\text{ edges in }G(n,p)\notin M]\le \frac{1}{c^2}$$

Let $m_*=\mathrm{argmin}_m \Pr[G(n,m)\in A]$.

and $m^*=\mathrm{argmax}_m \Pr[G(n,m)\in A]$ (arg max 取得极大值的集合)

$$\begin{aligned}
\Pr[G(n,p)\in A] &= \sum_{m=0}^{\binom{n}{2}} \Pr[G(n,m)\text{has A}]\Pr[|G(n,p)|=m]\\
&\geqslant \sum_{m\in M} \Pr[G(n,m)\in A]\,Pr[G(n,p)\text{has m edges}]\\
&\geqslant \sum_{m\in M} \Pr[G(n,m_*)\in A]\Pr[G(n,p)\text{has m edges}]\\
&= \Pr[G(n,m_*)\in A]\sum_{m\in M} \Pr[G(n,p)\text{has m edges}]\\
&\geqslant (1-\frac{1}{c^2})\Pr[G(n,m_*)\in A]\\
&\Rightarrow \ \mathrm{liminf}\Pr[G(n,p)\in A]\ge a(1-\frac{1}{c^2}).
\end{aligned}$$

We've seen $\liminf \Pr\left[G\left(n,p\right)\in A\right]\geq a\left(1-\frac{1}{c^2}\right)$.

We need to give a matching upper bound.

$$\begin{aligned}
\Pr\left[G\left(n,p\right)\in A\right] &= \sum_{m=0}^{\binom{n}{2}} \Pr\left[G\left(n,p\right)\text{has m edges}\right]\Pr\left[G\left(n,m\right)\text{ has }A\right]\\
&\leqslant c^{-2}+\sum_{m\in M} \Pr\left[G\left(n,p\right)\text{has m edges}\right]\Pr\left[G\left(n,m\right)\text{ has }A\right]\\
&\leqslant c^{-2}+\Pr[G(n,m^*)\text{ has }A].
\end{aligned}$$

Hence,

$$\limsup \Pr\left[G\left(n,p\right)\in A\right]\leq a+c^{-2}.$$

Together, we get that $\lim \Pr\left[G\left(n,p\right)\in A\right]=a$. □

Question 1. Can we conclude from $G\left(n,p\right)$ to $G\left(n,m\right)$?

For the direction $G\left(n,p\right)\rightarrow G\left(n,m\right)$, we have to be a bit more careful. We need monotonicity; consider the counterexample of A being the property of having exactly m edges.

Lemma 2.12. (cf. [1] Prop 1.30.) *Let A be a* ***monotone*** *graph property. Assume $0\leq m=m(n)\leq N=\binom{n}{2}$ (where N potential edges). Assume that for every sequence p(n) such that*

$$p=\frac{m}{N}+O\left(\sqrt{\frac{m\left(N-m\right)}{N^3}}\right)$$

We have if

$$\lim \Pr\left[G\left(n,p\in A\right)\right]=a,$$

then $\lim \Pr\left[G\left(n,m\right)\in A\right]=a$.

Proof omitted.

3 Evolution of Random Graphs

Remark. Section 3.1-3.3 is based on [2] chapter 2.

3.1 Subcritical phase

Around $p=1/n$, $G(n,p)$ changes a lot.

Proposition 3.1. (thm 2.1 [2]) *If* $p=o(1/n)$ *(resp.* $m=o(1)$*), then (asymptotic almost surely) a.a.s.* $G(n,p)$ *is a forest.*

Proof. Let X be the number of cycles in $G(n,p)$.

$$\begin{aligned} E[X] &= \sum_{k=3}^{n} \binom{n}{k}\frac{(k-1)!}{2}\,p^k \\ &\leqslant \sum_{k=3}^{n} \frac{n^k}{k!}\,\frac{(k-1)!}{2}\,\frac{1}{n^k\,\omega^k} \\ &= \sum_{k=3}^{n} \frac{1}{2\,k}\,\frac{1}{\omega^k} = o\!\left(\frac{1}{\omega^3}\right) = o(1) \end{aligned}$$

$p\,n\to 0 \Longleftrightarrow p=o(1/n)$ iff $p=\frac{1}{n\,\omega(n)}$ with $\omega(n)\to\infty$.

Note that $\sum \frac{1}{2\,k}\,\frac{1}{\omega^k} \leq \sum \frac{1}{6}\,\frac{1}{\omega^3}\,(1+o(1))$.

By Markov's inequality, $\Pr[X\geq 1] \leq \frac{E[X]}{1} = o(1)$ □

注记. 树是没有环的连接图, 森林是多个不连接的树.

Proposition 3.2. (thm 2.2 [2]) *If* $p=o(n^{-1.5})$ *(resp.* $m=o(n^{-0.5})$*) then a.a.s.* $G(n,p)$ *is a collection of isolated vertices and edges.*

(也就是仅仅有孤立的点和孤立的线段)

Proof. Let X be the number of path of length 2 in $G(n,p)$.

$E[X]=\binom{n}{3}3\,p^2=\frac{n\,(n-1)\,(n-2)}{2}\,p^2=o(n^3\cdot n^{2\cdot -1.5})=o(1)$.

By Markov's inequality a.s.s. $G(n,p)$ has no paths of length 2. □

The first moment method:

X_n is a sequence of r.v., suppose $(X_n)\in\mathbb{B}\cup\{0\}$ then If $E[X_n]=o(1)$, $\Pr[X_n=0]\to 1$.

Lemma 3.3. (First moment method) *Let X be a non-negative integer valued random variable. Then*

$$\Pr[X>0] \leqslant E[X].$$

Proposition 3.4. (thm 2.3 [2]) *If $p=\omega(n^{-1.5})$ (little ω notation), then a.a.s. $G(n, p)$ contains a path of length 2.*

Lemma 3.5. (thm 2.5, [2]) *Fix $k\geq 3$, if $p=o(n^{-\frac{k}{k-1}})$, then a.a.s. $G(n,p)$ contains no tree with k vertices.*

Proof. Let $p=\frac{1}{n^{k/(k-1)}\omega}$ where $\omega(n)\to\infty$ arbitrary slowly. Let X be the numbers of trees with k vertices.

$E[X]=\binom{n}{k}k^{k-2}p^{k-1}\leq n^k\frac{k^{k-2}}{k!}p^{k-1}=c_k n^k p^{k-1}=c_k\cdot\omega^{-(k+1)}=o(1)$.

Note that k^{k-2} is the number of labelled (带标记的树) trees on k vertices, Cayley's formula. (cf. Proofs of the book.) □

Theorem 3.6. (lemma 2.10, [2]) *If $p\leq\frac{1}{n}-\frac{\omega}{n^{4/3}}$ where $\omega\to\infty$. Then a.a.s. every (connected) component in $G(n,p)$ contains at most one cycle.*

Proof. w.t.s. there are no connected components with two cycles. If there is such a conn. components, then there are three possible forms:

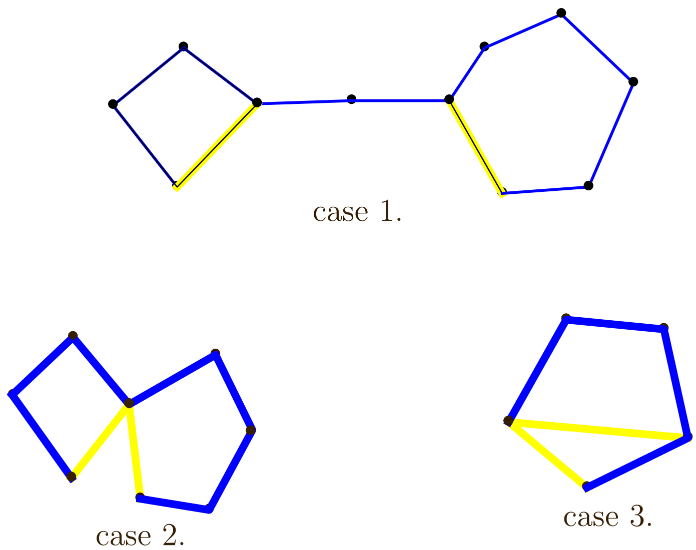

Figure 3.1. Three possible forms

And we can think of it has a path with two additional edges from the endpoints.

The number of connected component with two edges on k vertices is $\leq k!\cdot k^2$.

If X counts such components:

$$\begin{aligned}
E[X] &\leqslant \sum_{k=4}^{n} \binom{n}{k} k^2\, k!\, p^{k+1} \\
&\leqslant \sum \frac{n^k}{k!} k^2\, k!\left(\frac{1}{n} - \frac{\omega}{n^{4/3}}\right)^{k+1} \\
&\leqslant \sum k^2 (1 - n^{1/3})^{k+1} \cdot \frac{1}{n} \\
&\leqslant \int_0^\infty \frac{x^2}{n} \exp\left(-\frac{\omega\, x}{n^{1/3}}\right) d\, x = \frac{2}{\omega^3} = o(1).
\end{aligned}$$

Note that $1 - x \le e^{-x}$. □

We have already seen following results.

Subcritical Phase: $G(n, p = \frac{c}{n})$, $c < 1$.

$p = o(n^{-3/2})$: no V, at most $-$ shapes;

$p = o(n^{-\frac{k}{k-1}})$: no trees with k vertices.

$p^* = n^{-\frac{k}{k-1}}$ is a threshold function for the appearance of trees with k vertices. (Critical Point)

Remark 3.7. p^* is a *threshold*(阈值) for monotone property A if

$p \gg p^* \Longrightarrow G(n, p) \in A$, a.a.s.

$p \ll p^* \Longrightarrow G(n, p) \notin A$, a.a.s.

Lemma 3.8. (lemma 2.10, [2]) *If $p \le \frac{1}{n} - \frac{\omega}{n^{4/3}}$, $\omega \to \infty$ a.a.s. there are no components with two cycles.*

Lemma 3.9. (lemma 2.11, [2]) *If $p = \frac{c}{n}$, with $c \neq 1$. Then a.a.s. the number of vertices in unicyclic (单圈, 只有一个圈) components is $O(\omega)$ for any $\omega \to \infty$.*

(for $\omega \to \infty$ as slowly as you like)

Remark 3.10. the result is equivalent to there are some constant M(c), such that # bounded by M(c).

Proof. X_k is the number of vertices in unicyclic components with k vertices.

$E\, X_k \le \binom{n}{k} k^{k-2} \binom{k}{2} k\, p^k\, (1-p)^{k(n-k) + \binom{k}{2} - k}$.

Recall $\binom{n}{k} \leq \frac{n^k}{k!} e^{-\frac{k(k-1)}{2n}}$ and $k! \leq (\frac{k}{e})^k$ $(1-p)^{kn-k^2+\binom{k}{2}-k} = (1-p)^{kn-\binom{k}{2}-2k}$.

Then

$$\begin{aligned}
\mathrm{E}[X_k] &\leqslant \frac{n^k}{k!} e^{-\frac{k(k-1)}{2n}} k^{k+1} (c/n)^k e^{-ck+\frac{ck(k-1)}{2n}+\frac{ck}{2n}} \\
&\leqslant \frac{e^k}{k^k} e^{-\frac{k(k-1)}{2n}} k^{k+1} c^k e^{-ck+\frac{k(k-1)}{2n}+c/2} \\
&= \frac{e^k}{k^k} e^{f(k)/n} c^k e^{-ck} e^{c/2} \\
&\leqslant k (c e^{1-c})^k e^{c/2(1+o(1))}
\end{aligned}$$

$$\mathrm{E}\left[\sum_{k=3}^{n} X_k\right] \leq \sum_{k=3}^{n} k (c e^{1-c})^k e^{c/2} (1+o(1)) = O(1).$$

By Markov's inequality, when $\omega \to \infty$,

$$\Pr\left[\sum_{k=3}^{n} X_k \geq \omega\right] = O\left(\frac{1}{\omega}\right) \to 0 \qquad \square$$

Lemma 3.11. (lemma 2.12, [2]) *Let $p = c/n$ with $c \neq 1$ constant, $\alpha = c - 1 - \log c$, $\omega(n) \to \infty$ and $\omega = o(\log \log n)$.*

1. *a.a.s. there is a isolated tree of size $\frac{1}{\alpha}(\log n - \frac{5}{2} \log \log n) - \omega$.*
2. *a.a.s. there is no isolated tree of size at least $\frac{1}{\alpha}(\log n - \frac{5}{2} \log \log n) + \omega$*

注: 如果上式没有ω 则可能存在也可能不存在.

Then we have following theorem by last 3 lemmas,

Theorem 3.12. (lemma 2.9, [2]) *If $m = \frac{1}{2} c n$, where $0 < c < 1$ is a constant, then a.a.s. the order of the largest component of a random graph $G(n, m)$ is $O(\log(n))$.*

3.2 Supercritical phase

Theorem 3.13. (cf. thm 2.14, [2]) *If $p = c/n, c > 1$, then a.s.s. (w.h.p.) $G(n, p)$ consists of a unique giant component, with $\left(1 - \frac{x}{c} + o(1)\right) n$ vertices and $\left(1 - \frac{x^2}{c^2} + o(1)\right) \frac{cn}{2}$ edges. Here $0 < x < 1$ is the solution of the equation $x e^{-x} = c e^{-c}$. The remaining components are of order at most $O(\log n)$.*

Proof. cf. pp. 34-37, thm 2.14 [2]. $\square$

Structure of the supercritical giant component is given by the following theorem.

Theorem 3.14. (Ding, Lubetzky, Peres, thm1, [8]) *c_1 - the largest component in $G(n, p = \lambda/n)$, $\lambda > 1$. Let $\mu < 1$ is s.t. $\mu e^{-\mu} = c e^{-c}$.*

c_1 is contiguous(类似于同构) to the following:

1. *$\Lambda \sim N\left(\lambda - \mu, \frac{1}{n}\right)$, $D_u \sim \mathrm{Poisson}(\Lambda)$ for $u \in [n]$ be i.i.d. conditioned on $\sum_{D_u | D_u \geq 3} D_u$ being even. $N_k = |\{u | D_u = k\}|$, $N = \sum_{k \geq 3} N_k$. Select a random multigraph $\mathcal{K}$ on N vertices, uniformly over all multigraphs with N_k vertices of degree k ($k \geq 3$).*

2. *Replace each edge of $\mathcal{K}$ by a path of $\mathrm{Geo}(1 - \mu)$ length, i.i.d.*

3. *Attach to each vertex an independent $\mathrm{Poisson}(\mu)$ Galton Watson tree.*

That is: $P(\widetilde{C}_1 \in A) \to 0, n \to \infty$ implies $\Pr(C_1 \in A) \to 0, n \to \infty$ for any graph property A.

3.3 Thresholds and second moment method

This section is based on [2] Sec 5.1.

3.3.1 Threshold

Definition 3.15. (Threshold) *A function $p^*(n)$ is a threshold for a monotone graph property A is*

1. *$p = o(p^*) \Rightarrow \Pr[G(n,p) \in A] = o(1)$*

2. *$p = \omega(p^*) \Rightarrow \Pr[G(n,p) \in A] = 1 - o(1)$*

Example 3.16. p=1/n is a threshold function for having a component with at least 2 cycles.

- $p \ll 1/n$ all components are unicyclic a.a.s..

- $p \gg 1/n$ ass there is a component with more than one cycle.

Definition 3.17. *Let A be a monotone increasing graph property. For $0 < a < 1$. $p_a := \min\{0 \leq p \leq 1 | \Pr[G(n,p) \in A] \geq a\}$.*

Definition 3.18. (sharp threshold) *A has a sharp threshold if for every constant $\epsilon > 0$, consider the gap $\frac{p_{1-\epsilon} - p_\epsilon}{p_{1/2}} = o(1)$. Otherwise, it's a coarse threshold.*

Example 3.19. If $p = \frac{\ln n + \omega(n)}{n}$ then a.a.s. $G(n,p)$ is connected. if $p = \frac{\ln n - \omega(n)}{n}$ then a.a.s. there are isolated vertices.

Note that $p_\epsilon \geq \frac{\ln b - \ln\ln n}{n}$, then

$$\frac{p_{1-\epsilon} - p_\epsilon}{p_{1/2}} = \frac{2\ln\ln n}{n} \Big/ \frac{\ln n}{n} = o(1).$$

Then we get a sharp threshold.

3.3.2 The second moment method

(cf. Section 21.1 in [2] or cf. [9])

Proposition 3.20. *X is non-negative and integer value. If* $\mathrm{Var}\,[x] = o(E^2[X])$, *then* $\Pr\,[X \geq 1] = 1 - o(1)$.

Proof. $\Pr\,[X = 0] \leq \Pr\,[|X - E[X] \geq E[X]] \leq \frac{\sigma^2}{E^2[X]} = o(1)$. ☐

If $X = \sum_{i=1}^k X_i$ is a sum of indicator RV.

$$\mathrm{Var}[X] = \sum_{i=1}^k \mathrm{Var}[X_i] + \sum_{1 \leqslant i \neq j \leqslant k} \mathrm{Cov}(X_i, X_j)$$

Note that $\mathrm{Var}[X_i] = p(1-p) \leqslant p = E[X_i]$, then we have

$$\mathrm{Var}[X] \leqslant E[X] + \sum_{1 \leqslant i \neq j \leqslant k} \mathrm{Cov}(X_i, X_j)$$

Write $i \sim j$ (**adjacent**) if X_i and X_j are not independent, and $i \nsim j$ if X_i, X_j are independent.

Denote

$$\boxed{\Delta = \sum_{i \sim j} \Pr\,[X_i = 1 \text{ and } X_j = 1]}$$

Recall that

$$\mathrm{Cov}(X_i, X_j) = E\,[X_i\, X_j] - E[X_i]\, E[X_j] \leq E\,[X_j\, X_j] = \Pr\,[X_i = 1 \text{ and } X_j = 1].$$

$$\begin{aligned} \mathrm{Var}[X] &\leqslant \mathrm{E}[X] + \sum_{i \sim j} \mathrm{Cov}(X_i, X_j) \\ &\leqslant \mathrm{E}[X] + \sum_{i \sim j} \mathrm{E}[X_i X_j] \\ &= \mathrm{E}[X] + \Delta. \end{aligned}$$

Example 3.21. (threshold for K_4) X counts copies of K_4 in $G(n,p)$.

$$E\,X = \binom{n}{4} p^6 = \Theta\,(n^4\,p^6).$$

intuition: $E\,X = 1$, then we need $n^4\,p^6 = 1$ i.e. $p = n^{-2/3}$.

Claim: $p = n^{-2/3}$ is a threshold function for the appearance of K_4.

Proof. $X = \sum_{s \leq v, |s| = 4} X_s$ where X_s is an indicator for $G[s] = K_4$.

$$E[X] = E[\sum X_s] = \sum E[X_s] = \binom{n}{4} p^6 \sim \frac{n^4}{24} p^6$$

if $p \ll n^{-2/3}$, $E[X] = o(1)$ by Markov's inequality a.a.s. $X = 0$.

The positive part: $p(n) = \omega(n^{-2/3})$. $E[X] = \Theta\,(n^n\,p^6) = \omega(1)$.

$$\mathrm{Var}[X] \leq E[X] + \Delta$$

X_{s_1} and X_{s_2} are dependent iff $2 \leqslant |s_1 \cap s_2| \leqslant 3$, the following picture shows those two cases of dependency.

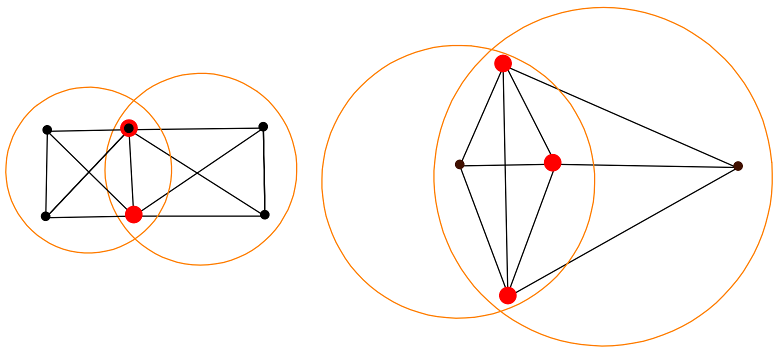

Figure 3.2. two acceptable cases

s1 s2至少两个共同顶点:

$$\Delta = \sum_{\substack{2 \leq |s_1 \cap s_2| \leq 3 \\ |s_1| = |s_2| = 4, s_1, s_2 \leq v}} \Pr\,[X_{s_1} = 1 \text{ and } X_{s_2} = 1]$$

Then number of pairs s_1, s_2 s.t. $|s_1 \cap s_2| = 2$, $\binom{n}{4}\binom{4}{2}\binom{n-4}{2} = \Theta(n^6)$. the probability $\Pr\,[X_{s_1} = X_{s_2} = 1] = p^{11}$ for all such pairs.

If $|s_1 \cap s_2| = 3$, then the number of pairs is $\binom{n}{4}\binom{4}{3}\binom{n-4}{1} = \Theta(n^5)$,

$$\Pr\,[X_{s_1} = 1 \wedge X_{s_2} = 1] = p^9$$

Hence, $\mathrm{Var}[X] \le E[X] + \Delta = \Theta\left(n^4 p^6\right) + O(n^6 p^{11} + n^5 p^9) = o(E^2[X])$

by the second moment method,

$$\Pr[X \ge 1] = 1 - o(1).$$

Example 3.22. (Kite graph) 风筝 5个顶点和7条边. $E[X] = \binom{n}{5} p^7$, $p^* = n^{-5/7}$ 吗?

take $p = n^{-5/7} \log n$ a.a.s. no copy of kite in $G(n,p)$, since ass there is no copy of K_4 !

$E[X] = \binom{n}{4} p^5$ hence if $p = n^{-5/7}$, then $E(X) = o(1)$ ass there is no copy of K4.

the expected degree is $(n-1)\,p \approx n^{2/7}$.

<u>这个错误是风筝有很多线</u>!

How to determine a certain subgraph appear ass? We will first define the inherited density.

3.3.3 Density of graphs and threshold

The density of a graph H, is defined by $d(H) = \frac{e(H)}{v(H)}$.

Definition 3.23. *The inherited density of a graph H is*

$$m(H) = \max\{d(H_0) | H_0 \subseteq H, v(H_0) > 0\}.$$

If H itself is the densest subgraph of it, we call H a balanced graph. If H is denser than all proper subgraph then it is strictly balanced graph.

Example 3.24. 在上述风筝图H中, $m(H) = \frac{e(K_4)}{v(K_4)} = \frac{6}{4} = 3/2$.

Theorem 3.25. (Bollobas, '81, cf. [2] thm 5.3) *Let H be a fixed graph, $e(H) > 0$. The threshold function for the appearance of H as a subgraph of $G(n,p)$ is $p(n) = n^{-1/m(H)}$, i.e.*

$$\lim_{n \to \infty} \Pr[H \subseteq G(n,p)] = \begin{cases} 0 & \text{if} \quad p = o(n^{-1/m(H)}) \\ 1 & \text{if} \quad p = \omega(n^{-1/m(H)}) \end{cases}$$

Proof. Cf. pp. 74-75 and Proof of thm 5.2 in [2]. ◻

Remark 3.26. The probability of not having a copy of H in $G(n, p \gg n^{-1/m(H)})$ is bounded by

$$\exp\left(-\frac{1}{1-p}\,\Phi(H)\right) \leq e^{-c\Phi(H)}.$$

Question 2. (Shamir & Schmult conjecture) Every vertex is in a copy of H? (每个顶点是否都在 H 这样的子图内?)

- JKV (2008)
- Fractional versions etc. (Simi)

3.4 Exercises

Exercise 3.1. (cf. [2] Lemma 21.4, 21.5) Let X be an integral non-negative random variable. Show:

a)

$$\Pr[X=0] \leqslant \frac{E[X^2]}{(EX)^2} - 1;$$

b)

$$\Pr[X=0] \leqslant 1 - \frac{(EX)^2}{E[X^2]};$$

c) Which bound is stronger?

Proof. (a) Note that $\{X=0\} \subseteq \{|X - EX| \geq EX\}$, then by Chebyshev's inequality,

$$\Pr[X=0] \leq \Pr[|X - EX| \geq EX] \leq \frac{\operatorname{Var}(X)}{(EX)^2} = \frac{E[X^2]}{(EX)^2} - 1.$$

(b) Note that $X = X I_{X>0}$, where $I_{X>0}$ is the indicator of $X>0$. This is true since if $X=0$ then both sides are 0, while if $X>0$ then both sides are X. By Cauchy-Schwarz inequality,

$$EX = E(X I_{X>0}) \leq \sqrt{E[X^2]\,E(I_{X>0}^2)} = \sqrt{E[X^2]\,E(I_{X>0})}.$$

Rearranging this and using the fundamental bridge, we have

$$\Pr[X>0] = E(I_{X>0}) \geq \frac{(EX)^2}{E[X^2]}.$$

Hence,

$$\Pr[X=0] \leq 1 - \frac{(EX)^2}{E[X^2]}.$$

(c) The inequality in (b) is stronger than the one in (a). Recall that the RHS of the inequality in (a) is $\frac{\operatorname{Var}(X)}{(EX)^2}$, while the the one in (b) is $\frac{\operatorname{Var}(X)}{E[X^2]}$. On the other hand, $\operatorname{Var}(X) \geq 0$ because of the non-negativity of X, and then

$$\operatorname{Var}(X) = E[X^2] - (EX)^2 \geq 0,$$

i.e. $E[X^2] \geq (E X)^2$. Hence we have

$$\frac{\mathrm{Var}(X)}{E[X^2]} \leq \frac{\mathrm{Var}(X)}{(E X)^2},$$

which means (b) is stronger than (a). ☐

Exercise 3.2. (cf [2] Thm 1.12) Let $\tilde{G} = (G_n)_{n=1}^{\infty}$ be a random graph process. Show that a.a.s. triangles appear in $\tilde{G}$ before it becomes connected.

Proof. Note that having a triangle and being connected are both monotone increasing graph properties, and we can conclude from $G(n, p)$ to $G(n, m)$. To prove that a.a.s. triangles appear before it becomes connected, it is enough to show that the thresholds of those properties satisfy $p_T^* \leq p_C^*$ (p_C^*, p_T^* are thresholds of having a triangle and connectivity respectively). We would like to use Erdős–Rényi's famous result for the threshold of connectivity that $p_C^*(n) = \frac{\ln n}{n}$ is a sharp threshold for the connectedness of $G(n, p)$.

For the property of having a triangle, we claim its threshold is $p_T^*(n) = \frac{1}{n}$. Recall that if $p = o(1/n)$, then a.a.s. $G(n, p)$ is a forest, hence there is no triangles (by Markov's inequality). The only thing we need to show is that if $n p \to \infty$, then a.a.s. $G(n, p)$ contains at least one triangle.

Assume that $n p = \omega(n)$, where $\omega(n) \to \infty$ as slowly as possible. Let X be the number of triangles in $G(n, p)$. Then

$$E[X] = \binom{n}{3} p^3 \geq \frac{\omega^3}{6} (1 - o(1)) \to \infty.$$

For simplicity, we denote $N = \binom{n}{3}$. Let $T_1, \ldots, T_N$ denote all triangles. Then we can calculate the following

$$\begin{aligned}
E[X^2] =& \sum_{i,j=1}^{N} \Pr[T_i, T_j \in G(n, p)] \\
=& \sum_{i=1}^{N} \Pr[T_i \in G(n, p)] \sum_{j=1}^{N} \Pr[T_j \in G(n, p) | T_i \in G(n, p)] \\
=& N \Pr[T_1 \in G(n, p)] \sum_{j=1}^{N} \Pr[T_j \in G(n, p) | T_1 \in G(n, p)] \\
=& E[X] \cdot \sum_{j=1}^{N} \Pr[T_j \in G(n, p) | T_1 \in G(n, p)] \\
=& E[X] \cdot \Bigg(1 + \sum_{T_1, T_j \text{share 1 edges}} \Pr[T_j \in G(n, p) | T_1 \in G(n, p)] + \\
& \sum_{T_1, T_j \text{share 0 edges}} \Pr[T_j \in G(n, p) | T_1 \in G(n, p)] \Bigg) \\
=& E[X] \cdot \left\{ 1 + 3(n-3) p^2 + \left[3 \cdot \binom{n-3}{2} + \binom{n-3}{3} \right] p^3 \right\} \\
\leq& E[X] \cdot \left(1 + \frac{3 \omega^2}{n} + E[X] \right).
\end{aligned}$$

From above **exercise 3.1(a)**, we have

$$\Pr[X=0] \leq \frac{E(X^2)}{E^2[X]} - 1 \leq \frac{E[X]\,(1+\frac{3\,\omega^2}{n})}{E^2[X]} = \frac{1+\frac{3\,\omega^2}{n}}{\frac{\omega^3}{6}\,(1-o(1))} = o(1).$$

Hence, a.a.s. $G(n,p)$ contains at least one triangle. To sum up, we conclude that a.a.s. triangles appear in $\tilde{G}$ before it becomes connected since the thresholds satisfy $p_T^* \leq p_C^*$. □

Exercise 3.3. (cf. [10] Solution 1.2) Let $\varepsilon > 0$ be constant. Show that a.a.s. $G(n, m = (1+\varepsilon)n)$ is not planar.

Proof. Let denote n, e, f the number of vertices, edges and faces of a planar graph, by Euler's formula, we have $n - e + f \geq 2$. If a graph G has girth at least k then $f \leq \frac{2\,e}{k}$ and therefore $e \leq \frac{k}{k-2}\,(n-2) \leq \frac{k}{k-2} \cdot n$. Choose $k := \frac{4+2\,\epsilon}{\epsilon}$. Then we have

$$|E(G)| = e \leq \left(1 + \frac{\epsilon}{2}\right) n \qquad (1)$$

Let $G(n,m)$ with $m \geq (1+\epsilon)\,n$ and $N_i, N_{\leq i}$ denote the number of cycles of length i and at most i respectively in $G(n,m)$. We call a cycle is *short* if it has length at most k. We claim the expected number of short cycles is bounded by a constant which doesn't depend on n.

Claim. $E[N_{\leq k}] \leq c$ where c does not depend on n.

Proof of the claim. Suppose $X_{i_1, i_2, \ldots, i_j}$ be the indicator variable for the event that the cycle $i_1, i_2, \ldots, i_j$ appears in G. For every $j \leq k$, we have

$$E[N_j] = \sum_{i_1, \ldots, i_j} E[X_{i_1, \ldots, i_j}] \leq n^j \left(\frac{m}{\binom{n}{2}}\right)^j \leq \left(\frac{m}{n-1}\right)^j 2^j \leq (2+2\,\epsilon)^j \left(\frac{n}{n-1}\right)^j$$
$$\leq (2+2\,\epsilon)^j\, e^{\frac{j}{n-1}} \leq (2\,e + 2\,e\,\epsilon)^j.$$

Hence $N_{\leq k} \leq k\,(2\,e + 2\,e\,\epsilon)^k$. Setting $c := k\,(2\,e + 2\,e\,\epsilon)^k$ proves the claim.

Let c be as above, then by Markov's inequality, we have

$$\Pr[N_{\leq k} \geq \log n] \leq \frac{E[N_{\leq k}]}{\log n} \leq \frac{c}{\log n} = o(1).$$

So a.a.s. G contains at most $\log n$ short cycles. We remove one edge per short cycle and let G' denote the resulting graph. Hence a.a.s. G' contains at least

$$m - \log n \geq (1+\epsilon)\,n - \log n \geq (1 + \frac{3\,\epsilon}{4})\,n \qquad (2)$$

edges.

Note that the girth of G' is at least k. If G' is planar then by inequality (1), $|E(G')| \leq (1 + \frac{\epsilon}{2})\,n$. But a.a.s. G' has more edges (as shown in (2)) and thus cannot be planar. This means that a.a.s G is non-planar as well. □

Exercise 3.4. (cf. [10] Solution 1.4) Prove that asymptotically almost surely:

$$\alpha(G(n, 1/2)) = \omega(G(n, 1/2)) = (1 - o(1))2\log_2 n.$$

Proof. Because the probability of having an edge or non-edge are equal, so we just need to prove one of statements. Let us consider the statement for clique number.

Let X denote the number of cliques of size k in $G(n,1/2)$. Then

$$E[X] = \binom{n}{k} 2^{-\binom{k}{2}}$$

The second moment of X is

$$E[X^2] = \sum_{i=0}^{k} \binom{n}{k}\binom{k}{i}\binom{n-k}{k-i} 2^{-2\binom{k}{2}+\binom{1}{2}}.$$

and the remaining term of the variance is given by

$$(EX)^2 = \sum_{i=0}^{k} \binom{n}{k}\binom{k}{i}\binom{n-k}{k-i} 2^{-2\binom{k}{2}}.$$

By the second moment method (cf. **exercise 3.1(a)**),

$$\Pr[|X - E[X]| \geq E[X]] \leq \frac{\mathrm{Var}(X)}{(EX)^2} = \frac{\sum_{i=2}^{k} \binom{n}{k}\binom{k}{i}\binom{n-k}{k-i} 2^{-2\binom{k}{2}} (2^{\binom{i}{2}} - 1)}{\sum_{i=0}^{k} \binom{n}{k}\binom{k}{i}\binom{n-k}{k-i} 2^{-2\binom{k}{2}}}$$

$$= \frac{\sum_{i=2}^{k} \binom{k}{i}\binom{n-k}{k-i}(2^{\binom{i}{2}} - 1)}{\binom{n}{k}} = o(1).$$

Hence, $X \sim E[X]$ for $n \to \infty$. Now set $k = \lceil 2\log_2 n \rceil$. Then

$$E[X] = \binom{n}{k} 2^{-\binom{k}{2}} \leq \left(\frac{en}{k}\right)^k 2^{-\frac{k^2}{2}+\frac{k}{2}} = \left(\frac{en\,2^{-k/2}\sqrt{2}}{k}\right)^k \leq \left(\frac{e\sqrt{2}}{k}\right)^k = o(1)$$

On the other hand, we could set for example $k = \lceil 2\log_2 n - 2\log_2 2\log_2 n \rceil$, and then

$$E[X] \geq \left(\frac{n}{k}\right)^k 2^{-k^2/2+k/2} = \left(\frac{n\,2^{-k/2}\sqrt{2}}{k}\right)^k \geq \left(\frac{2\sqrt{2}\log_2 n}{2\log_2 n - 2\log_2 2\log_2 n}\right)^k \geq \sqrt{2}^k \to \infty. \qquad \square$$

Exercise 3.5. (cf. [2] Thm 1.7) Let A be a monotone increasing graph property, s.t. the empty graph does not have A and the complete graph has A.

a) Show that there exists p^* such that $\Pr[G(n,p^*) \in A] = 1/2$.

b) Use staged exposure to prove $\Pr[G(n,kp) \notin A] \leq (\Pr[G(n,p) \notin A])^k$.

c) Let ω be a function tending to infinity and $\omega = o(n)$. Deduce that

$$\lim \Pr[G(n,wp^*) \in A] = 1, \quad \lim \Pr[G(n,p^*/\omega) \in A] = 0$$

and hence p^* is a *threshold function* for A.

Proof. Suppose A be a monotone increasing graph property, s.t. the empty graph does not have A and the complete graph has A.

(a) Note that

$$\Pr[G(n,p) \in A] = \sum_{G \in A} p^{|E(G)|} (1-p)^{\binom{n}{2} - |E(G)|}$$

This is a polynomial in p. Since A is monotone increasing, so that increasing p increases the likelihood that $G(n,p) \in A$, i.e. the probability $\Pr[G(n,p) \in A]$ increases from 0 to 1. Then for every given $0 < \epsilon < 1$ we define $p(\epsilon)$ by $\Pr[G(n, p(\epsilon)) \in A] = \epsilon$, and $p(\epsilon)$ exists because of above monotone property of $\Pr[G(n,p) \in A]$. In particular, we set $\epsilon = 1/2$, then there exists $p^* = p(1/2)$ such that $\Pr[G(n,p^*) \in A] = 1/2$.

(b) Let $G_1, \ldots, G_k$ be independent random graphs of $G(n,p)$. Then the union graph $\cup_{i=1}^k G_i$ is distributed as $G(n, 1-(1-p)^k)$. Recall $1-(1-p)^k \leq kp$, and by the coupling argument, we have $G(n, 1-(1-p)^k) \subseteq G(n, kp)$, which means that $G(n, kp) \notin A$ implies $\cup_{i=1}^k G_i \notin A$. Hence

$$\Pr[G(n, kp) \notin A] \leq \Pr[G(n, 1-(1-p)^k) \notin A] = (\Pr[G(n,p) \notin A])^k.$$

(c) Now we will show $p^* = p(1/2)$ is a threshold for A. Let ω be a function tending to infinity with $\omega = o(n)$. Setting $k = \omega$, we get

$$\Pr[G(n, \omega p^*) \notin A] \leq (\Pr[G(n,p^*) \notin A])^{\omega} = \left(\frac{1}{2}\right)^{\omega} = o(1),$$

i.e.

$$\lim \Pr[G(n, \omega p^*) \in A] = 1.$$

On the other hand for $p = p^*/\omega$,

$$\frac{1}{2} = \Pr[G(n,p^*) \notin A] \leq (\Pr[G(n, p^*/\omega) \notin A])^{\omega}.$$

So

$$\Pr[G(n, p^*/\omega) \notin A] \geq (1/2)^{1/\omega} = 1 - o(1),$$

i.e.

$$\lim \Pr[G(n, p^*/\omega) \in A] = 0. \qquad \square$$

4 Long Paths and Hamilton cycles

Note 4.1. This section is almost based on [1] Section 3. One can also refer to [2] Section 6.2, 6.3. An alternative reference is [11].

A **trail** is a walk in which all edges are distinct. A **path** is a trail in which all vertices (and therefore also all edges) are distinct.

Theorem 4.2. (Ajtai, Komlós, Szemeral, 1981) *There is a function $\alpha(c)$: $(1,\infty)\to(0,1)$, such that $\lim_{c\to\infty}\alpha(c)=1$. A random graph $G\left(n,p=\frac{c}{n}\right)$ a.a.s. has a path of size at least $\alpha(c)\cdot n$.*

Recall that a path is Hamilton, if its length is $n-1$. A cycle is Hamilton if its length is n. A graph G is Hamiltonian if it contains a Hamilton cycle.

Definition 4.3. *Let $G=(V,E)$ and $e=\{u,v\}\notin E$. If adding e to E results in a graph $G'=G+e$ s.t. $\ell(G')>\ell(G)$, or if G' is Hamiltonian, then e is called a booster.*

Remark 4.4. If G is Hamiltonian, then every $e\notin E$ is a booster. Adding n boosters results in a Hamiltonian graph. (增加n次booster就一定得到哈密顿图.)

Proposition 4.5. *G is a connected graph, $P=(v_0,v_1,\ldots,v_\ell)$ is a longest path in G. If $\{v_0,v_\ell\}\notin E$, then it is a booster.*

Proof. Let G' be the graph G by adding the edge $\{v_0,v_\ell\}$. If G' contains the cycle C is Hamiltonian, then we've done. If not, we could find a longer path in G'. In fact, there exists $w\in V, w\notin P$. From connectivity, w is connected to C, say by $w_1,\ldots,$ $w_k=v_i\in C$. Consider the path

$$w_1,\ldots,w_k=v_i,v_{i+1},\ldots,v_l,v_0,\ldots,v_{i-1}.$$

It contains all vertices of P and $w\notin P$. Hence we have a longer path. So $\{v_0.v_\ell\}$ is a booster. □

4.1 Rotation

This Algorithm is given by Posa in 1976.

We want to prove the following result:

Proposition 4.6. $G\left(n,p=C\frac{\log n}{n}\right)$ *is a.a.s. Hamiltonian for larger enough constant C.*

Algorithm 4.1

1. Assume $P=(v_0,v_1,\ldots,v_\ell)$ is a longest path. (不是最长找到最长)
2. For every edge $e\in E$ containing v_ℓ, the other endpoint is on p, otherwise we can extend. If $e=(v_1,v_l)$ with $1\le i\le l-2$, then we can rotate p at v_i by adding the edge e (to p) and remove (v_i,v_{i+1}) (from p). new path $p'=(v_0,v_1\ldots,v_{i-1},$ $v_\ell,v_{\ell-1},\ldots,v_{i+1})$.

Definition 4.7. *For a set of vertices U, $N(U)$, the external neighborhood of U, s.t.*

$$N(U)=\{v\in V\setminus U|\,\exists u\in U,\{u,v\}\in E\}.$$

P is a path, $R\subseteq P$. Denote $R^-=\{v_i|\,v_{i+1}\in R\}$ and $R^+=\{v_i|\,v_{i-1}\in R\}$.

Remark. $|R^-|\leq|R|$ and $|R^+|\leq|R|$. If $v_l\in R$, then $|R^+|=|R|-1$.

Lemma 4.8. (Rosa) *G is a graph. P a longest path. Let R be the set of endpoints of paths obtained from p by a sequences of rotations. Then $N(R)\subseteq R^-\cup R^+$.*

Proof. take any $v\in R$, $u\in V(G)\setminus(R\cup R^-\cup R^+)$. We want to show $\{v,u\}\notin E$.

If $u\notin V(P)$. Consider Q, a path obtained from P by a sequence of rotations and Q's endpoint is v. If $\{u,v\}\in E$, we could have extend Q, contradicting the maximality of P.

If $u\in V(P)$. If $u\in V(G)\setminus(R\cup R^-\cup R^+)$, then u has the exact same neighbourhoods in every path p' obtained from p by sequences of rotations. Because when rotating at v_i only v_i,v_{i+1} and v_ℓ change their neighbors, but $v_\ell\in R$, $v_{i+1}\in R$, and $v_i\in R^-\cup R^+$. If $\{u,v\}\in E$ then we can rotate at u, and a neighbor of u in Q, w, becomes an endpoint. But then since w was a neighbor of u in P, $u\in R^-\cup R^+$. □

4.2 Expansion

Definition 4.9. (expander) $k\in\mathbb{N}$, $t>0$ *a real number. G is called a (k,t)-expander if for any set U, $|U|\leq k$,*

$$|N(U)|>t|U|.$$

Proposition 4.10. *Let G be a $(k,2)$-expander. Then G has a path of length at least $3k-1$. If G is connected and non-Hamiltonian, then G has at east $(k+1)^2/2$ boosters.*

Note 4.11. Random graph is a good expander.

Proof. Let $P=(v_0,\ldots,v_l)$ be longest path in G. Let R be the set of endpoints of paths obtained from P by sequences of Posa rotations. Then $|R^-|\leq R$ and $|R^+|\leq|R|-1$. By Posa's lemma 4.8,

$$N(R)\subseteq R^-\cup R^+.$$

$$|N(R)|\leq|R^-|+|R^+|\leq 2|R|-1$$

Thus $|R|>k$. Let $R_0\subset R$ s.t. $|R_0|=k$. Now

$$N(R_0)\subset R\cup R^-\cup R^+,$$

hence $N(R_0)\subseteq V(P)$. $|V(R_0)|\geq 2|R_0|=2k$. However, R_0 and $N(R_0)$ are in P. $|V(P)|\geq k+2k=3k$. Hence, the length of P is at least $3k-1$.

Assume that G is connected and non-Hamiltonian. If $v\in R$, then $(v_0,v)\notin E(G)$, thus (v_0,v) is a booster.

Since $|R|>k$, we found at least $k+1$ boosters. Thinking of terminal points as starting points, we get for each such vertex, at least $k+1$ boosters. Since each booster was counted at most twice, we have found at least $(k+1)^2/2$ boosters. $\square$

Proof of the negative part of the threshold for Hamiltonicity.

Proposition 4.12. *Let $k\geq 1$ be a fixed integer.*

1. *If $p(n)=\frac{\log(n)+(k-1)\log\log n-\omega(n)}{n}$ then a.a.s. $\delta(G(n,p))\leq k-1$. (δ is the minimal degree.)*

2. *If $p(n)=\frac{\log(n)+(k-1)\log\log n+\omega(n)}{n}$ then a.a.s. $\delta(G(n,p))\geq k$.*

In particular, if $p(n)=\frac{\log n+\log\log n-\omega(n)}{n}$ then a.a.s. $\delta(G(n,p))\leq 1$, and G does not contain a Hamilton cycle.

Proof. (Exercise 2.1, cf. Chapter 3 in [2]) For part a) use the first and second moment method, for part b) a first moment calculation together with a union bound.

1. Let X_d be the number of vertices of degree d in $G(n,p)$. The degree of an individual vertex of $G(n,p)$ is a Binomial random variable with parameters $n-1$ and p. Hence,

$$\mathrm{E}[X_d]=n\binom{n-1}{d}p^d(1-p)^{n-1-d}.$$

Therefore, as $n\to\infty$,

$$\mathrm{E}[X_d]\to\begin{cases}0 & \text{if } p\ll n^{-(d+1)/d},\\ \infty & \text{if } p\gg n^{-(d+1)/d} \text{ and } pn-\ln n-d\ln\ln n\to-\infty,\\ 0 & \text{if } pn-\ln n-d\ln\ln n\to\infty.\end{cases} \tag{4.1}$$

Let X be the number of vertices of degree $k-1$ in $G(n,p)$, and $p=\frac{\ln n+(k-1)\ln\ln n-w(n)}{n}$, then $E[X]\to\infty$. We have to compute the second moment. For this we need to estimate,

$$\begin{aligned}\Pr[d(i)=d(j)=k-1] &= p\left(\binom{n-2}{k-2}p^{k-2}(1-p)^{n-k}\right)^2\\ &\quad+(1-p)\left(\binom{n-2}{k-1}p^{k-1}(1-p)^{n-k-1}\right)^2\\ &= \Pr[d(i)=k-1]\cdot\Pr[d(j)=k-1](1+o(1)).\end{aligned}$$

The first line here accounts for the case where $\{i,j\}$ is an edge and the second line deals with the case where it is not. Thus,

$$\begin{aligned}\mathrm{Var}[X] &= \sum_{i=1}^{n}\sum_{j=1}^{n}\{\Pr[d(i)=d(j)=k-1]-\Pr[d(i)=k-1]\Pr[d(j)=k-1]\}\\ &\leqslant E[X]+\sum_{i\neq j=1}^{n}o(1)\end{aligned}$$

Note that $\frac{\mathrm{Var}[X]}{E^2[X]} \leqslant \frac{1}{E[X]} + o(1) = o(1)$. Therefore, by the second moment inequality, we have

$$\Pr[X \geqslant 1] = 1 - o(1).$$

This means that a.a.s. for $G \sim G(n,p)$ one has $\delta(G) \leqslant k-1$.

2. For $p = \frac{\ln n + (k-1)\ln\ln n + w(n)}{n}$, we have $E[X_{k-1}] \to 0$ by (4.1). Hence, a.a.s. there is no degree $k-1$ vertices. Similarly, $E[X_i] \to 0$ for all $0 \leqslant i \leqslant k-1$, $E[X_{k-1}] \to 0$. i.e., a.a.s. there is no degree $\leqslant k-1$ vertices. Thus, $\delta(G) \geqslant k$. ☐

$G \sim G(n,p)$, we can represent G as $G = G_1 \cup G_2$, $G_i \sim G(n, p_i)$ we want $p_2 = c/n$ for large enough $c > 0$.

Since $(1-p_1)(1-p_2) = 1-p$, we get $p_1 \geq p - p_2$, thus $p_1 = \frac{\log(n) + (k-1)\log\log n + \omega(1)}{n}$.

Lemma 4.13. *Let $p = \frac{\log(n) + (k-1)\log\log n + \omega(1)}{n}$. then a.a.s. $G(n,p)$ is an (n/4,2) expander.*

We claim the following two propositions, and then finish the proof.

Proposition 4.14. (Proposition 3.13 in [1]) *Let $p = \frac{\log n + \log\log n + \omega(1)}{n}$, then a.a.s. G(n,p) has the following properties:*

1. *$\delta(G) \geqslant 2$. Let $\mathrm{Small} = \{v \in [n] | d(v) \leqslant (\log n)^{7/8}\}$.*
2. *no vertex of small lies on a cycle of length $\leqslant 4$;*

   G does not contain a path between two vertices of Small of length at most 4.
3. *for every pair of sets A, B s.t. $|A|, |B| \geqslant \frac{n}{\sqrt{\log n}}$, there is an edge between A and B.*
4. *Every set $V_0 \subseteq [n]$ of size $\leqslant \frac{2n}{\log^{3/8} n}$ spans at most $3|V_0|\log^{5/8} n$ edges.*

Proof. (Exercise 2.2) (a) From Proposition 4.12, we have $\delta(G) \geqslant 2$ for $k = 1$.

(b) Recall that for any fixed vertex v, its degree $d(v)$ in $G(n,p)$ is $\mathrm{Bin}(n-1,p)$ distributed. Note that for every constant $c > 0$ and for every $k \leqslant (\ln n)^{\varepsilon}$, we have

$$\begin{aligned}
\Pr[\mathrm{Bin}(n-c,p) \leq k] \leq & (k+1)\Pr[\mathrm{Bin}(n-c,p) = k] \\
= & (k+1)\binom{n-c}{k} p^k (1-p)^{n-c-k} \\
\leq & \ln n (np)^k e^{-pn} e^{(k+c)p} \\
\leq & \ln n (2\ln n)^k \frac{1}{n} \\
\leq & n^{-0.9},
\end{aligned}$$

We claim that a.a.s. no vertex of Small lies on a triangle. For $u,v,w\in V[G]$ distinct denote $A_{u,v,w}$ be the event $u\in\text{Small}$ and $(u,v),(v,w),(w,u)\in E(G)$. Then we have

$$\begin{aligned}\Pr[A_{u,v,w}] &= p^3\,(p\,(\Pr[\text{Bin}(n-3,p)\leqslant\log^{7/8}n-2])^2+\\ &\quad +p^3\,((1-p)\Pr[\text{Bin}(n-3,p)\leqslant\log^{7/8}n-1])^2\\ &\leqslant p^3n^{-0.9\times2}\leqslant n^{-4.7}.\end{aligned}$$

Then, $\Pr[\exists u,v,w\,A_{u,v,w}]=\Pr[\bigcup_{u,v,w}A_{u,v,w}]\leqslant\binom{n}{3}n^{-4.7}=o\left(\frac{1}{n}\right)$. Similarly for the case C_4. Hence, a.a.s. no vertex of Small lies on a triangle or a C_4.

We first claim that a.a.s. Small is an independent set. For any two vertices $u\neq v$ of G, let $A_{u,v}$ be the event that $u,v\in\text{Small}$ and $(u,v)\in E[G]$. Then we have

$$\Pr[A_{u,v}]=p\,(\Pr[\text{Bin}(n-2,p)\leqslant\log^{7/8}n-1])^2\leqslant pn^{-1.8}\leqslant n^{-2.7}.$$

Hence, we have

$$\Pr[\text{Small is not independent}]=\Pr\left[\bigcup_{u\neq v}A_{u,v}\right]\leqslant\binom{n}{2}n^{-2.7}=o(1).$$

We claim that a.a.s. every two vertices in Small do not have a common neighbor (i.e., of distance 3). For $u,v,w\in V(G)$ distinct denote $A_{u,v,w}$ be the event $u,v\in\text{Small}$ and $(u,w),(v,w)\in E(G)$. Then we have

$$\begin{aligned}\Pr[A_{u,v,w}] &= p^2\,(p\,(\Pr[\text{Bin}(n-3,p)\leqslant\log^{7/8}n-2])^2+\\ &\quad +p^2\,((1-p)\Pr[\text{Bin}(n-3,p)\leqslant\log^{7/8}n-1])^2\\ &\leqslant p^2n^{-1.8}\leqslant n^{-3.7}.\end{aligned}$$

Then

$$\Pr[\exists u,v,w\,A_{u,v,w}]=\Pr\left[\bigcup_{u,v,w}A_{u,v,w}\right]\leqslant\binom{n}{3}n^{-3.7}=o(1).$$

We may similarly prove the case of distance four for any two vertices $u,v\in\text{Small}$. Let $A_{u,v,w,x}$ be the event, then $\Pr[\exists u,v,w,x\,A_{u,v,w,x}]=o(1)$. Hence, a.a.s. no vertices of Small are of distance at most 4.

(c) Let X be the event that there is no edge between in A and B. Then

$$\begin{aligned}\Pr[X] &= \Pr\left[\exists A,B\text{ disjoint},|A|,|B|\geqslant\frac{n}{\sqrt{\log n}},E_G(A,B)=0\right]\\ &\leqslant \binom{n}{\frac{n}{\sqrt{\log n}}}(1-p)^{n^2/\log n}\\ &\leqslant \left(e\sqrt{\log n}\right)^{\frac{2n}{\sqrt{\log n}}}e^{-\frac{pn^2}{\log n}}\\ &= o(1)\end{aligned}$$

Hence, for every vertex subsets A, B of sizes at least $\frac{n}{\sqrt{\log n}}$, there is an edge between a vertex in A and a vertex in B.
(d) Denote $N = \frac{2n}{\log^{3/8} n}$, then we have

$$\begin{aligned}
\Pr[\exists V_0 \subset V(G),\ |V_0| \leqslant N,\ E(V_0) \geqslant 3|V_0| \log^{5/8} n] &\leqslant \sum_{k \leqslant N} \binom{n}{k} \Pr\left[\mathrm{Bin}\left(\binom{k}{2}, p\right) \geqslant 3k \log^{5/8} n\right] \\
&\leqslant \sum_{k \leqslant N} \binom{n}{k} \binom{\binom{k}{2}}{2k \log^{5/8} n} p^{3k \log^{5/8} n} \\
&\leqslant \sum_{k \leqslant N} \left(\frac{en}{k}\left(\frac{ekp}{6 \log^{5/8} n}\right)^{3\log^{5/8} n}\right)^k
\end{aligned}$$

If $k \leqslant \sqrt{n}$, then

$$\left(\frac{en}{k}\left(\frac{ekp}{6 \log^{5/8} n}\right)^{3\log^{5/8} n}\right)^k \leqslant \left(enn^{-\frac{1}{3} 3\log^{5/8} n}\right)^k = o\left(\frac{1}{n}\right).$$

If $\sqrt{n} < k \leq N$, then

$$\begin{aligned}
\left(\frac{en}{k}\left(\frac{ekp}{6 \log^{5/8} n}\right)^{3\log^{5/8} n}\right)^k &= \left(\frac{en}{k} \frac{ekp}{6 \log^{5/8} n}\left(\frac{ekp}{6 \log^{5/8} n}\right)^{3\log^{5/8} n - 1}\right)^k \\
&\leqslant \left(\log n\left(\frac{ekp}{6 \log^{5/8} n}\right)^{2\log^{5/8} n}\right)^k \\
&\leqslant (\log n\, 0.95^{2\log^{3/8} n})^{\sqrt{n}} \\
&= o\left(\frac{1}{n}\right).
\end{aligned}$$

Therefore, it holds a.a.s. □

Proposition 4.15. *If G satisfies properties $(1)-(4)$ above, then G is an $\left(\frac{n}{4}, 2\right)$-expander for large enough n.*

Notation 4.16. *$\Delta(G)$ is the maximum degree of a vertex in G, and $\delta(G)$ is the minimum degree.*

Lemma 4.17. *Let G_1 be an $(n/4, 2)$ expander, and let $G_2 \sim G(n, p_2)$ with $p_2 = 80/n$. (80>64) Then a.a.s. $G = G_1 \cup G_2$ is Hamiltonian.*

Proof. (Lemma 4.17) G_1 is connected. If C is a connected component of G_1, $N(C) = \varnothing$, thus $c > n/4$. Choose $V_0 \subseteq C$, $|V_0| = n/4$. $N(V_0) \geq n/2$ and $N(V_0) \subseteq C$. Then $|C| \geq \frac{3}{4} n$. There is room for just one connected component.
Represent $G_2 = \cup_{j=1}^{2n} G_{2,j}$, where $(1 - 2p)^{2n} = 1 - p_2 = 1 - 80/n$. $p \geq p_2/(2n) = 40/n^2$.
Denote by $H_j = G_1 \cup \cup_{k=1}^{j} G_{2,k}$, we say that round j is successful if H_{j-1} is Hamiltonian or $E(G_{2,j})$ hits a booster of H_{j-1}.

We need to have at least n successful rounds. Consider round j, either H_{j-1} is Hamiltonian or since $G_1 \subseteq H_{j-1}$ and G_1 is connected and it is $(n/4, 2)$ expander, H_{j-1} has at least $(n/4)^2/2 = n^2/32$ boosters.

$\Pr[\text{round } j \text{ is successful}] \geq 1-(1-p)^{n^2/32} \geq 1-e^{-pn^2/32} \geq 1-e^{-40/32} = 1-e^{-3/4} \geq 2/3$.

Let X be a random variable counting successful rounds. X stochastically dominated a binomial random variable $\mathrm{Bin}(2n, 2/3)$, a.a.s. $X \geq n$. □

Theorem 4.18. (cf. Theorem 3.6 in [1]) *The threshold for Hamiltonicity is at*

$$p(n) = \frac{\log n + \log\log n}{n}.$$

5 The Chromatic Number of Random Graphs

Note 5.1. This section is almost based on [1] Section 5.1–5.4.

Let G be a graph. A set $I \subseteq V(G)$ is independent if the induced subgraph $G[I]$ is empty. The size of a largest independent set is denoted by $\alpha(G)$, is called the independence number of G.

A partition $V = \cup C_1 \cup C_2 \cup \cdots \cup C_k$ is a k-coloring of G is each C_i is an independent set. Equivalently, a function $c: v \rightarrow \{1,\ldots,k\}$ is a k-coloring if for every edges $e = (u,v)$ we have $c(u) \neq c(v)$.

If there is a k-coloring of G then G is k-colorable.

The chromatic number of G, $\chi(G)$, is the smallest k for which G is k-colorable.

Example 5.2. $\alpha(K_n) = 1$, $\chi(K_n) = n$. $\alpha(K_{m,n}) = \max\{m,n\}$, $\chi(K_{m,n}) = 2$.

Then $\chi(G) \geqslant \frac{V(G)}{\alpha(G)}$ because each color class is of size $\leqslant \alpha(G)$.

greedy coloring: a coloring of the vertices of a graph formed by a greedy algorithm that considers the vertices of the graph in sequence and assigns each vertex its first available color.

按序列顺序染色（每次新增染色数尽量少），不过这样反而会增加染色数。

Example 5.3. the following graph can be 2-color (black and orange), but for greedy coloring, it can be 3-color instead (1,2,3).

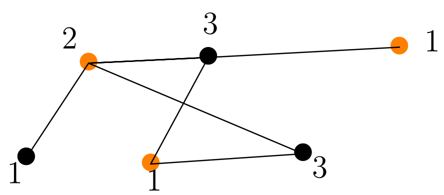

Figure 5.1. coloring and greedy coloring

A greedy coloring may require n colors even though $\chi(K_{n,n} - M) = 2$. (M is a matching)

For any graph G there is an ordering of the vertices s.t. the greedy coloring with this ordering uses $\chi(G)$ colors.

5.1 Coloring $G(n, 1/2)$

We will first consider the lower bound for $\chi(G(n, 1/2))$, $f(k) = \binom{n}{k}(1/2)^{\binom{k}{2}}$ the expected number of independent sets of size k. Then

$$\frac{f(k+1)}{f(k)} = \frac{n-k}{k+1} 2^{-k}.$$

The maximum $f(k)$ occurs when k is roughly $\log_2 n$. Let $k^* = \max\{k \mid f(k) \geqslant 1\}$.

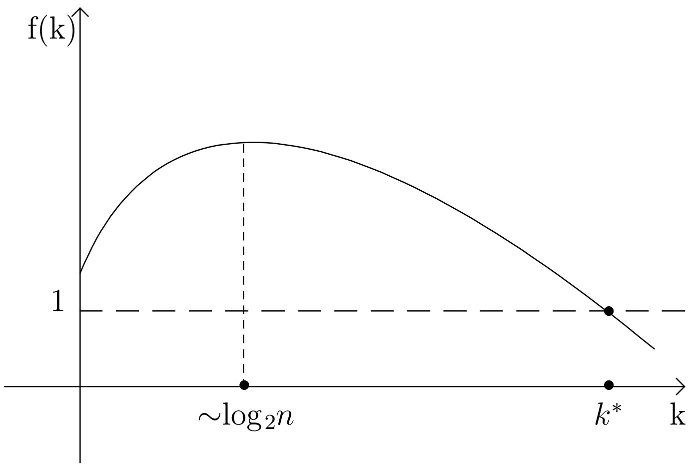

Figure 5.2. graph of f(k)

then $k^* = 2\log_2 n - 2\log_2\log_2 n + \Theta(1)$, If $k \approx k^*$, $\frac{f(k+1)}{f(k)} = \tilde{\Theta}\left(\frac{1}{n}\right)$, the tilde means up to polylogarithmic factors.

Proposition 5.4. *a.a.s. in $G(n, 1/2)$, $\alpha(G) < k^* + 2$.*

Proof. $f(k^*+1) < 1$ by definition of k^*. $f(k^*+2) = f(k^*+1) \cdot \tilde{\Theta}(1/n) = o(1)$. So the expected number of independent sets of size k^*+2 is $o(1)$. By Markov inequality, a.a.s. there is no such set. □

Corollary 5.5. *a.a.s. in $G(n, 1/2)$, then*

$$\chi(G) \geqslant \frac{n}{2\log_n n - 2\log_2\log_2 n + \Theta(1)}$$

Now we consider the upper bound.

Theorem 5.6. (Bollobas, 1988) *a.a.s. in $G(n,1/2)$, $\chi(G)=\frac{n}{2\log_2 n}(1+o(1))$.*

5.1.1 Jensen Inequalities

A set Ω, a random subset $R\subseteq\Omega, r\in\Omega$ and $P[r\in R]=P_r$ independently of other choices.

A family of subsets $A_i\subseteq\Omega$, $\{A_i\}_{i\in I}$. We are interested in estimating $\Pr[\text{no } A_i \text{ is a subset of } R]$.

Let $X_i=\delta_{A_i\subseteq R}$ (indicator function), then

$$\Pr[X_i=1]=E[X_i]=\prod_{r\in A_i} P_r$$

Let $X=\sum_{i\in I} X_i$,

$$\mu=E[X]=\sum_{i\in I} E[X_i]=\sum_{i\in I}\prod_{r\in A_i} P_r.$$

If the set A_i are pairwise disjoint, then the events $P[A_i\subset R]$ are independent.

$$\Pr[X=0]=\prod_{i\in I}\Pr[X_i=0]=\prod_{i\in I}\left(1-\prod_{r\in A_i}P_r\right)\leqslant\prod_{i\in I}\exp\left(-\prod_{r\in A_i}P_r\right)=\exp(-\mu).$$

Remark 5.7. The Poisson paradigm for a counting random variable X, $E[X]=\mu$, then $\Pr[X=0]\approx e^{-\mu}$ (as in the poisson distribution).

If we consider dependence, i.e. A_i are not disjoint. The Janson inequality gives us a measurement of dependence.

Define $i\sim j$ if $A_i\cap A_j\neq\varnothing$. $\Delta=\sum_{i\sim j}\Pr[A_i\subseteq R \text{ and } A_j\subseteq R]$.

Example. $\Omega=E(K_n)$ (edges of K_n), $R=G(n,p)$.

Theorem 5.8. (The Janson inequality) *with the above notation,*

$$\Pr[X=0]\leqslant e^{-\mu+\frac{\Delta}{2}}$$

If in addition, $\Delta\geqslant\mu$, then

$$\Pr[X=0]\leqslant e^{-\frac{\mu^2}{2\Delta}}.$$

Proof cf. Alon and Spencer's book: The probabilistic method ([3] chapter 8.2).

5.2 Chromatic number of $G(n,1/2)$

Proposition 5.9. *G is a graph s.t. for some $1\leqslant k\leqslant m\leqslant n$, every set of m vertices spans an independent set of size k. Then $\chi(G)\leqslant\frac{n}{k}+m$.*

Proof. Start with $G' = G$, $i = 1$.

1. as long as $|V(G')| \geqslant m$, find an independent set of size k and color it with color i. Remove these k vertices from G'. $i = i + 1$;

2. color the remaining vertices with distinct new colors.

The first phase used at most $\frac{n}{k}$ colors, and the second phase used $< m$ colors, all in all at most used $\leqslant \frac{n}{k} + m$. $\square$

Set $m = \frac{n}{(\log_2 n)^2}$. we want to show that every set of m vertices spans an independent set of size $(1 - o(1))2\log_2 n$.

Lemma 5.10. *$G \sim G(n, 1/2)$, let $k^* = \max\left\{k \middle| \binom{m}{k} 2^{-\binom{k}{2}} \geqslant 1\right\}$. Set $k = k^* - 3$, then*

$$\Pr[\alpha(G) < k] = \exp\left(-\Omega\left(\frac{m^2}{k^4}\right)\right).$$

Proof. $f(k, m) = \binom{m}{k} 2^{-\binom{k}{2}} \geqslant m^{3+o(1)}$ since for $k \approx k^*$, $\frac{f(k+1, m)}{f(k, m)} = m^{-1+o(1)}$, and $k = k^* - 3$. (In fact, we can generalize it for $k \leqslant k^* - 3$.)

Then we want to apply the Janson inequality, $\Omega = E(K_m)$, R is the set of m edges in $G(m, 1/2)$, $\Pr[r \in R] = \frac{1}{2}$ for all $r \in R$.

Let $S \subseteq [m]$, $|S| = k$, $A_s = \binom{s}{2} = \{\{s_1, s_2\} | s_1 \neq s_2 \in S\}$ is the set of $\binom{k}{2}$ pairs in S. Then

$$X_s = \begin{cases} 1 & A_s \subseteq R \Longleftrightarrow S \text{ spans an independent set} \\ 0 & \text{otherwise} \end{cases}.$$

$X_s = 1$ iff S is an independent set in $G(m, 1/2)$. Let $X = \sum_{S \subseteq [m], |S| = k} X_s$ count k-independent sets in $G(m, 1/2)$.

$$\Pr[\alpha(G(m, 1/2)) < k] = \Pr[X = 0].$$

Then

$$\mathrm{E}[x] = \mu = \sum_{S \subseteq m, |S| = k} \mathrm{E}[X_S] = \binom{m}{k} 2^{-\binom{k}{2}} \geqslant m^{3+o(1)}.$$

$S \sim S'$, with $2 \leqslant |S \cap S'| \leqslant k - 1$ for $|S| = |S'| = k$, both subsets of $[m]$.

$$\begin{aligned}
\Delta = \sum_{S \sim S'} \Pr[(X_S = 1) \wedge (X_{S'} = 1)] &= \sum_{S \subseteq [m], |S| = k} \sum_{i=2}^{k-1} \binom{k}{i} \binom{m-k}{k-i} \left(\frac{1}{2}\right)^{\binom{k}{2} + \binom{k}{2} - \binom{i}{2}} \\
&= \binom{m}{k} \sum_{i=2}^{k-1} \binom{k}{i} \binom{m-k}{k-i} 2^{-2\binom{k}{2} + \binom{i}{2}} \\
&= \binom{m}{k} 2^{-\binom{k}{2}} \sum_{i=2}^{k-1} \binom{k}{i} \binom{m-k}{k-i} 2^{-\binom{k}{2} + \binom{i}{2}}
\end{aligned}$$

$$= \mu \sum_{i=2}^{k-1} g(i)$$

where $g(i) = \binom{k}{i}\binom{m-k}{k-i}\left(\frac{1}{2}\right)^{\binom{k}{2}-\binom{i}{2}}$. We want to argue that most mass of $g(i)$ is on $g(2)$. Namely, $\sum_{i=2}^{k-1} g(i) = (1+o(1))g(2)$.

In fact, $g(i+1)/g(i)$ shows that $g(i)$ is unimodal (first decreasing then increasing), then

$$\sum_{i=2}^{k-1} g(i) = (1+o(1))(g(2)+g(k-1)),$$

Then $g(2) = \Theta\left(\frac{k^4}{m^2}\mu\right)$, $g(k-1) = \tilde{O}\left(\frac{1}{m}\right)$, $\Rightarrow$ $g(2) \gg g(k-1)$. Hence, $\sum_{i=2}^{k-1} g(i) = (1+o(1))(g(2))$.

Hence, $\Delta = \mu g(2)(1+o(1)) = \Theta\left(\frac{k^4}{m^2}\mu^2\right)$. $\frac{\Delta}{\mu} = \Theta\left(\frac{k^4}{m^2}\mu\right) \gg 1$. Then by the second case of the Janson inequality, $\Pr[X=0] \leqslant \exp\left(-\frac{\mu^2}{2\Delta}\right)$.

$\Pr[\alpha(G(n, 1/2) < k)] = \Pr[X = 0] \leqslant e^{-\frac{\mu^2}{2\Delta}} = \exp\left(-\Theta\left(\frac{m^2}{k^4\mu^2}\mu^2\right)\right) = \exp\left(-\Theta\left(\frac{m^2}{k^4}\right)\right)$. □

Proof. (Theorem 5.6) By the last lemma, $G \sim G(n, 1/2)$, for any set V_0 of m vertices

$$\Pr[\alpha(G(V_0)) < k] = \exp\left(-\Omega\left(\frac{m^2}{k^4}\right)\right).$$

By the union bound,

$$\Pr[\exists V_0 \subseteq [n], |V_0| = m | \alpha(G[V_0]) < k] \leqslant \binom{n}{m}\exp\left(-\Omega\left(\frac{m^2}{k^4}\right)\right) \leqslant 2^n \exp\left(-\frac{cn^2}{(\log_2 n)^{2\times 4}}\right) = o(1)$$

Hence, a.a.s., $G(n, 1/2)$ has the properties of the claim with $m = \frac{n}{(\log_2 n)^2}$ and $k = k^*(m) - 3$,

$$\chi(G(n,1/2)) \leqslant \frac{n}{k} + m = \frac{n}{2\log_2 m(1+o(1))} + m = \frac{n}{(1+o(1))2\log_2 m} + \frac{n}{(\log_2 n)^2} = \frac{n}{2\log_2 n}(1+o(1)).$$ □

Theorem 5.11. (thm 7.9 in [2]) *The chromatic number of a greedy coloring of $G(n,p)$ is*

$$\chi^*(G) = \frac{n}{\log_2 n}(1+o(1)) = (2+o(1))\chi(G).$$

6 Logic of Graphs

$G(n, \frac{1}{2})$ is a.a.s. connected, but not k-colorable, not planar, Hamiltonian, containing any fixed H as a subgraph. So $G(n, \frac{1}{2})$ has a large set of graph properties.

Which property is "natural"?

Other property:

– The number of vertices is even.

– The number of edges is even.

6.1 First order language of graphs

Definition 6.1. (First order language)

- x, y, x_1, x_1 *variables represent vertices*
- $=, \sim$ *equality , adjacency* $x = y$ $x \sim z$
- $\wedge, \vee, \neg, \rightarrow, \leftrightarrow, ..$
- $\exists, \forall$

Example 6.2. $\exists x, y, z, x \neq y \wedge y \neq z \wedge z \neq x \wedge x \sim y \wedge y \sim z \wedge z \sim x.$

$G \vDash \varphi$ (G is a model of φ.) $\Longleftrightarrow G$ contains a triangle.

$\forall x, \exists y, x \sim y$ means no isolated vertices.

And only "containing any fixed H as a subgraph" is a first order language.

Remark 6.3. If φ is a sentence (a formula w.o. free variables), then it represents a graph property.

Definition 6.4. *If for a graph property P there is a sentence φ s.t. $G \vDash \varphi \Leftrightarrow G \in P$, then we call P a first order property.*

Theorem 6.5. (GKLT '69, Fagin '76) *The 0-1 laws holds for the first order language of graphs and $G(n, \frac{1}{2})$.*

That is: for every first order sentence φ, s.t.

$$\lim_{n \to \infty} \Pr\left[G\left(n, \frac{1}{2}\right) \vDash \varphi\right] \in \{0, 1\}.$$

We can replace $\frac{1}{2}$ with any constant $p \in (0,1)$.

If $p \to 0$ slowly enough then the $0-1$ laws holds for $G(n,p)$ and $G(n,1-p)$ as well.

Proof. The (k,ℓ)-extension property is $\forall x_1, x_2, \ldots, x_k, y_1, \ldots, y_\ell$ (Distinct x_i and y_j).

$\to \exists z, z \neq x \wedge z \neq x_2 \wedge \cdots \wedge z \neq y_\ell \wedge z \sim x_1 \wedge \cdots \wedge z \sim x_k \wedge z \nsim y_1 \wedge \cdots \wedge z \nsim y_\ell$.

Then (k,ℓ)-extension property is a 1st order property.

Claim: fix k, ℓ, and $p \in (0,1)$, a.a.s. $G(n,p)$ is a model for the (k,ℓ)-extension property.

Proof of the Claim: Given $u_1, \ldots, u_k$ and $v_1, \ldots, v_\ell$, the probability that there is no witness there vertices is

$$(1-p^k(1-p)^\ell)^{n-k-\ell} < c^n$$

for some constant $0 < c < 1$. Taking a union bound over all choices of $k+l$ vertices, we get

$$\binom{n}{k+l} c^n \leqslant n^{k+l} c^n \underset{n\to\infty}{\longrightarrow} 0.$$

This proves the claim.

Fix k, l, $G(n, p = \text{const})$ a.a.s. has the (k',ℓ')-extension property for all $0 \leqslant k' \leqslant k$ and $0 \leqslant l' \leqslant l$.

Then there is another union bound.

Notice that no finite graph can be a model for all the extension property together. Since a graph with no vertices can't satisfy the $(n,0)$-extension property.

Lemma. *there is a unique graph over $\mathbb{N}$ obeying (k,l)-extension properties for all $k, l \in N$.*

Proof. (lemma) (existence) (0,0)-ext, (1,0), (0,1), (2,0), ... are countable infinite list of extension property. If a.a.s. $G(n,p) \vDash \varphi_1$ .... $(G,p) \vDash \varphi_a$, then $G(n,p) \vDash \varphi_1 \wedge \cdots \wedge \varphi_a$. (resp.?) Start with empty graph.

Then the natural numbers make a list of all pairs X_i, Y_j where X_i, Y_j are disjoint finite subsets of $\mathbb{N}$.

We go over this list pick a vertex $z \notin X_i \cup Y_j$ that was not picked before and make it adjacent to all the vertices in X.

This (infinite) process defines a countable graph. By construction, it is a model for all (k,l)-extension property. This countable graph is called the Rado graph. $\square$

By the lemma we have one countable model (and no finite model).

$\varphi \in \mathrm{FOL}$, $G \vDash \varphi$ or $G \vDash \neg\varphi$.

There is a proof for $\varphi \ldots$, which is the end of the story. $\square$

Acknowledgement

This note is distributed to participants of the course Random Graphs by Dr. Simcha Haber (Simi) at Bar-Ilan University via Zoom in Spring 2020. I attended the course since the third lesson after Simi changed the teaching way into Zoom due to the corona virus pandemic.

Zoom recordings links are listed as follows:

Lectures	Date & Videos	Notes
1	08-mar-2020	@BIU
2	15-mar-2020	
3	22-mar-2020	
4	29-mar-2020	
5	19-apr-2020	
6	26-apr-2020	
7	03-may-2020	
8	10-may-2020	
9	17-may-2020	whiteboard
10	24-may-2020	
11	31-may-2020	
12	07-jun-2020	
13	14-jun-2020	
14	21-jun-2020	

Table 1. Zoom recordings links

References

1. Michael Krivelevich, *Topics in Random Graphs*, 2010, Internet Version.

2. ALAN FRIEZE and MICHAŁ KARONSKI, *Introduction to Random Graphs*, CMU Math. online version.

3. Noga Alon, Joel H. Spencer, The Probabilistic Method, 4th Edition, 2016, Internet Version.

4. P. Erdös, **Some remarks on the theory of graphs**, Bull. Amer. Math. Soc. **53** (1947), 292-294.

5. Theorem of the week, Theorem 25: Erdős's lower bound for the Ramsey numbers.

6. Bollobás, B., Catlin, P. A., & Erdös, P. (1980). Hadwiger's Conjecture is True for Almost Every Graph. European Journal of Combinatorics, 1(3), 195–199. (pdf version)

7. Robertson, N., Seymour, P. D., & Thomas, R. (1993). HADWIGER'S CONJECTURE FOR K6-FREE GRAPHS. Combinatorica, 13(3), 279–361.(pdf-version)

8. Ding, J., Lubetzky, E., & Peres, Y. (2014). Anatomy of the giant component: The strictly supercritical regime. European Journal of Combinatorics, 35, 155–168. (arXiv)

9. Zhao, Y., *Lecture Notes on Second Moment Method*, (online version).

10. Michael Krivelevich, *Exercises on Topics in Random Graphs*, ETH Zurich, (online teaching web).

11. Michael Krivelevich, *Hamiltonicity threshold in random graphs*, (online version)