Random Graph

Summer School 2016, Jiangsu Normal University

Yusheng Li

Tongji University

Contents

Chapter 1

Probabilistic Method and Random Graphs

Probabilistic method primarily used in combinatorics and pioneered by Paul Erdős, for proving the existence of a prescribed kind of mathematical object. This method has now been applied to other areas of mathematics such as number theory, algebra, analysis, and geometry, as well as in computer science (e.g. randomized rounding). When each element of a finite set Ω is assigned a non-negative weight, over which the sum is one, we have a probability space. The method works by showing that if one randomly chooses objects from a specified class, the probability that the result is of the prescribed kind is more than zero. The *basic probabilistic method* is by calculating the expected value of some random variable.

In basic probabilistic method, we only use expectation of random variable. In addition to basic probabilistic method, common tools used in the probabilistic method include Markov's inequality, the Chernoff bound, and the Lovász local lemma, etc.

1.1 Random graphs

The main reason that the probabilistic method becomes a main tool in Ramsey theory is that there are many Ramsey problems that traditional combinatorial methods do not work well. The standard texts on random graphs, we refer that of Alon and Spencer (2008), Bollobás (2001), and Jason, Łuczak, and Ruciński (2000). Random graphs began with some sporadic papers of Erdős in the 1940s and 1950s, in which he used random methods to show the existence of graphs with seemingly contradictory properties. A paper "On the Evolution of Random Graphs" of Erdős and Rényi in 1960 was very important for the development of the theory of random graphs. Among all these results, Erdős gave an exponential lower bound for the diagonal Ramsey number $r(n, n)$. Today random methods have been used in many other areas.

Every probability space whose points are graphs gives a notion of a random graph. For a family of graphs $\mathcal{G} = \{G_1, G_2, \ldots\}$ with probabilities $\Pr(G_i)$ such that $0 \leq \Pr(G_i) \leq 1$ and $\sum_{i\geq 1} \Pr(G_i) = 1$, we have a probability space of random graphs. Each G_i is called a *random graph* of $\mathcal{G}$ with probability $\Pr(G_i)$. We shall consider the probability space that consists of graphs on a fixed set $V = [n] = \{1, 2, \ldots, n\}$, where the vertices in V are *distinguishable*, so edges on V are distinguishable, too. Note that the complete graph K_n on vertex set V has

$$\binom{n}{1} + \binom{n}{2}2 + \cdots + \binom{n}{k}2^{\binom{k}{2}} + \cdots + \binom{n}{n}2^{\binom{n}{2}}$$

subgraphs. The general term corresponds the subgraphs that have exactly k vertices, and the last term $2^{\binom{n}{2}}$ corresponds all spanning subgraphs.

Let us label all edges of K_n on vertex set $V = [n]$ as $e_1, e_2, \ldots, e_m$, where $m = \binom{n}{2}$. Note that the number of graphs on vertex set $[n]$ is 2^m since the edges are distinguishable. The space $\mathcal{G}(n; p_1, \ldots, p_m)$ is defined for $0 \leq p_i \leq 1$ as follows. To get a random element of this space, one selects the edge e_i independently, with probability p_i. Putting it another way, the ground set of $\mathcal{G}(n; p_1, \ldots, p_m)$ is the set of all 2^m graphs on $V = [n]$. For a specific graph H in the space with $E(H) = \{e_j : j \in S\}$, where $S \subseteq \{1, \ldots, m\}$ is the index set of edges of H, the probability that H appears is

$$\Big(\Pi_{j\in S}\, p_j\Big)\Big(\Pi_{j\notin S}(1 - p_j)\Big).$$

That is to say, each of the edges of H has to be selected and none of $\overline{H}$ is allowed to be selected. Write $q_j = 1 - p_j$ and $G(p_1, \ldots, p_m)$ for a random element in $\mathcal{G}(n; p_1, \ldots, p_m)$. Then

$$\Pr(G(p_1, \ldots, p_m) = H)) = (\Pi_{j \in S} p_j)\,(\Pi_{j \notin S} q_j)\,.$$

Since the vertices (and edges) are distinguishable, the event $G(p_1, \ldots, p_n) = H$ is different from that $G(p_1, \ldots, p_m)$ is isomorphic to H. To see that $\mathcal{G}(n; p_1, \ldots, p_m)$ is truly a probability space, let us verify that

$$\begin{aligned}\sum_H \Pr\Big(G(p_1, \ldots, p_m) = H\Big) &= \sum_{S \subseteq [m]} (\Pi_{j \in S} p_j)\,(\Pi_{j \notin S} q_j) \\ &= \Pi_{j=1}^m (p_j + q_j) = 1.\end{aligned}$$

We shall concentrate on the case $p_1 = p_2 = \cdots = p_m = p$, for which the probability space $\mathcal{G}(n; p_1, \ldots, p_m)$ is written as $\mathcal{G}(n, p)$.

In space $\mathcal{G}(n, p)$ the probability of a specific graph H in the space with k edges is $p^k(1-p)^{m-k}$: Each of the k edges of H has to be selected and none of $\overline{H}$ is allowed to be selected. Write $G_{n,p}$ or simply G_p for a random element of $\mathcal{G}(n, p)$, then

$$\Pr(G_p = H) = p^{e(H)} q^{m-e(H)}.$$

In the space $\mathcal{G}(n, 0)$, the probability that the empty graph $\overline{K}_n$ appears is one, and the probability that any other graph appears is zero. Similarly, in the space $\mathcal{G}(n, 1)$, the only graph that appears is K_n. Other than these two extremal cases, for $0 < p < 1$, any graph on vertex set $[n]$ can appear with a positive probability. As p increases from 0 to 1, random graph G_p *evolves* from empty to full.

In their original paper on random graphs in 1960, Erdős and Rényi let $\mathcal{G}(n, e)$ be the random graph with vertex set $V = [n]$ and precisely e edges. For $0 \le e \le m = \binom{n}{2}$ with e fixed, the space $\mathcal{G}(n, e)$ consists of all $\binom{m}{e}$ spanning subgraphs with exactly e edges: which can be turned into a probability space by taking its elements to be equiprobable. Thus, write G_e for a random graph in the space $\mathcal{G}(n, e)$, for a specific graph H in the space, then

$$\Pr[G_e = H] = \binom{m}{e}^{-1},$$

where the event $G_e = H$ means that G_e is precisely H, but not only isomorphic to H in general.

It is interesting, as expected, that for $e \sim p\binom{n}{2}$ the spaces $\mathcal{G}(n, e)$ and $\mathcal{G}(n, p)$ are close to each other as $n \to \infty$. In most proofs for existence, the calculations are easier in $\mathcal{G}(n, p)$ than in $\mathcal{G}(n, e)$. So we will work on the probability model $\mathcal{G}(n, p)$ exclusively.

Another point of view may be convenient, in which one colors all edges of the complete graph K_n with probability p, randomly and independently. Thus random graph G_p is viewed as a random coloring of edge set of K_n. The coloring of edge set of K_n is also said a coloring of K_n in short. Recalling the definition of Ramsey numbers, we see why *the relation between random method and Ramsey theory is so natural and tight.*

In Ramsey theory, we need to consider the probability that a given subgraph appears. Let F be a given graph on k vertices, and let $S \subseteq [n]$ with $|S| = k$. Let A_S be the event that the subgraph induced by S contains F as a subgraph, then the event $\cup_S A_S$ signifies the event that F appears in G_p as a subgraph, its probability is hard to calculate since the events A_S have a complex interaction. It is often to bound this probability from above by the expectation of the number of copies of F in the random graph. To get the expectation, let us look the number of copies of F in K_k first. This is closely related to the automorphism group of F. A permutation (or a bijection) ϕ of $V(F)$ is called an *automorphism* of F when $uv \in E(F)$ if and only if $\phi(u)\phi(v) \in E(F)$ for any pair of vertices u and v. It is straightforward to verify the set of all automorphisms of F forms a group under the operation of composition. This group is called the *automorphism group* of F, denoted by $\mathcal{A}(F)$. For example, $\mathcal{A}(K_k)$ is the symmetric group S_k of order $k!$, and $\mathcal{A}(C_k)$ is the dihedral group D_k of order $2k$. The following results are simple consequence of the definitions.

For any graph F, $\mathcal{A}(F) = \mathcal{A}(\overline{F})$;

Let k be the order of graph F. Then $|\mathcal{A}(F)|$ is a divisor of $k!$, and it is $k!$ if and only if F is isomorphic to K_k or $\overline{K}_k$.

In random graph space, the vertices (and hence edges) of the graphs are labeled, and counting different copies of a subgraph is related the situation caused by different labeling of the vertices of the subgraph.

So we define that two graphs F_1 and F_2 to be *identical* if $V(F_1) = V(F_2)$ and $E(F_1) = E(F_2)$, where the equalities mean the same sets. Clearly, identical graphs are isomorphic, but the inverse statement is not generally true. For example, given a labeling of a star $K_{1,k}$ as $V = \{u, v_1, \ldots, v_k\}$ with the center u and $E = \{uv_1, \ldots, uv_k\}$, exchanging the labels of any pair vertices v_i and v_j will yield identical graphs.

Theorem 1.1 *Let F be a graph of order k. Then the number of copies of F from a set of k labels such that no two resulting graphs are identical is $k!/|\mathcal{A}(F)|$.*

Proof. Let $\{v_1, v_2, \ldots, v_k\}$ be the set of labels. Certainly there are $k!$ labeling of F from this set of labels with the number of resulting labeled graphs that may be identical. Let $F_1, F_2, \ldots, F_{k!}$ be the labeled graphs obtained from F. Then the relation "F_i is identical to F_j" is an equivalence relation. For a given labeled graph F_i, each automorphism gives rise to a labeled graph that is identical to F_i, and conversely. Hence each equivalence class so determined contains $|\mathcal{A}(F)|$ elements, thus implying that there are $k!/|\mathcal{A}(F)|$ equivalent classes in all. This proves the theorem. □

For example, if we label the vertices of a star $K_{1,3}$ as $1, 2, 3, 4$, then any equivalence class is uniquely determined by the label of its center. So there are 4 such classes, and each class contain 6 copies of $K_{1,3}$ with the same label of the center.

In a random graph space $\mathcal{G}(n, p)$, we need to consider the number of copies of F in a labeled complete graph.

Corollary 1.1 *Let F be a graph of order k. Then the number of copies of F in a labeled complete graph of order k is $k!/|\mathcal{A}(F)|$.*

Let F be a graph of order k. Let $S \subseteq [n]$ with $|S| = k$ and let X_S be the number of copies of F on S. Then $X = \sum_S X_S$ is the number of copies of F in G_p. We have

$$E(X_S) = \frac{k!}{|\mathcal{A}(F)|} p^{e(F)},$$

and

$$E(X) = \binom{n}{k}\frac{k!}{|\mathcal{A}(F)|}p^{e(F)} = \frac{(n)_k}{|\mathcal{A}(F)|}p^{e(F)},$$

where $(n)_k = n(n-1)\cdots(n-k+1)$ is the falling factorial.

Similar formulas hold for the number of *induced* subgraphs. Let Y be the number of induced F in G_p. Then

$$E(Y) = \frac{(n)_k}{|\mathcal{A}(F)|}p^{e(F)}q^{\binom{k}{2}-e(F)}.$$

Recall that A_S signifies the event that the subgraph induced by S in G_p contains F as a subgraph, then

$$\Pr(A_S) \leq \frac{k!}{|\mathcal{A}(F)|}p^{e(F)}.$$

Hence

$$\Pr(F \subset G_p) = \Pr(\cup A_S) \leq \binom{n}{k}\frac{k!}{|\mathcal{A}(F)|}p^{e(F)} = \frac{(n)_k}{|\mathcal{A}(F)|}p^{e(F)}, \qquad (1.1)$$

where the upper bound is exactly $E(X)$. This can be seen also by the fact that X takes only nonnegative integral values and

$$\begin{aligned}\Pr(\cup A_S) &= \Pr(X \geq 1) = \sum_{i\geq 1}\Pr(X=i)\\ &\leq \sum_{i\geq 1} i\Pr(X=i) = E(X).\end{aligned}$$

It seems to be necessary to point out that F is not a random element in $\mathcal{G}(n,p)$ and the above discussion is about appearance of F as a subgraph.

It is worth remarking that $p = p(n)$ is often a function. The space $\mathcal{G}(n,p)$ is of great interest for fixed values of p as well; in particular, $\mathcal{G}(n,1/2)$ could be viewed as the space: it consists of all 2^m graphs on $V = [n]$, and the probability of any graph is equiprobable. This is just a classical probability space. Thus $G_{n,1/2}$ is also obtained by picking any of the 2^m graphs on $V = [n]$ at random with probability 2^{-m}. No matter how p is fixed or a function, we tend to be interested in what happens as $n \to \infty$.

Now we have obtained a space of random graphs, every graph invariant becomes a random variable; the nature of such a random variable depends crucially on p. For instant, the number $X_k(G)$ of complete graphs of order k in G is a random variable on our space of random graphs.

To be proficient in the probabilistic method one must have a feeling for asymptotic calculation. For the sake of convenience, we state some simple inequalities that will be used in the calculations. The following precise formula is called *Stirling formula.*

Lemma 1.1 *For all $n \geq 1$,*

$$n! = \sqrt{2\pi n}\Big(\frac{n}{e}\Big)^n e^{\theta/(12n)},$$

where $0 < \theta = \theta_n < 1$. Thus

$$\sqrt{2\pi n}\left(\frac{n}{e}\right)^n < n! < \sqrt{2\pi n}\left(\frac{n}{e}\right)^n e^{1/(12n)}.$$

Lemma 1.2 *For any positive integers $N \geq n$,*

$$\left(\frac{N}{n}\right)^n \leq \binom{N}{n} \leq \left(\frac{eN}{n}\right)^n.$$

If $n = o(\sqrt{N})$ as $n \to \infty$, then $\binom{N}{n} \sim \frac{N^n}{n!}$.

Proof. The first inequality comes from $\frac{N}{n} \leq \frac{N-i}{n-i}$, and the second does from $\binom{N}{n} \leq \frac{N^n}{n!}$ and Stirling's formula, and then it suffices to see that $(N)_n/N^n$ is equal to

$$\exp\Big[\sum_{i=1}^{n-1} \log\Big(1 - \frac{i}{N}\Big)\Big] = \exp\Big[-\sum_{i=1}^{n-1}\frac{i}{N} - O\Big(\frac{n^2}{N}\Big)\Big] \to 1,$$

and the desired asymptotical formula follows. □

The following simple fact from calculus is often used.

Lemma 1.3 *For any $0 \leq x \leq 1$ and $n \geq 0$,*

$$(1-x)^n \leq e^{-nx}.$$

If $x = x(n) \to 0$ and $x^2 n \to 0$ as $n \to \infty$, then

$$(1-x)^n \sim e^{-nx}.$$

We have obtained an upper bound of $r(n,n)$ in Chapter 1. This and Stirling's formula give

$$r(n,n) \le \binom{2n-2}{n-1} = (1+o(1))\frac{4^{n-1}}{\sqrt{\pi n}}.$$

Thomason (1988) slightly improved the upper bound as $\frac{\exp(c\sqrt{\log n})}{\sqrt{n}}\binom{2n-2}{n-1}$, where $\exp(c\sqrt{\log n}) = o(n^\epsilon)$ for any $\epsilon > 0$. Recently, Conlon improved the upper bound as $n^{-c\log n/\log\log n}\binom{2n-2}{n-1}$. However, this does not change the following limit as

$$\overline{\lim_{n\to\infty}} r(n,n)^{1/n} \le 4.$$

We do not have any general lower bound for $r(n,n)$ yet.

1.2 Elementary examples

This section is devoted to the methodology, in which we use basic probabilistic method to estimate $r(m,n)$. All lower bounds in this section have been improved. However, these bounds are from "almost all" argument, that is to say, the probability of graphs in the corresponding space offering such bounds tends to 1. This method is often associated with the *basic method*, in which we use expectation of random variables. These graphs imply these bounds directly or indirectly (after some vertices deleted).

In the original proof of the exponent lower bound for $r(n,n)$ in 1947, Erdős did not use the formal probabilistic language. So his paper has been considered as an informal starting point of random graphs. But in two papers published in 1959 and 1961, in which he gave a lower bound $c(n/\log n)^2$ for $r(3,n)$, he even wrote probabilities in the titles.

Theorem 1.2 *For* $n \ge 1$,

$$r(n,n) > \frac{n}{e\sqrt{2}}2^{n/2}.$$

Proof. Consider the random graphs in $\mathcal{G}(N, 1/2)$, or color K_N randomly and independently with probability $p = 1/2$, where N is a positive integer to be chosen. Let S be a set of n vertices and let A_S be the event that S is monochromatic. Then

$$\Pr[A_S] = 2\left(\frac{1}{2}\right)^{\binom{n}{2}} = 2^{1-\binom{n}{2}},$$

as for S to hold all $\binom{n}{2}$ edges must be colored the same. Consider the event $\bigcup A_S$ over all n-sets on $[N]$. We use the simple fact that the probability of a disjunction is at most the sum of the probability of the events. Thus

$$\Pr[\cup A_S] \leq \sum \Pr[A_S] = \binom{N}{n} 2^{1-\binom{n}{2}}.$$

If this probability is less than one, then the event $B = \bigcap \overline{A}_S$ has positive probability. Therefore B is not the null event. Thus there is a point in the probability space for which B holds. But a point in the probability space is precisely a coloring of the edges of K_N. And the event B is precisely that under this coloring there is no monochromatic K_n. Hence $r(n, n) > N$.

We need to find the maximum possible N such that $\Pr[\cup A_S] < 1$. From Stirling formula, we have

$$\binom{N}{n} 2^{1-\binom{n}{2}} \leq \frac{N^n}{n!} 2^{1-\binom{n}{2}} < \frac{2}{\sqrt{2\pi n}} \Big(\frac{e\sqrt{2}N}{n2^{n/2}}\Big)^n.$$

This can be ensured by setting $N = \left\lfloor \frac{n}{e\sqrt{2}} 2^{n/2} \right\rfloor$ such that the fraction in the parenthesis is at most one. Therefore $r(n, n) \geq N + 1 > \frac{n}{e\sqrt{2}} 2^{n/2}$. □

The original proof of Erdős used the counting argument as follows. Fix a set $S \subseteq [N]$ with $|S| = n$. Among all graphs on $[N]$, the proportion of the graphs in which S is a clique or an independent set is $2^{1-\binom{n}{2}}$, which can be seen as each edge in S has tow possibilities: to appear or not. Since the number of sets S is $\binom{N}{n}$, if

$$\binom{N}{n} 2^{1-\binom{n}{2}} < 1,$$

then there is a graph of order N which contains no K_n or independent set of order n, and thus $r(n,n) > N$.

Erdős in fact used the space $\mathcal{G}(N, 1/2)$, which is a classical probability space as mentioned. It is interesting to see that this space is the only one that counting argument works!

Theorem 1.3 *Let m, n and N be positive integers. If for some $0 < p < 1$,*

$$\binom{N}{m} p^{\binom{m}{2}} + \binom{N}{n} (1-p)^{\binom{n}{2}} < 1,$$

then $r(m,n) > N$. Hence $r(m,n) \geq c\left(\frac{n}{\log n}\right)^{(m-1)/2}$, where $c = c(m) > 0$ is a constant.

Proof. Consider random graphs G_p in $\mathcal{G}(N,p)$. Let S be a set of m vertices and let A_S be the event that S induces a complete graph. Then $\Pr[A_S] = p^{\binom{m}{2}}$, and

$$\Pr[\cup A_S] \leq \sum \Pr[A_S] = \binom{N}{m} p^{\binom{m}{2}}.$$

Let T be a set of n vertices and let B_T be the event that T induces an independent set. Then

$$\Pr[\cup B_T] \leq \sum \Pr[B_T] = \binom{N}{n} (1-p)^{\binom{n}{2}}.$$

Thus

$$\Pr[(\cup A_S) \cup (\cup B_T)] < 1.$$

So there exists a graph on N vertices such that there is neither an induced K_m nor an induced $\overline{K}_n$, thus $r(m,n) > N$.

The above result is ineffective in bounding $r(3,n)$. We now examine the lower bound of $r(4,n)$. We shall give details in calculation for choosing a suitable value of p, and that of N as large as possible for large n. To make the condition in Theorem 1.3 satisfied, we roughly estimate $\binom{N}{n}$ as $(eN/n)^n$, and $(1-p)^{\binom{n}{2}}$ as $e^{-p\binom{n}{2}}$, hence $\binom{N}{n}(1-p)^{\binom{n}{2}}$ as

$$\left(\frac{eN}{n}\right)^n \exp\left\{-p\binom{n}{2}\right\} = \left(\frac{eN}{ne^{p(n-1)/2}}\right)^n.$$

We have known that $r(4,n) \le (1+o(1))n^3/(\log n)^2$ in the last chapter, thus $e^{p(n-1)/2} = n^{a+o(1)}$ for some constant a, so we take $p = (c_1 \log n)/(n-1)$. Then

$$\binom{N}{4}p^6 \sim \frac{1}{24}N^4\Big(\frac{c_1 \log n}{n}\Big)^6 \sim \frac{c_1^6}{24}\Big(N\Big(\frac{\log n}{n}\Big)^{3/2}\Big)^4 < 1,$$

so $N \sim c_2(n/\log n)^{3/2}$ for some constant c_2.

Formally, we let $p = (c_1 \log n)/(n-1)$ and $N = \lfloor c_2(n/\log n)^{3/2} \rfloor$, where c_1 and c_2 are positive constants to be chosen satisfying that $c_1^6 c_2^4 < 24$. Then

$$\binom{N}{4}p^6 < \frac{N^4}{24}p^6 \le \frac{c_1^6 c_2^4}{24}\left(\frac{n}{n-1}\right)^6 \le c_3 < 1$$

for large n, where c_3 is a constant. For the second term, we estimate that $(1-p)^{\binom{n}{2}} < e^{-pn(n-1)/2} = n^{-c_1 n/2}$ and

$$\binom{N}{n}(1-p)^{\binom{n}{2}} < \left(\frac{eN}{n}\right)^n n^{-c_1 n/2} = \left(\frac{eN}{n^{1+c_1/2}}\right)^n.$$

In order to make the above tending to zero, we have to take $c_1 \ge 1$. On the other hand, in order to take c_2 as large as possible with $c_1^6 c_2^4 < 24$, we have to take c_1 as small as possible. So it has to be $c_1 = 1$.

Now, we may hope to optimize the constant c_2. Since we need only $c_2 < 24^{1/4}$, so $c_2 = 24^{1/4} - \epsilon$ will be ok. Thus we have

$$r(4,n) \ge (24^{1/4} - o(1))\left(\frac{n}{\log n}\right)^{3/2}.$$

For general $m \ge 4$, by taking $p = (m-3)\log n/(n-1)$, the similar calculation yields

$$r(m,n) \ge c\left(\frac{n}{\log n}\right)^{(m-1)/2}.$$

□

Hereafter we will choose p with some foresight. It is often to replace $\log n/(n-1)$ in the expression of p by $\log n/n$. Strictly speaking, only if $f(n)-g(n) \to 0$, we have $e^{f(n)} \sim e^{g(n)}$. But the replacement works as we use $p = \log n/n$ in above calculation for the lower bound of $r(4,n)$,

$$\binom{N}{n}(1-p)^{\binom{n}{2}} < \left(\frac{eN}{n}\right)^n n^{-(n-1)/2} = \left(\frac{eN}{n^{3/2}}\right)^n n^{1/2} \to 0.$$

Also, when $N = \lfloor f(n) \rfloor$ or $N = \lceil f(n) \rceil$ and $f(n) \to \infty$, we often ignore the fact that $f(n)$ may not be an integer by simply taking $N = f(n)$ if this does not affect the proof essentially.

We have seen that the property of G_p is sensitive with the value of p. To ensure that G_p contains no K_m (with a positive probability, more precisely, or $\binom{N}{m}p^{\binom{m}{2}}$ is small), it is better to take smaller p. But it is better to take a bigger p to ensure that there is no induced $\overline{K}_n$ (i.e, $\binom{N}{n}(1-p)^{\binom{n}{2}}$ is small). Our task is to balance both sides to obtain a larger N as possible.

We shall improve the obtained lower bounds for $r(n,n)$ and $r(m,n)$ by the proofs so called *deletion method.*

Theorem 1.4 *As $n \to \infty$,*

$$r(n,n) \geq (1-o(1))\frac{n}{e}2^{n/2}.$$

Proof. Consider the random graphs in $\mathcal{G}(N, 1/2)$. Let X be the number of clique or independent set of size n. Then

$$X = \sum X_S,$$

the sum over all $n-$set S, where X_S is the indicator random variable of the event A_S that S is a clique or independent. That is

$$X_S = \begin{cases} 1 & \text{if } S \text{ induces } K_n \text{ or } \overline{K_n} \\ 0 & \text{otherwise.} \end{cases}$$

Therefore

$$E[X_S] = \Pr[A_S] = 2\left(\frac{1}{2}\right)^{\binom{n}{2}}.$$

By linearity of expectation

$$E[X] = \sum E[X_S] = \binom{N}{n} 2^{1-\binom{n}{2}}.$$

There is a point in the probability space for which X does not exceed its expectation. That is, there is a graph with at most $E(X)$ sets S that induce a K_n or a $\bar{K}_n$. Fix that graph. For each such S select a point $x \in S$ and delete it from the vertex set. The remaining vertex set V^* have neither K_n nor $\bar{K}_n$. Thus

$$r(n,n) > |V^*| \geq N - E(X).$$

The rest of the proof is to find N such that $|V^*|$ as large as possible. By taking $N = \lfloor \frac{n2^{n/2}}{e} \rfloor$, from the Stirling formula, we have

$$\binom{N}{n} 2^{1-\binom{n}{2}} < \left(\frac{eN}{n}\right)^n 2^{1-\binom{n}{2}} < 2\left(\frac{e\sqrt{2}N}{n2^{n/2}}\right)^n \leq 2^{n/2+1},$$

which is $o(N)$. Thus $r(n,n) \geq (1-o(1))N$. □

Theorem 1.5 *For any positive integer m, n and N, and any real number $0 < p < 1$,*

$$r(m,n) > N - \binom{N}{m} p^{\binom{m}{2}} - \binom{N}{n} (1-p)^{\binom{n}{2}}.$$

Thus for fixed $m \geq 3$, if n is large, then

$$r(m,n) \geq c\left(\frac{n}{\log n}\right)^{m/2},$$

where $c = c(m) > 0$ is a constant.

Proof. Note that in G_p in $\mathcal{G}(N,p)$ the expectation of the number of clique of size m is $\binom{N}{m} p^{\binom{m}{2}}$, and expectation of number of independent sets of size n is $\binom{N}{n}(1-p)^{\binom{n}{2}}$. By the same argument in the proof for Theorem 1.4 gives the first lower bound. For the second, we set $N =$

$a(n/\log n)^{m/2}$ and $p = (m-2)\log n/(n-1)$ such that $a - \frac{(m-2)^{\binom{m}{2}}}{m!}a^m >$ 0. This is possible since $a^m = o(a)$ as $a \to 0^+$. Then

$$N_1 = \binom{N}{m} p^{\binom{m}{2}} \sim \frac{(m-2)^{\binom{m}{2}} a^m}{m!} \left(\frac{n}{\log n}\right)^{m/2},$$

and

$$N_2 = \binom{N}{n} (1-p)^{\binom{n}{2}} < \left(\frac{eN}{n}\right)^n e^{-pn(n-1)/2} = \left(\frac{eN}{n^{m/2}}\right)^n \to 0.$$

So if $c < a - \frac{(m-2)^{\binom{m}{2}} a^m}{m!}$, then

$$r(m,n) \geq N - N_1 - N_2 > c \left(\frac{n}{\log n}\right)^{m/2},$$

completing the proof. □

Let $f(n)$ and $g(n)$ be positive functions. Clearly, in order to show $f(n) \geq cg(n)$ for all positive n, we may assume that n is sufficiently large in the proof. Theorem 1.5 can be generalized as follows.

Theorem 1.6 *Let F be a graph with $m \geq 3$ vertices and $e(F) \geq m$ edges. Then*

$$r(F, K_n) \geq c \left(\frac{n}{\log n}\right)^{e(F)/(m-1)},$$

for all $n \geq 1$, where $c = c(F) > 0$ is a constant.

The above theorem and the linear formula of $r(T_m, K_n)$ in Chapter 2 give the following conclusion.

Corollary 1.2 *For a fixed connected graph G, $r(G, K_n)$ can have a linear upper bound on n if and only if G is tree.*

In Theorem 1.2, we proved that if $N = \frac{n}{e\sqrt{2}} 2^{n/2}$, then almost all graphs in $\mathcal{G}(N, 1/2)$ satisfy $\omega(G_p) < n$ and $\alpha(G_p) < n$. Can we construct one (in fact, one family of graphs)? Unfortunately, nobody can give a constructive proof of a lower bound like $r(n,n) \geq a^n$ for any constant $a > 1$. This reveals that finite discrete structures are much more complicated than they seem to be. The graphs of large order we can image must have some uniform description, but typical random graphs do not have such description. This is not very unusual in mathematics. Most real numbers that we can write down are rational, which are recycling, but almost all real numbers are irrational.

Let us have a brief remark on "almost all" arguments in the proofs of several theorems in this section. They are in fact *weighted counting arguments*. The probability space we are dealing is finite, which can be defined as a finite set Ω together with nonnegative weights on the elements that sum to 1. An event A is a subset of Ω. The probability $\Pr(A)$ is the sum of the weights of the elements of A. In particular, in the space $\mathcal{G}(N,p)$, the graphs with around $p\binom{N}{2}$ edges have larger weights. For example, we can say that almost all graphs in $\mathcal{G}(N, 1/N^3)$ are empty as the probability of any non-empty graph is negligible.

We have seen that

$$\sqrt{2} \leq \underline{\lim} r(n,n)^{1/n} \leq \overline{\lim} r(n,n)^{1/n} \leq 4.$$

There is a big gap between the two bounds. It is likely that in fact, the limit of $r(n,n)^{1/n}$ is 2, but both the lower bound $\sqrt{2}$ and the upper bound 4 has been standing for more than half century without improvement.

Problem 1.1 *Prove or disprove that* $\lim_n r(n,n)^{1/n}$ *exists. Is the limit* 2 *if it exists?*

1.3 References

N. Alon and J. Spencer, *The Probabilistic Method, 3rd ed.*, Wiley-Interscience, New York, 2008.

B. Bollobás, *Random Graphs, 2nd Edition*, Cambridge University Press, London-New York, 2001.

D. Conlon, A new upper bound for diagonal Ramsey numbers, *Ann. of Math.*, **170** (2009), 941-960.

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

P. Erdős, Graph theory and probability, *Canad. J. Math.*, **11** (1959), 34-38.

P. Erdős, Graph theory and probability II, *Canad. J. Math*, **13** (1961), 346-352.

P. Erdős and A. Rényi, On the evolution of random graphs, *Publ. Math. Inst. Hungar. Acad. Sci.*, **5** (1960), 17-61.

S. Jason, T. Łuczak, and A. Ruciński, *Random Graphs*, Wiley-Interscience, New York, 2000.

M. Molloy and B. Reed, *Graph Coloring and the Probabilistic Method*, Springer, Berlin, 2002.

Chapter 2

Concentration

The probability space we consider in graph Ramsey theory has only finite many possible outcomes, and the random variable is often nonnegative. Markov's inequality gives an upper bound for the probability that such a random variable is greater than or equal to some positive constant. It is named after the Russian mathematician Andrey Markov, although it appeared earlier in the work of Pafnuty Chebyshev (Markov's teacher). Chebyshev's inequality guarantees that in any data sample or probability distribution, "nearly all" values are close to the mean. The inequalities have great utility because it can be applied to completely arbitrary distributions (unknown except for mean and variance). In probability theory, the Chernoff bound, named after Herman Chernoff, gives exponentially decreasing bounds on tail distributions of sums of independent random variables. It is a sharper bound than the known first or second moment based tail bounds such as Markov's inequality or Chebyshev inequality, which only yield power-law bounds on tail decay but it requires that the variates be independent - a condition that neither the Markov nor the Chebyshev's inequalities require This chapter contains various Chernoff's inequalities with the detailed proofs, and some of their applications, particularly in Ramsey theory. The last two sections are on martingales and concentration of parameters of dense random graphs, which the beginners of readers can skip.

2.1 The Chernoff's inequality

Let X be a discrete random variable. Then the expected value of X is defined to be $E(X) = \sum_i a_i \Pr(X = a_i)$, where the summation is taken over all values a_i that X can take.

Theorem 2.1 (Markov's Inequality) *Let $a > 0$ and let X be a nonnegative random variable. Then*

$$\Pr(X \geq a) \leq \frac{E(X)}{a}.$$

Proof. Suppose that $\{a_i\}$ is the set of all values that X takes. Then

$$\begin{aligned} E(X) &= \sum_i a_i \Pr(X = a_i) \geq \sum_{a_i \geq a} a_i \Pr(X = a_i) \\ &\geq a \sum_{a_i \geq a} \Pr(X = a_i) = a \Pr(X \geq a), \end{aligned}$$

as required. □

The following is exactly what we used to obtain lower bounds of Ramsey numbers in the last chapter.

Corollary 2.1 *If a random variable X only takes nonnegative integer values and $E(X) < 1$, then $\Pr(X \geq 1) < 1$ hence $\Pr(X = 0) > 0$.*

For a positive integer k, the *kth moment* of a real-valued random variable X is defined to be $E(X^k)$, and so the first moment is simply the expected value. Denote by $\mu = E(X)$, and define the *variance* of X as $E((X - \mu)^2)$, which is denoted by σ^2. Call

$$\sigma = \sqrt{E((X - \mu)^2)}$$

as the *standard deviation* of X. A basic equality is as follows.

$$\sigma^2 = E(X^2) - \mu^2.$$

Theorem 2.2 (Chebyshev's Inequality) *Let X be a random variable and let a be a positive number. Then*

$$\Pr(|X - \mu| \geq a) \leq \frac{\sigma^2}{a^2}.$$

Proof. By Markov's inequality, for any $a > 0$,

$$\begin{aligned}\sigma^2 &= E((X-\mu)^2)) \geq a^2 \Pr((X-\mu)^2 \geq a^2) \\ &= a^2 \Pr(|X-\mu| \geq a).\end{aligned}$$

It follows by the required statement. □

In importance, the second moment $E(X^2)$ is second to the first moment $E(X)$.

Lemma 2.1 (Second Moment Method) *If X is a random variable, then*

$$\Pr(X=0) \leq \frac{\sigma^2}{\mu^2} = \frac{E(X^2)-\mu^2}{\mu^2},$$

where $\mu = E(X)$. In particular, $\Pr(X=0) \to 0$ *if* $E(X^2)/\mu^2 \to 1$.

The proof follows from Chebyshev's Inequality and the trivial fact that $\Pr(X=0) \leq \Pr(|X-\mu| \geq \mu)$ immediately. Intuitively, if σ grows more slowly than μ grows, than $\Pr(X=0) \to 0$ since σ "pulls" X close to μ thus far away from zero.

The Chebshev's inequality is in fact the Markov's inequality on random variable $|X-\mu|$. However, Chebshev's inequality states the probability of a random variable X apart from $E(X)$ is bounded. When this is the case, we say that X is *concentrated.* A concentration bounds is used to show that a random variable is very close to its expected value with high probability, so it behaves approximately as one may "expect" it to be. When S_n is the sum of n independent variables, each variable equals to 1 with probability p and -1 with probability $1-p$, respectively, the bound can be sharper. Such random variables are bounded in Chernoff's inequality. Most of the results in this chapter may be found in, or immediately derived from, the seminal paper of Chernoff (1952) while our proofs are self-contained. A set of random variables $X_1, X_2, \ldots$ are said to be mutually independent means each X_i is independent of any Boolean expression formed from other $(X_j)'s$. In any form of Chernoff bounds, we have an assumption as follows.

Assumption A : On the independence of variables in *Chernoff bound* or *Chernoff inequality.* Let $X_1, X_2, \ldots$ be mutually independent variables and they have the same binomial distribution. Set

$$S_n = \sum_{i=1}^{n} X_i.$$

All concentration bounds in the remaining part of this section are Chernoff bounds of different forms, which estimate the probability of

$$\Pr(S_n \geq n(\mu + \delta)) = \Pr\Big(\frac{S_n}{n} \geq \mu + \delta\Big),$$

where $\mu = E(X_i)$. The symmetric bound on $\Pr(S_n \leq n(\mu - \delta))$ can be obtained similarly.

Theorem 2.3 *Under Assumption A, suppose*

$$\Pr(X_i = 1) = \Pr(X_i = -1) = \frac{1}{2}$$

for $i = 1, 2, \ldots$. *Then for any* $\delta > 0$,

$$\Pr(S_n \geq n\delta) < \exp\{-n\delta^2/2\},$$

and for any $a > 0$,

$$\Pr(S_n \geq a) < \exp\{-a^2/(2n)\}.$$

Proof. Let $\lambda > 0$ be arbitrary. Then

$$E(e^{\lambda X_i}) = \frac{e^{\lambda} + e^{-\lambda}}{2}.$$

Note that

$$\begin{aligned} E(e^{\lambda S_n}) &= E(e^{\lambda X_1})E(e^{\lambda X_2}) \ldots E(e^{\lambda X_n}) = \left(\frac{e^{\lambda} + e^{-\lambda}}{2}\right)^n \\ &= \left(\sum_{j=0}^{\infty} \frac{\lambda^{2j}}{(2j)!}\right)^n < \left(\sum_{j=0}^{\infty} \frac{1}{j!}\left(\frac{\lambda^2}{2}\right)^j\right)^n = e^{n\lambda^2/2}, \end{aligned}$$

where we use the fact that $(2j)! \geq 2^j j!$ for all $j \geq 0$ with strict inequality when $j \geq 2$. Now by Markov's inequality,

$$\begin{aligned}\Pr(S_n \geq n\delta) &= \Pr(e^{\lambda S_n} \geq e^{\lambda n\delta}) \\ &\leq \frac{E(e^{\lambda S_n})}{e^{\lambda n\delta}} < \exp\{n(\lambda^2/2 - \lambda\delta)\},\end{aligned}$$

for all $\lambda > 0$. Setting $\lambda = \delta$, we obtain the desired result. □

For large n, the central limit theorem implies that S_n is approximately normal with zero mean and standard deviation $\sqrt{n}$. For any fixed u,

$$\lim_{n\to\infty} \Pr(S_n \geq u\sqrt{n}) \int_u^\infty \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt < e^{-u^2/2}.$$

However, the Chernoff bound holds for all positive n and a.

Since X_i is often an indicator variable of some random event, so X_i takes 1 when the event appears and 0 otherwise. The following form of Chernoff bound may be used in more cases.

Theorem 2.4 *Under Assumption A, suppose*

$$\Pr(X_i = 1) = \Pr(X_i = 0) = \frac{1}{2}$$

for $i = 1, 2, \ldots$. *Then for any* $\delta > 0$,

$$\Pr(S_n \geq n(1+\delta)/2) < \exp\{-n\delta^2/2\}.$$

Namely

$$\Pr(S_n \geq n(1/2+\delta) < \exp\{-2n\delta^2\}.$$

Proof. Set $Y_i = 2X_i - 1$ and $T_n = \sum_{i=1}^n Y_i = 2S_n - n$. Then

$$\Pr(Y_i = 1) = \Pr(Y_i = -1) = \frac{1}{2},$$

and $\{Y_i\}$ satisfies Assumption A. Note that $T_n \geq n\delta$ if and only if $S_n \geq n(1+\delta)/2$. Applying Theorem 2.3 to $\{Y_i\}$ and T_n, we have

$$\Pr(S_n \geq n(1+\delta)/2) = \Pr(T_n \geq n\delta) < \exp\{-n\delta^2/2\}$$

as claimed. □

Under Assumption A, suppose

$$\Pr(X_i = 1) = p, \quad \text{and} \quad \Pr(X_i = 0) = 1 - p$$

for $i = 1, 2, \ldots$. Then we say that the sum $S_n = \sum_{i=1}^n X_i$ has binomial distribution, denoted by $B(n, p)$. Involved in Theorem 2.4 is special binomial distribution $B(n, 1/2)$. For general case , the calculation is slightly more complicated, but the technique is the same. As usual, denote by q for $1 - p$.

Theorem 2.5 *Under Assumption A, suppose*

$$\Pr(X_i = 1) = p \quad \textit{and} \quad \Pr(X_i = 0) = q$$

for $i = 1, 2, \ldots$. *Then there exists* $\delta_0 = \delta_0(p) > 0$ *so that if* $0 < \delta < \delta_0$, *then*

$$\Pr(S_n \geq n(p + \delta)) < \exp\{-n\delta^2/(3pq)\}.$$

Proof. Let $a = p + \delta$. By the same argument as used before,

$$\begin{aligned}
\Pr(S_n \geq na) &= \Pr(e^{\lambda S_n} \geq e^{\lambda na}) \leq \frac{1}{e^{\lambda na}} E(e^{\lambda S_n}) \\
&= \frac{1}{e^{\lambda na}}(pe^{\lambda} + q)^n = (pe^{\lambda(1-a)} + qe^{-\lambda a})^n
\end{aligned}$$

for all $\lambda > 0$. Let $c = 1 - a = q - \delta > 0$, then $a + c = 1$. By taking $\lambda = \log(aq/cp)$, we have

$$\min_{\lambda > 0}(pe^{\lambda c} + qe^{-\lambda a}) = e^{-\lambda a}(pe^{\lambda} + q) = \left(\frac{cp}{aq}\right)^a \frac{q}{c} = \left(\frac{p}{a}\right)^a \left(\frac{q}{c}\right)^c.$$

Setting $0 < \delta < 1 - p$, and expanding in powers of δ, with the fact that

$$\log(1 + x) = x - \frac{x^2}{2} + \frac{x^3}{3} + O(x^4),$$

we find

$$\begin{aligned}
\log\left(\frac{p}{a}\right)^a &= (p + \delta)\log\left(1 - \frac{\delta}{p + \delta}\right) \\
&= -\delta - \frac{\delta^2}{2(p + \delta)} - \frac{\delta^3}{3(p + \delta)^2} + o(\delta^3),
\end{aligned}$$

and

$$\begin{aligned}\log\left(\frac{q}{c}\right)^c &= (q-\delta)\log\left(1+\frac{\delta}{q-\delta}\right)\\ &= \delta-\frac{\delta^2}{2(q-\delta)}+\frac{\delta^3}{3(q-\delta)^2}+o(\delta^3).\end{aligned}$$

Adding them by terms, the first sum vanishes, and the second is

$$\begin{aligned}\frac{-\delta^2}{2}\left(\frac{1}{p+\delta}+\frac{1}{q-\delta}\right) &= \frac{-\delta^2}{2}\left(\frac{1}{p(1+\delta/p)}+\frac{1}{q(1-\delta/q)}\right)\\ &= \frac{-\delta^2}{2}\left(\frac{1}{pq}-\frac{(q^2-p^2)\delta}{p^2q^2}+o(\delta)\right)\\ &= \frac{-\delta^2}{2pq}+\frac{(q-p)\delta^3}{2p^2q^2}+o(\delta^3),\end{aligned}$$

and the third is

$$\begin{aligned}\frac{\delta^3}{3}\left(\frac{1}{(q-\delta)^2}-\frac{1}{(p+\delta)^2}\right) &= \frac{\delta^3}{3}\left(\frac{1}{q^2}-\frac{1}{p^2}+o(1)\right)\\ &= \frac{-(q-p)\delta^3}{3p^2q^2}+o(\delta^3).\end{aligned}$$

We have for small $\delta > 0$

$$\log\left[\left(\frac{p}{a}\right)^a\left(\frac{q}{c}\right)^c\right]=\frac{-\delta^2}{2pq}+\frac{(q-p)\delta^3}{6p^2q^2}+o(\delta^3)<\frac{-\delta^2}{3pq}.$$

Thus

$$\Pr(S_n \geq n(p+\delta)) < \exp\{-n\delta^2/(3pq)\},$$

completing the proof. □

From above proof for $p > q$ and Theorem 2.4 for $p = q = 1/2$, we see that if $p \geq 1/2$, the bound can be slightly better as

$$\Pr(S_n > n(p+\delta)) < \exp\{-n\delta^2/(2pq)\}.$$

We now write out a symmetric form for Theorem 2.5, and omit those for Theorem 2.3 and Theorem 2.4.

Theorem 2.6 *Under Assumption A, suppose*

$$\Pr(X_i = 1) = p \quad \textit{and} \quad \Pr(X_i = 0) = q$$

for $i = 1, 2, \ldots$. *Then there exists* $\delta_0 = \delta_0(p) > 0$ *so that if* $0 < \delta < \delta_0$, *then*

$$\Pr(S_n \leq n(p - \delta)) < \exp\{-n\delta^2/(3pq)\}.$$

Therefore

$$\Pr(|S_n - np| > n\delta)) < 2\exp\{-n\delta^2/(3pq)\}.$$

□

From the above proof, we have

$$\begin{aligned} \Pr(S_n \geq na) &\leq \left(\left(\frac{p}{a}\right)^a \left(\frac{q}{c}\right)^c\right)^n \\ &= \exp\left\{n\left(a\log\frac{p}{a} + (1-a)\log\frac{q}{1-a}\right)\right\}, \end{aligned}$$

where $a = p + \delta$ and $c = 1 - a$. Set $k = na$, then $k > np$ and

$$\Pr(S_n \geq k) \leq \exp\left\{n\left((k/n)\log\frac{p}{k/n} + (1-k/n)\log\frac{q}{1-k/n}\right)\right\}.$$

Let $H(x)$ signify the entropy function

$$H(x) = x\log\frac{p}{x} + (1-x)\log\frac{q}{1-x}, \quad 0 < x < 1,$$

then

$$\Pr(S_n \geq k) \leq \exp\{nH(k/n)\},$$

which is valid also for $k = np$ since $H(p) = 0$. The following form of Chernoff's inequality was used by Beck (1983).

Theorem 2.7 *Under Assumption A, suppose*

$$\Pr(X_i = 1) = p \quad \textit{and} \quad \Pr(X_i = 0) = q$$

for $i = 1, 2, \ldots$. *If* $k \geq np$, *then*

$$\Pr(S_n \geq k) \leq \left(\frac{np}{k}\right)^k \left(\frac{nq}{n-k}\right)^{n-k}.$$

Consequently,

$$\Pr(S_n \geq k) \leq \left(\frac{npe}{k}\right)^k.$$

Proof. The right hand side of the first inequality is just $\exp\{nH(k/n)\}$. For the second inequality, simply note that

$$\left(\frac{nq}{n-k}\right)^{n-k} \leq \left(\frac{n}{n-k}\right)^{n-k} = \left(1 + \frac{k}{n-k}\right)^{n-k} < e^k.$$

Thus the required result follows. □

2.2 Applications of Chernoff's bounds

Let us first see that a. a. graphs are nearly regular.

Theorem 2.8 *Let* $0 < p < 1$ *and* $\epsilon > 0$ *be fixed. Then almost all graphs* G *in* $\mathcal{G}(n, p)$ *satisfy*

$$|\deg(v) - (n-1)p| \leq \epsilon(n-1)p$$

for each vertex v.

Proof. Let G be a random graph in $\mathcal{G}(n, p)$ and let v ba a fixed vertex of G. Then $\deg(v)$ has binomial distribution $B(n-1, p)$. From Chernoff's Theorems, we have

$$\begin{aligned} \Pr(|\deg(v) - (n-1)p| > \epsilon(n-1)p) &< 2\exp(-(n-1)\epsilon^2/(3pq)) \\ &\sim 2\exp(-n\epsilon^2/(3pq)). \end{aligned}$$

Hence we bound the probability that there is at least one vertex v such that $|\deg(v) - (n-1)p| > (n-1)p)$ by $(2 + o(1))n\exp(-n\epsilon/(3pq))$, which tends to zero as $n \to \infty$. □

The condition that fixed p can be weakened as $p = (\log n/n)\omega(n)$ with $\omega(n) \to \infty$, see Alon and Spencer (2008).

Let us enjoy an application of Chernoff bound that is of Erdős style, in which the authors disproved a conjecture with almost all graphs.

A suspended path in graph G is a path $(x_0, x_1, \ldots, x_k)$ in which $x_1, \ldots, x_{k-1}$ have degree two in G. A graph H is a *subdivision* of G if H is obtained from G by replacing each edge of G with a suspended path, that is to say, H is obtained by adding vertices on the edges of G.

A often used measure for sparseness of graphs is K_r-freeness as we have met in Chapter 3. The simplest K_3-free graphs are bipartite graphs. However, there are K_3-free graphs whose chromatic number can be arbitrarily large, see Mycielski's construction (1955) in the exercises. A more general measure for sparseness is to forbid some subdivision. Hajós conjectured that every graph G with $\chi(G) \geq r$ contains a subdivision of K_r as a subgraph. This conjecture is trivial for $r = 2, 3$, and it is confirmed by Dirac (1952) for $r = 4$, and it is open for $r = 5, 6$. Catlin (1979) disproved the conjecture for $r \geq 7$ by a constructive proof, but the disproof of Erdős and Fajtlowicz (1981) was more powerful. Let $\gamma(G)$ denote the largest r such that G contains a subdivision of K_r as a subgraph. Hajós conjecture is equivalent to that $\gamma(G) \geq \chi(G)$.

Theorem 2.9 *Almost all graphs $G = G_p \in \mathcal{G}(n, 1/2)$ satisfy*

$$\chi(G) \geq \frac{n}{2\log_2 n}, \quad \text{and} \quad \gamma(G) \leq \sqrt{6n}.$$

Proof. Set $k = \lfloor 2\log_2 n \rfloor$. Since

$$\Pr(\alpha(G) \geq k) \leq \binom{n}{k} 2^{-\binom{k}{2}} < \left(\frac{e\sqrt{2}n}{k2^{k/2}}\right)^k \to 0,$$

and $\alpha(G)\chi(G) \geq n$ for any graph G, the first statement follows immediately. Set $r = \lceil \sqrt{6n} \rceil$. Then $n \leq r^2/6$. There are

$$\binom{n}{r} \leq \left(\frac{en}{r}\right)^r \leq \left(\frac{er}{6}\right)^r$$

potential K_r subdivisions, one for each $r-$element subset of $V(G)$. If we fix such a subset X, then, since each subdivided edge has to use a distinct vertex of $V(G) \setminus X$, there are $\binom{r}{2}$ suspended pathes in a subdivision, and at most $n-r$ of them are of length two or more, which are "really" subdivided edges. So the subgraph induced by X contains at least

$$\binom{r}{2} - (n-r) \geq \binom{r}{2} + r - \frac{r^2}{6} \geq \frac{2}{3}\binom{r}{2}$$

edges. But the number of edges in subgraph induced by X, denoted by $e(X)$, has binomial distribution $B(N, 1/2)$, where $N = \binom{r}{2}$. From Chernoff bound in the last section,

$$\Pr\Big(e(X) \geq N(1+\delta)/2\Big) \leq \exp\{-N\delta^2/2\},$$

by taking $\delta = 1/3$ hence $\frac{2}{3}\binom{r}{2} = \binom{r}{2}(1+\delta)/2$, we have

$$\Pr\left(e(X) \geq \frac{2}{3}\binom{r}{2}\right) \leq \exp\{-N\delta^2/2\} = \exp\left\{-\frac{1}{18}\binom{r}{2}\right\}.$$

Thus we bound the probability that our random graph G contains a subdivision of K_r as follows.

$$\begin{aligned} \Pr(\gamma(G) \geq r) &\leq \sum_X \Pr\left(e(X) \geq \frac{2}{3}\binom{r}{2}\right) \\ &\leq \binom{n}{r}\exp\left\{-\frac{1}{18}\binom{r}{2}\right\} \leq \left(\frac{er\exp\{-(r-1)/36\}}{6}\right)^r, \end{aligned}$$

which tends to zero as $n \to \infty$. □

Since for almost all G in $\mathcal{G}(n, 1/2)$,

$$\chi(G) - \gamma(G) \geq \frac{n}{2\log_2 n} - \sqrt{6n} \to \infty$$

as $n \to \infty$, so Hajós conjecture failed badly, and almost all graphs in $\mathcal{G}(n, 1/2)$ are counterexamples.

We shall have another application of Chernoff's bounds for Ramsey numbers $r_k(K_{m,n})$ and $r(K_{m,n}, K_n)$.

Recall that $r_k(G)$ is the smallest integer N such that in any $k-$coloring of edges of K_N, there is a monochromatic G. Chung, Erdős, and Graham (1975) propose a problem to determine $r_k(K_{m,n})$. We now give a lower bound for it as k and m are fixed and $n \to \infty$, in which $\sqrt{n \log n}$ can be replaced by $\sqrt{n}\,\omega(n)$, where $\omega(n) \to \infty$.

Theorem 2.10 *Let integers k, $m \geq 1$ be fixed. Then, there is constant $C = C(k, m) > 0$ such that*

$$r_k(K_{m,n}) \geq k^m n - C\sqrt{n \log n}$$

for all large n.

Proof. Set $N = k^m n - C\sqrt{n \log n}$, where C is a constant to be determined. Then

$$n = \Big(\frac{1}{k^m} + \frac{C\sqrt{n \log n}}{k^m N}\Big)N > (k^{-m} + \delta_n)N = (p + \delta_n)N,$$

where $p = k^{-m}$ and $\delta_n = \frac{C}{2k^{2m}}\sqrt{\frac{\log n}{n}}$. Let us color the edges of K_{N+m} with k colors randomly and independently, such that each edge is assigned in each color with probability $1/k$. Consider a fixed color, say color A, and an arbitrary but fixed set U of m vertices. Let $v_1, v_2, \ldots, v_N$ be the N vertices outside U. For each j, define a random variable X_j such that $X_j = 1$ if the edges between v_j and U are all in color A and 0 otherwise. Then $\Pr(X_j = 1) = k^{-m} = p$. Set $S_N = \sum_{i=1}^N X_j$. Clearly S_N has the binomial distribution $B(N, p)$ and the event $S_N \geq n$ means that there is a monochromatic $K_{m,n}$ in color A (in which U is the m-vertex part). Hence

$$\Pr(\exists \text{ monochromatic } K_{m,n}) \leq k\binom{N+m}{m}\Pr(S_N \geq n).$$

By virtue of Chernoff bound (Theorem 2.5)

$$\Pr(S_N \geq n) \leq \Pr(S_N \geq (p + \delta_n)N) \leq \exp\{-N\delta_n^2/(3pq)\}.$$

From the facts that

$$-\frac{N\delta_n^2}{3pq} \sim \frac{-C^2 \log n}{12k^m(k^m-1)}$$

and

$$k\binom{N+m}{m} = O(n^m) = O\Big(\exp\{m \log n\}\Big),$$

we have that the probability that there exists monochromatic $K_{m,n}$ tends to zero as $N \to \infty$ if $C \geq k^m\sqrt{12m}$, which guarantees the existence of an edge-coloring of K_{N+m} with no monochromatic $K_{m,n}$, implying that $r_k(K_{m,n}) > N + m$ for all large n. □

Theorem 2.11 *Let integer $m \geq 2$ be fixed. Then there exists a constant $c = c(m) > 0$ such that*

$$r(K_{m,n}, K_n) \geq c\frac{n^{m+1}}{(\log n)^m}.$$

Proof. The lower bound is obtained through a simple application of Chernoff bound (Theorem 2.7). Let

$$N = \left\lfloor \frac{n^{m+1}}{3(2m \log n)^m} \right\rfloor$$

and $p = (2m \log n)/n$. The probability that m chosen vertices in $G(N,p)$ are connecting with another fixed vertex is p^m. So the probability that they have at least n common neighbors is $\Pr(S \geq n)$, where S has the binomial distribution $B(N-m, p^m)$. Then $n > Np^m$ and Theorem 2.7 yields

$$\Pr(K_{m,n} \subseteq G(N,p)) \leq \binom{N}{m}\left(\frac{(N-m)p^m e}{n}\right)^n$$
$$< \frac{N^m}{m!}\left(\frac{Np^m e}{n}\right)^n < c_1\frac{n^{m(m+1)}}{(\log n)^{m^2}}\left(\frac{e}{3}\right)^n,$$

where $c_1 = c_1(m) > 0$ is a constant. Hence $\Pr(K_{m,n} \subseteq G(N,p)) \to 0$. At the time, by standard estimates that $\binom{N}{n} \leq (Ne/n)^n$ and $1 -$

$p < e^{-p}$, we obtain a bound of the probability that $G(N,p)$ has an independent set of size at least n as follows

$$\Pr(\alpha(G(N,p)) \geq n) \leq \binom{N}{n}(1-p)^{n(n-1)/2}$$
$$\leq \left(\frac{Ne}{n}e^{-p(n-1)/2}\right)^n \leq \left(\frac{c_2}{3(2m\log n)^m}\right)^n,$$

where $c_2 = c_2(m) > 0$ is a constant, so $\Pr(\alpha(G(N,p)) \geq n) \to 0$. Hence the probability that $G(N,p)$ contains neither $K_{m,n}$ as a subgraph nor an independent set of size n is positive (in fact, close to 1). Thus $r(K_{m,n}, K_n) > N$. $\square$.

In Chapter 3, we proved that for any fixed $m \geq 1$,

$$r(K_m + \overline{K_n}, K_n) \leq (1+o(1))\frac{n^{m+1}}{(\log n)^{m-1}}.$$

So the upper bound and lower bound are just a $\log n$ factor away. In Chapter 8 we shall show that the obtained lower bound is the right order of $r(K_{m,n}, K_n)$.

2.3 Martingales on random graphs $\star$

Most parameters of a random graph are concentrated around their expectations. To describe such phenomena, martingale is a powerful tool, which may liberate us from drudgery computations.

Let X and Y be random variables on a probability space Ω. Given $Y = y$ with $\Pr(Y = y) > 0$, we define a conditional expectation $E(X|Y = y)$ as

$$E(X|Y=y) = \sum_x x \Pr(X = x|Y = y),$$

which is a number depending on y. As Y is random, we have a new random variable $E(X|Y)$. For an element $\omega \in \Omega$, if $Y(\omega) = y$, then $E(X|Y)$ takes value $E(X|Y = y)$ at ω.

Lemma 2.2 $E[E(X|Y)] = E[X]$.

Proof. From the definition, we have

$$\begin{aligned}
E[E(X|Y)] &= \sum_y E[X|Y=y]\Pr(Y=y) \\
&= \sum_y \left(\sum_x x \Pr[X=x|Y=y]\right)\Pr(Y=y) \\
&= \sum_x x \left(\sum_y \Pr[X=x|Y=y]\Pr(Y=y)\right) \\
&= \sum_x x \Pr(X=x) = E(X)
\end{aligned}$$

as asserted. □

A *martingale* is a sequence $X_0, X_1, \ldots, X_m$ of random variables so that for $0 \le i < m$,

$$E(X_{i+1}|X_i) = X_i;$$

namely, $E(X_{i+1}|X_i = x) = x$ for any given $X_i = x$.

Imagine one walks on a line randomly, at each step he moves one unit to the left or right with probability p, or stands still with probability $1-2p$. Let X_i be the position of i step. This is a martingale as the expected position after $i+1$ steps equals the actual position after i steps.

Let us look some martingales used in graph theory. The first is called the *edge exposure martingale* on chromatic numbers, in which we reveal G_p *one edge-slot at a time.* Let the random graph space $\mathcal{G}(n,p)$ be the underlying probability space. Set $m = \binom{n}{2}$, and label the potential edges on vertex set $[n]$ by $e_1, e_2, \ldots, e_m$ in any manner. As follows, we define $X_0(H), X_1(H), \cdots, X_m(H)$ for a given graph H on vertex set $[n]$, which are random variables if H is a random graph in $\mathcal{G}(n,p)$. Let $X_0(H) = E(\chi(G_p))$. For general i,

$$X_i(H) = E[\chi(G_p)|e_j \in E(G_p) \text{ iff } e_j \in E(H), 1 \le j \le i].$$

In other words, $X_i(H)$ is the expected value of $E[\chi(G_p)]$ under the condition that the set of the first i edges of G_p equals that of H while the remaining edges are not seen and considered to be random. Note that $X_0 = E(\chi(G_p))$ and $X_m = \chi(H)$.

Figure 1 shows why this is a martingale on the random space $\mathcal{G}(3, 1/2)$. Of cause, we can consider any graph parameters other than χ.

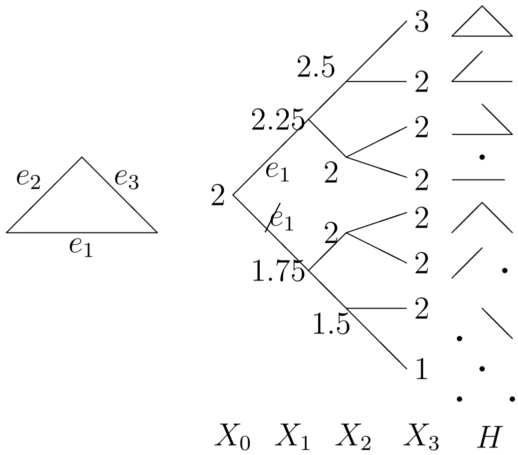

Fig. 5.1 An edge exposure martingale

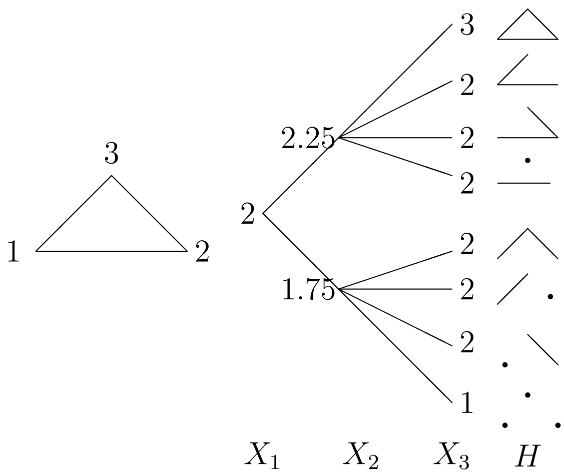

Fig. 5.2 A vertex exposure martingale

The second is called the *vertex exposure martingale* on chromatic numbers, in which we reveal G_p *one vertex-slot at a time*. Let the random graph space $\mathcal{G}(n,p)$ be the underlying probability space. We define $X_1 = E(\chi(G_p))$ and

$$X_i(H) = E[\chi(G_p)|E_i(G_p) = E_i(H)],$$

where $E_i(H)$ is the edge set induced by vertex set $\{1, \cdots, i\}$. In other words, $X_i(H)$ is the expected value of $E[\chi(G_p)]$ under the condition that the set of the edges of G_p induced by the first i vertices equals that of H while the remaining edges are not seen and considered to be random. Note that $X_1 = E(\chi(G_p))$ and $X_n = \chi(H)$. Note that the vertex exposure martingale is a subsequence of the edge exposure martingale.

In Fig. 5.1, The probability space is $\mathcal{G}(3, 1/2)$, so $X_0 = E(\chi(G_p)) = 2$, and $X_1(H) = 2.75$ if $e_1 \in E(H)$, and $X_1(H) = 1.75$ otherwise. Thus $E(X_1|X_0) = 2 = X_0$. The random variables X_2 and X_3 take 4 values and 8 values, respectively, and $E(X_{i+1}|X_i) = X_i$.

Lemma 2.3 *Let Y be a (discrete) random variable such that $E(Y) = 0$ and $|Y| \le 1$. Then $E(e^{tY}) \le (e^t + e^{-t})/2$ for all $t \ge 0$.*

Proof. For a fixed $t \ge 0$, set

$$h(y) = \frac{e^t + e^{-t}}{2} + \frac{e^t - e^{-t}}{2} y, \quad -1 \le y \le 1.$$

Note that the function $f(y) = e^{ty}$ is convex, and $h(y)$ is a line through the point $(-1, f(-1))$ and $(1, f(1))$ as $f(-1) = h(-1)$ and $f(1) = h(1)$, hence $e^{ty} \le h(y)$, and

$$E(e^{tY}) \le E(h(Y)) = \frac{e^t + e^{-t}}{2}$$

as $E(Y) = 0$, and thus the assertion follows. □

Theorem 2.12 (Azuma's Inequality) *Let $X_0, X_1, \cdots, X_m$ be a martingale with*

$$|X_{i+1} - X_i| \le 1$$

for all $0 \le i < m$*, and let* $\lambda > 0$*. Then*

$$\Pr[X_m - X_0 \ge \lambda\sqrt{m}] < e^{-\lambda^2/2},$$

and

$$\Pr[X_m - X_0 \le -\lambda\sqrt{m}] < e^{-\lambda^2/2}.$$

Proof. We may assume that $X_0 = 0$ by translation. Set $Y_i = X_i - X_{i-1}$, then $|Y_i| \le 1$ and $E(Y_i|X_{i-1}) = 0$. Then Lemma 2.3 yields that

$$E(e^{tY_i}|X_{i-1}) \le \frac{e^t + e^{-t}}{2} \le e^{t^2/2}$$

for any $t > 0$, where the last inequality is clear. Hence by Lemma 2.2, we have

$$\begin{aligned}
E(e^{tX_m}) &= E\left[e^{tX_{m-1}}e^{tY_m}\right] \\
&= E\left[E\left(e^{tX_{m-1}}e^{tY_m}|X_{m-1}\right)\right] \\
&= \sum_x E\left(e^{tX_{m-1}}e^{tY_m}|X_{m-1} = x\right)\Pr(X_{m-1} = x) \\
&= \sum_x e^{tx}E\left(e^{tY_m}|X_{m-1} = x\right)\Pr(X_{m-1} = x) \\
&\le e^{t^2/2}\sum_x e^{tx}\Pr(X_{m-1} = x) \\
&= e^{t^2/2}E(e^{tX_{m-1}}).
\end{aligned}$$

This and the induction gave $E(e^{tX_m}) \le e^{mt^2/2}$. Using Markov's Inequality, we obtain

$$\begin{aligned}
\Pr(X_m \ge \lambda\sqrt{m}) &= \Pr(e^{tX_m} \ge e^{t\lambda\sqrt{m}}) \\
&\le \frac{E(e^{tX_m})}{e^{t\lambda\sqrt{m}}} \le \frac{e^{mt^2/2}}{e^{t\lambda\sqrt{m}}}.
\end{aligned}$$

The assertion follows by letting $t = \lambda/\sqrt{m}$. □

2.4 Parameters of random graphs ⋆

We are ready to discuss some parameters of random graph G_p for fixed p. It is easy to see some parameters are concentrated around their expectations. The following result was due Shamir and Spencer (1987).

Theorem 2.13 *Let n and p be arbitrary and let $G_p \in \mathcal{G}(n,p)$. Then*

$$\Pr\left(|\chi(G_p) - E(\chi(G_p))| > \lambda\sqrt{n-1}\right) < 2e^{-\lambda^2/2}.$$

Proof. Consider the vertex exposure martingale $X_1, \cdots, X_n$ on $\mathcal{G}(n,p)$ with the parameter $\chi(G)$. A single vertex can always be given a new color so Azuma's Inequality can apply. □

Similarly, we have

$$\Pr\left(|\omega(G_p) - E(\omega(G_p))| > \lambda\sqrt{n-1}\right) < 2e^{-\lambda^2/2},$$

and

$$\Pr\left(|e(G_p) - E(e(G_p))| > \lambda\sqrt{m}\right) < 2e^{-\lambda^2/2},$$

where $m = \binom{n}{2}$. However, the proofs give no clue that what are these expectations.

Lemma 2.4 *Let $0 < p < 1$, $a = 1/p$ and $\epsilon > 0$ be fixed, and let $f(x) = \binom{n}{x}p^{\binom{x}{2}}$ for $0 \le x \le n$. Define a positive integer k such that*

$$f(k-1) > 1 \ge f(k).$$

Then as $n \to \infty$,

$$\lceil \omega_n - \epsilon \rceil \le k \le \lfloor \omega_n + \epsilon \rfloor + 1,$$

where

$$\omega_n = 2\log_a n - 2\log_a \log_a n + 2\log_a(e/2) + 1,$$

and $f(k-4) > c\left(\frac{n}{\log_a n}\right)^3 = n^{3-o(1)}$, where $c > 0$ is a constant.

Proof. It is easy to know that $k \to \infty$ and $k = o(\sqrt{n})$, thus by Stirling's formula, we have

$$f(k) = \binom{n}{k} p^{\binom{k}{2}} \sim \frac{n^k}{k!} p^{k(k-1)/2} \sim \frac{1}{\sqrt{2\pi k}} \left(\frac{en}{k} p^{(k-1)/2} \right)^k.$$

So if $\delta > 0$ fixed, for all large n,

$$\frac{en}{k} p^{(k-1)/2} \leq 1 + \delta$$

as $f(k) \leq 1$. This is equivalent to that

$$k \geq 2 \log_a n - 2 \log_a k + 2 \log_a e + 1 - 2 \log_a (1 + \delta).$$

Let us set $k \sim 2 \log_a n$ first. Then the difference between the right hand side in the above inequality and ω_n is

$$2 \log_a \frac{2 \log_a n}{k} - 2 \log_a (1 + \delta) \to -2 \log_a (1 + \delta),$$

so $k - \omega_n \geq -2 \log_a (1+\delta) + o(1) \geq -\epsilon$ if we take δ small enough. Hence $k \geq \omega_n - \epsilon$.

Similarly, from

$$f(k-1) \sim \frac{1}{\sqrt{2\pi(k-1)}} \left(\frac{en}{k-1} p^{(k-2)/2} \right)^{k-1},$$

we have $\frac{en}{k-1} p^{(k-2)/2} \geq 1$, which gives

$$k \leq 2 \log_a n - 2 \log_a (k-1) + 2 \log_a e + 2.$$

Furthermore, by taking $k \sim 2 \log_a n$ first, we obtain $k \leq \omega_n + 1 + o(1) \leq \omega_n + \epsilon + 1$, the desired upper bound for k follows.

Finally, note that

$$f(k-2) > \frac{f(k-2)}{f(k-1)} = \frac{k-1}{n-k+2} a^{k-2} \sim p^2 \frac{k}{n} a^k > \frac{c\,n}{\log n},$$

the assertion for $f(k-4)$ follows immediately. □

Lemma 2.5 *For fixed* $0 < p < 1$, $a = 1/p$ *and* $\epsilon > 0$, *almost all graphs* $G_p \in \mathcal{G}(n,p)$ *satisfy*

$$\omega(G_p) < \lfloor \omega_n + \epsilon \rfloor < 2\log_a n,$$

where ω_n *is defined in Lemma 2.4.*

Proof. Let X_r be the number of r-cliques, where r is referred as an integer. Then

$$E(X_r) = f(r) = \binom{n}{r} p^{\binom{r}{2}} \le \frac{n^r}{r!} p^{r(r-1)/2} < \frac{1}{\sqrt{2\pi r}} \left(\frac{en}{r} p^{(r-1)/2} \right)^r.$$

We shall find some $r = r(n) \to \infty$ such that $E(X_r) \to 0$. This is certainly true if $enp^{(r-1)/2}/r \le 1$ (hence $r \to \infty$). The same argument in the proof of Lemma 2.4 applies that if $r = \lceil \omega_n + \epsilon \rceil$, then $E(X_r) \to 0$, thus $\Pr[\omega(G_p) \ge r] \to 0$ and $\Pr[\omega(G_p) \le \lfloor \omega_n + \epsilon \rfloor] \to 1$. □

Note that the above result can be stated as

$$\Pr\left(\omega(G_p) \le \lceil \omega_n + \epsilon \rceil - 1\right) \to 1.$$

Matula (1970, 1972, 1976) was the first to notice that for fixed values of p almost all $G_p \in \mathcal{G}(n,p)$ have clique numbers concentrated on (at most) two values,

$$\lfloor \omega_n - \epsilon \rfloor \le \omega(G_p) \le \lfloor \omega_n + \epsilon \rfloor.$$

Results asserting this phenomenon were proved by Grimmett and McDiarmid (1975); and these were further strengthened by Bollobás and Erdős (1976).

In order to reduce the difficulty of the proof and preserve the typical flavor, we slightly weaken the above lower bound $\lfloor \omega_n - \epsilon \rfloor$ by having its asymptotical form a little bit later. Let us discuss the chromatic numbers first. A technical lemma is as follows.

Lemma 2.6 *Let* k *be the integer defined in Lemma 2.4 and let* $\ell = k - 4$. *Let* $Y = Y(G)$ *be the maximum size of a family of edge-disjoint cliques of size* ℓ *in* $G \in \mathcal{G}(n,p)$. *Then*

$$E(Y) \ge \frac{c\,n^2}{\ell^4},$$

where $c > 0$ *is a constant.*

Proof. Let $\mathcal{L}$ denote the family of ℓ-cliques of G. Then by Lemma 2.4, we have

$$\mu = E(|\mathcal{L}|) = f(\ell) = \binom{n}{\ell} p^{\binom{\ell}{2}} \geq c_1 \binom{n}{\ell}^3 .$$

Let W denote the number of unordered pairs $\{A, B\}$ of ℓ-cliques of G with $A \sim B$, where $A \sim B$ signifies that $2 \leq |A \cap B| < \ell$. Let

$$\Delta = \sum_{A \sim B} \Pr(AB),$$

where the sum is taken over all ordered pairs $\{A, B\}$. Then $E(W) = \Delta/2$ and

$$\begin{aligned} \Delta &= \binom{n}{\ell} \sum_{i=2}^{\ell-1} \binom{\ell}{i} \binom{n-\ell}{\ell-i} p^{2\binom{\ell}{2} - \binom{i}{2}} \\ &= \mu \sum_{i=2}^{\ell-1} \binom{\ell}{i} \binom{n-\ell}{\ell-i} p^{\binom{\ell}{2} - \binom{i}{2}} = \mu \sum_{i=2}^{\ell-1} R_i . \end{aligned}$$

Setting $a = 1/p$, we have

$$\frac{R_{i+1}}{R_i} = \frac{(\ell - i)^2}{(i+1)(n - 2\ell + i + 1)} a^i .$$

If i is small, say bounded, then this ratio is $O((\log_a n)^2/n)$, and if i is large, say $\ell - i = O(1)$, then the ratio is at least $\sqrt{n}$. It is increasing on i, so

$$\Delta = \mu \sum_{i=2}^{\ell-1} R_i \leq 2\mu(R_2 + R_{\ell-1}).$$

Here

$$\begin{aligned} R_2 &= \binom{\ell}{2} \binom{n-\ell}{\ell-2} p^{\binom{\ell}{2} - 1} \\ &= \frac{\ell^2(\ell-1)^2}{2p(n-\ell+2)(n-\ell+1)} \mu \leq \frac{\ell^4}{2pn^2} \mu , \end{aligned}$$

and

$$R_{\ell-1} = \ell(n-\ell) p^{\binom{\ell}{2} - \binom{\ell-1}{2}} \leq n\ell p^{\ell-1},$$

thus

$$\Delta \leq 2\mu\left(\frac{\ell^4}{2pn^2}\mu + n\ell p^{\ell-1}\right) \leq C\,\frac{\mu^2\ell^4}{n^2}.$$

Let $\mathcal{C}$ be a random subfamily of $\mathcal{L}$ defined by setting for each $A \in \mathcal{L}$,

$$\Pr[A \in \mathcal{C}] = p_1,$$

where $0 < p_1 < 1$ will be determined. Then $E(|\mathcal{C}|) = \mu p_1$. Let W' be the number of unordered pairs $\{A, B\}$ of ℓ-cliques in $\mathcal{C}$ with $A \sim B$. Then

$$E(W') = E(W)p_1^2 = \frac{\Delta p_1^2}{2}.$$

Delete from $\mathcal{C}$ one set from each such pair $\{A, B\}$. This yields a set $\mathcal{C}^*$ of edge-disjoint ℓ-cliques of G and

$$E(Y) \geq E(|\mathcal{C}^*|) \geq E(|\mathcal{C}|) - E(W') = \mu p_1 - \frac{\Delta p_1^2}{2}.$$

By choosing $p_1 = \frac{\mu}{\Delta} < 1$, we have

$$E(Y) \geq \frac{\mu^2}{2\Delta} \geq \frac{c\,n^2}{\ell^4}$$

as asserted. □

Theorem 2.14 (Bollobás) *Let $0 < p < 1$, $a = 1/p$ be fixed, and let $m = \lceil n/\log_a^2 n\rceil$. Then for almost all graphs $G_p \in \mathcal{G}(n,p)$, each induced subgraph of order m of G_p has a clique of size at least $r = 2\log_a n - 7\log_a\log_a n$.*

Proof. Let S be an m-set of vertices. We shall bound the probability that S induces no r-clique by $e^{-m^{1+\delta}}$ for all large n (hence all large m), where $\delta > 0$ is a constant. So the probability that there exists an m-set with no r-clique is at most

$$\binom{n}{m}e^{-m^{1+\delta}} < \left(\frac{en}{m}\right)^m e^{-m^{1+\delta}} = \exp\left(m\log_e\frac{en}{m} - m^{1+\delta}\right),$$

which goes to zero, and the assertion follows.

Let X be the maximum number of pairwise edge-disjoint r-cliques sets in this graph (induced by S), where *edge-disjoint* means they share at most one vertex. We shall show that $X \geq 1$ holds almost surely. To do this, we invoke Azuma's Inequality. Consider the edge exposure martingale for X that results from revealing G one-edge slot at a time. We have $X_0 = E(X)$ and $X_{\binom{m}{2}} = X$. Clearly the Lipschitz condition $|X_{i+1} - X_i| \leq 1$ is satisfied, so Azuma's Lemma gives

$$\begin{aligned}
\Pr(X = 0) &\leq \Pr[X - E(X) \leq -E(X)] \\
&= \Pr\left[X - E(X) \leq -\lambda \binom{m}{2}^{1/2}\right] \leq e^{-\lambda^2/2} \\
&= \exp\left(-\frac{E^2(X)}{m(m-1)}\right),
\end{aligned}$$

where $\lambda = E(X)/\binom{m}{2}^{1/2}$. Hence it suffices to find $\delta > 0$ such that $E^2(X) \geq m^{3+\delta}$ for all large n.

Now, let t_0 be the integer such that $f(t_0 - 1) > 1 \geq f(t_0)$, where $f(x) = \binom{m}{x} p^{\binom{x}{2}}$, and let $t = t_0 - 4$. Then by Lemma 2.4, we have

$$t \geq 2\log_a m - 2\log_a \log_a m - 3 > 2\log_a n - 7\log_a \log_a n,$$

so $t > r$. Let T be the maximum number of edge-disjoint cliques of size t, Then $E(X) \geq E(T)$ and $E(T) \geq cm^2/t^4$ by Lemma 2.6, hence

$$E(X) \geq \frac{cm^2}{t^4} \sim \frac{cn^2}{16(\log_a n)^8},$$

implying that $E^2(X) \geq n^{4-o(1)} \geq n^{3+\delta}$ for any $1 > \delta > 0$ if n is large, which completes the proof. □

Theorem 2.15 (Bollobás) *Let $0 < p < 1$ and $\epsilon > 0$ be fixed. Denote $b = 1/q = 1/(1-p)$. Then almost all graphs $G_p \in \mathcal{G}(n,p)$ satisfy*

$$\frac{n}{2\log_b n} \leq \chi(G_p) \leq (1+\epsilon)\frac{n}{2\log_b n}.$$

Proof. The lower bound holds because almost all G_p satisfy $\alpha(G_p) \le 2\log_b n$ and $\chi(G)\alpha(G) \ge n$. The upper bound follows from the above theorem, which is applied for independent sets instead of cliques, because we can almost always select independent set of size $2\log_b n - 7\log_b\log_b n$ until we have only $n/\log_b^2 n < (\epsilon/2)n/(2\log_b n)$ vertices left. We first use at most

$$\frac{n}{2\log_b n - 7\log_b\log_b n} < \left(1+\frac{\epsilon}{2}\right)\frac{n}{2\log_b n}$$

colors, and then we can complete the coloring by using distinct new colors on each of the remaining vertices. □

Let us remark that Achlioptas and Naor recently obtained a result on sparser random graphs as follows. Given $d > 0$, let k_d be the smallest integer k such that $d < 2k\log k$. Then $\chi(G_p)$ for almost all $G_p \in \mathcal{G}(n, d/n)$ is either k_d or $k_d + 1$. This result improves an earlier result of Łuczak (1991) by specifying the form of k_d.

Theorem 2.16 *Let $0 < p < 1$ and $\epsilon > 0$ be fixed. Then almost all graphs $G_p \in \mathcal{G}(n,p)$ satisfy*

$$(1-\epsilon)2\log_b n \le \alpha(G_p) < 2\log_b n.$$

Proof. The upper bound is the complement of that in Lemma 2.5. The lower bound follows from Theorem 2.15 and the fact that $\alpha(G) \ge n/\chi(G)$. □

Theorem 2.17 *Let $0 < p < 1$ and $\epsilon > 0$ be fixed. Then almost all graphs $G_p \in \mathcal{G}(n,p)$ satisfy*

$$(1-\epsilon)2\log_a n \le \omega(G_p) < 2\log_a n.$$

Proof. This is complement of Theorem 2.16. □

For some graph parameter $f(G)$, we have seen that there is a function $g(n)$ such that almost all graphs G_p in $\mathcal{G}(n,p)$ satisfy that

$$(1-\epsilon)g(n) \le f(G_p) \le (1+\epsilon)g(n),$$

hence $f(G)$ concentrate in a small range. We shall call the function $g(n)$ as a *threshold* for the parameter f. We will discuss the threshold for probability $p = p(n)$ instead of fixed p, and will consider some other graph parameters in the next chapter.

2.5 References

D. Achlioptas and A. Naor, The two possible values of the chromatic number of a random graph, *Ann. of Math.*, **162** (2005), 1335-1351.

N. Alon and J. Spencer, *The Probabilistic Method, 3rd Edition*, Wiley-Interscience, New York, 2008.

B. Bollobás, The chromatic numbers of random graphs, *Combinatorica*, **8** (1988), 49-55.

B. Bollobás and P. Erdős, Cliques in random graphs, *Math. Proc. Cambridge Philos. Soc.*, **80** (1976), 419-427.

P. Catlin, Hajós' graph-coloring conjecture: variations and counterexamples, *J. Combin. Theory, Ser. B*, **26** (1979), 268-274.

H. Chernoff, A measure of the asymptotic efficiency for tests of a hypothesis based on the sum of observations, *Ann. Math. Statis.* , **23** (1952), 493-509.

F. Chung and R. Graham, On multicolor Ramsey numbers for bipartite graphs, *J. Combin. Theory Ser. B*, **18** (1975),164-169.

G. Dirac, A property of 4-chromatic graphs and some remarks on critical graphs, *J. Lond. Math. Soc.*, **27** (1952), 85-92.

P. Erdős and S. Fajtlowicz, On the conjecture of Hajós, *Combinatorica*, **1** (1981) , 141-143.

G. Grimmett and C. McDiarmid, On coloring random graphs, *Math. Proc. Cambridge Philos. Soc.*, **77** (1975), 313-324.

Y. Li and C. Rousseau, On the Ramsey number $r(H + \bar{K}_n, K_n)$, *Discrete Math.*, **170** (1997), 265-267.

T. Łuczak, The chromatic number of random graphs, *Combinatorica*, **11** (1991), 45-54.

T. Łuczak, A note on the sharp concentration of the chromatic number of random graphs, *Combinatorica*, **11** (1997), 295-297.

D.W. Matula, On the complete subgraphs of a random graph, in: *Proc. Second Chapel Hill Conference on Combinatory Mathematics and its Applications*, University of North Carolina, Chapel Hill, North Caroline, 1970.

D.W. Matula, The employee party problem, *Notices Amer. Math. Soc.*, **19** (1972), A-328.

D.W. Matula, The largest clique size in a random graph, *Tech. Rep.*, Dept. Comput. Sci., Southern Methodist University, Dallas, 1976.

J. Mycielski, Sur le coloring des graphs, *Coll. Math.*, **3** (1955), 161-162.

E. Shamir and J. Spencer, Sharp concentration of the chromatic number in random graph $G_{n,p}$, *Combinatorica*, **7** (1987), 121-130.

Chapter 3

Properties of Random Graphs

As we have seen that the random graph is important in probabilistic method for Ramsey theory. In this chapter, let us digress a bit from Ramsey theory to general properties of random graphs. This chapter contains some of most important topics on random graphs, which and Chapter 5 form a short path to the theory. A random graph is obtained by starting with a set of n vertices and adding edges between them at random. Different random graph models produce different probability distributions on graphs, for which the model in this chapter is classic. Erdős and Rényi showed that for many monotone-increasing properties of random graphs, graphs of a size slightly less than a certain threshold are very unlikely to have the property, whereas graphs with a few more graph edges are almost certain to have it. This is known as a phase transition. The second section is devoted to this topic. The last section covers some deeper discussion, for which the beginners of readers can skip.

3.1 Some behavior of almost all graphs

Given a graph property A, it is often associated with a family Q of graphs as

$$Q = Q(A) = \{G : G \text{ has } A\}.$$

Slightly abusing notation, we do not distinguish the property A and the family Q if no danger of confusion. We say that almost all (a.a.) graphs in $\mathcal{G}(n,p)$ have property Q if $\lim_{n\to\infty}\Pr[G_p \in Q] = 1$. In this case we also say that *almost surely* that $G_p \in \mathcal{G}(n,p)$ has property Q. We begin at a classic result of Erdős that almost all graphs seem to behave strangely even though they are sparse. In the following result, we denote by $\langle S\rangle$ the subgraph of G induced S.

Theorem 3.1 (Erdős) *For any $k \geq 1$, there exist positive constants $c = c(k)$ and $\epsilon = \epsilon(k)$ such that almost all graphs in $\mathcal{G}(n,p)$ with $p = c/n$ satisfy that $\chi(G) \geq k$, and yet $\chi(\langle S\rangle) \leq 3$ for any vertex subset S with $|S| \leq \epsilon n$.*

Proof. Let

$$H(x) = -\log\left(x^x(1-x)^{1-x}\right), \quad 0 < x < 1,$$

and let constants c and ϵ satisfy

$$c > 2k^2 H(1/k) \quad \text{and} \quad c^3 e^5 \epsilon < 3^3.$$

Set $p = c/n$ and consider $G = G_p$ in $\mathcal{G}(n,p)$. We show that almost all graphs in this space satisfy the conditions. If $\chi(G) \leq k$, then $\alpha(G) \geq n/k$. The expected numbers of such independent sets is

$$\binom{n}{n/k}(1-p)^{\binom{n/k}{2}}.$$

From Stirling formula, we estimate that

$$\binom{n}{n/k} = \frac{n!}{(n/k)!(n-n/k)!} \leq \exp\{nH(1/k)\},$$

and

$$(1-p)^{\binom{n/k}{2}} \leq \exp\{-\frac{pn}{2k}(\frac{n}{k}-1)\} = \exp\{-\frac{cn}{2k^2}(1-o(1))\}.$$

Therefore,

$$\binom{n}{n/k}(1-p)^{\binom{n/k}{2}} \leq \exp\left\{-n\left(\frac{c}{2k^2} - H(\frac{1}{k}) - o(1)\right)\right\},$$

which tends to zero by the condition satisfied by c.

Suppose some set S with $t \le \epsilon n$ vertices such that $\chi(\langle S \rangle) \ge 4$, we claim that $\langle S \rangle$ would have at least $3t/2$ edges. Suppose S is a minimal such set. For any $v \in S$, there would be a (proper) $3-$coloring of $S \backslash \{v\}$. If v has two or fewer neighbors in $\langle S \rangle$ then it would be extended to a $3-$coloring of S. Hence the minimum degree of $\langle S \rangle$ is at least 3 and the claim follows. The probability that some $t \le \epsilon n$ vertices have at least $3t/2$ edges is less than

$$\sum_{4 \le t \le \epsilon n} \binom{n}{t} \binom{\binom{t}{2}}{3t/2} \left(\frac{c}{n}\right)^{3t/2} = \sum_{4 \le t \le n^{1/4}} + \sum_{n^{1/4} < t \le \epsilon n} = s_1 + s_2.$$

We bound

$$\binom{n}{t} \le \left(\frac{en}{t}\right)^t \quad \text{and} \quad \binom{\binom{t}{2}}{3t/2} \le \left(\frac{et}{3}\right)^{3t/2}.$$

So each term is at most

$$\left(\frac{en}{t}\right)^t \left(\frac{et}{3}\right)^{3t/2} \left(\frac{c}{n}\right)^{3t/2} = \left(\frac{c^{3/2} e^{5/2} t^{1/2}}{3^{3/2} n^{1/2}}\right)^t.$$

Hence the first summation

$$s_1 \le n^{1/4} \left(\frac{c^{3/2} e^{5/2} n^{1/4}}{3^{3/2} n^{1/2}}\right)^4 \to 0.$$

Each term in the second summation is at most

$$\left(\frac{c^{3/2} e^{5/2} t^{1/2}}{3^{3/2} n^{1/2}}\right)^t \le \left(\frac{c^{3/2} e^{5/2}}{3^{3/2}} \epsilon^{1/2}\right)^{n^{1/4}}.$$

Let δ be the bracketed term $\frac{c^{3/2} e^{5/2}}{3^{3/2}} \epsilon^{1/2}$, then $\delta < 1$ by the choice of ϵ. Thus the second summation

$$s_2 \le \epsilon n \delta^{n^{1/4}} \to 0.$$

So almost surely no such set S exists, completing the proof. □

By the above theorem, we know that in random graphs, the neighbors of average vertices distribute evenly in every part of the vertex set. So their clique numbers and independence numbers are small, and chromatic number are big. Let $g(G)$ be the girth of G, which is the smallest length of a cycle in G. A historic result of Erdős (1959) is that both of $\chi(G)$ and $g(G)$ can be arbitrarily large.

Theorem 3.2 *For any fixed ℓ and k, there exists a graph G with $g(G) > \ell$ and $\chi(G) > k$.*

Proof. Fix $0 < \theta < 1/\ell$, let $p = n^{\theta-1}$. Consider random graphs in $\mathcal{G}(n,p)$. Let X be the number of cycles of length at most ℓ in G. Then

$$E(X) = \sum_{i=3}^{\ell} \frac{(n)_i}{2i} p^i \leq \sum_{i=3}^{\ell} \frac{n^{\theta i}}{2i} = o(n)$$

as $\theta\ell < 1$, where $(n)_i$ is the falling factorial $n(n-1)\cdots(n-i+1)$. On the other hand,

$$E(X) = \sum_i i \Pr(X = i) \geq \frac{n}{2} \Pr(X \geq n/2),$$

which implies that $\Pr(X \geq n/2) = o(1)$ by $E(X) = o(n)$.

Set $m = 3n^{1-\theta} \log n$. Then it is a routine procedure to show

$$\Pr(\alpha(G) \geq m) \leq \binom{n}{m} (1-p)^{\binom{m}{2}} < \left(ne^{-p(m-1)/2}\right)^m = o(1).$$

There exists a graph G of large order n such that $X(G) < n/2$ and $\alpha(G) < m$. By deleting a vertex from each cycle of length at most ℓ, we obtain a graph G^* of order at least $n/2$, which satisfies $g(G^*) > \ell$, $\alpha(G^*) \leq m$, and

$$\chi(G^*) \geq \frac{|V(G^*|}{\alpha(G^*)} \geq \frac{n/2}{m} \geq \frac{n^\theta}{6 \log n} > k,$$

completing the proof. □

3.2 Threshold functions

For fixed $0 < p \leq 1$, most graphs in $\mathcal{G}(n,p)$ are dense. As we have seen in the last chapter, Bollobás (1988) proved that the chromatic numbers $\chi(G_p)$ for $G_p \in \mathcal{G}(n,p)$ are concentrated at $n/(2\log_{1/q} n)$, where $q = 1 - p$. In this section, we investigate the concentration of edge probability function $p = p(n)$ associated with a property. We have seen that random graphs in $\mathcal{G}(n,p)$ behave sensitively on $p = p(n)$. A monumental discovery of Erdős and Rényi (1960) was that many natural graph theoretic properties become true in a very narrow range of $p = p(n)$.

A property Q is said to be *monotone increasing* if G has Q, then any graph from G by adding new edges has Q. The *monotone decreasing* property can be defined similarly. Thus the property of being connected is monotone increasing and that of being triangle-free is monotone decreasing. As mentioned in Section 1, a property Q is associated with a family of graphs. We call this family to be monotone increasing if so is Q. Also we do not distinguish the property and its associated family.

Lemma 3.1 *Let Q be a monotone increasing property. For $G_p \in \mathcal{G}(n,p)$, the function $\Pr(G_p \in Q)$ is increasing on p.*

Proof. Let $0 \leq p_1(n) < p_2(n) \leq 1$. We shall verify

$$\Pr(G_{p_1} \in Q) \leq \Pr(G_{p_2} \in Q).$$

Set $p = (p_2 - p_1)/(1 - p_1)$, then $p_2 = p + p_1 - pp_1$. Choose $G \in \mathcal{G}(n,p)$ and $G_1 \in \mathcal{G}(n,p_1)$, independently, and set $G_2 = G \cup G_1$. Namely G_2 is a graph on vertex set $V = [n]$ with edge set $E(G) \cup E(G_1)$, in which each edge e appears with probability

$$\Pr(e) = \Pr(e \in E(G) \cup E(G_1)) = p + p_1 - pp_1 = p_2$$

as the events that e appears in $E(G)$ and in $E(G_1)$ are independent. Thus G_2 is exactly a random graph of $\mathcal{G}(n,p_2)$. As Q is monotone increasing, if G_1 has Q so does G_2, thus $\Pr(G_{p_1} \in Q) \leq \Pr(G_{p_2} \in Q)$ as claimed. $\square$

Let Q be a monotone increasing property. Erdős and Rényi defined a function $f(n)$ with $0 \leq f(n) \leq 1$ as *threshold function* for Q if

$$\lim_{n\to\infty} \Pr\left(G_p \in Q\right) = \begin{cases} 0 & \text{if } p = f(n)/\omega(n) \\ 1 & \text{if } p = f(n)\omega(n), \end{cases}$$

where $0 < \omega(n) < 1/f(n)$ is a function which tends to infinity with n, as slowly as desired. For example, if $f(n) = \log n/n$, we may assume that $0 < \omega(n) < \log\log n$. Note that if $f(n)$ is a threshold function for Q, so is $cf(n)$ for any constant $c > 0$.

Clearly, the definition of $f(n)$ being a threshold function for a monotone increasing property Q is equivalent to

$$\lim_{n\to\infty} \Pr\left(G_p \in Q\right) = \begin{cases} 0 & \text{if } p \ll f(n) \\ 1 & \text{if } p \gg f(n), \end{cases}$$

where $p \ll f(n)$ means $p = o(f(n))$.

For obvious reason, the above threshold function is in fact *threshold probability function*. One can certainly define other threshold functions such as threshold edge function.

The definition of the threshold function for a monotone decreasing property is similar. The definitions means the situation whether or not a.a. G_p have Q changes suddenly even though $p = p(n)$ changes slightly in the moment.

Let $X = X(G)$ be a non-negative integral parameter of graph G. Since

$$\Pr(X \geq 1) = \sum_{k\geq 1} \Pr(X = k) \leq E(X),$$

so $E(X) \to 0$ implies that a.a. graphs in $\mathcal{G}(n,p)$ satisfy $X = 0$. And in many cases $E(X) \to \infty$ implies that a.a graphs in $\mathcal{G}(n,p)$ satisfy $X \geq 1$, which can be shown by Chebyshev's inequality often. For example, let X be the number of triangles in $G_p \in \mathcal{G}(n,p)$. Then

$$E(X) = \binom{n}{3} p^3 \sim \frac{1}{6}(np)^3.$$

As we will see in the next theorem that $f(n) = 1/n$ truly is a threshold function for triangle-containedness. Let $p = \gamma/n$ and let $\gamma \to 0$ or

$\gamma \to \infty$ signify $\omega(n)$ in the denominator or in the numerator in the definition, respectively. When γ reaches and passes 1, the structure of G_p changes radically. This is called the *double jump* because the structure of G_p is significantly different for $\gamma \ll 1$, $\gamma \sim 1$ and $\gamma \gg 1$.

Let us recall the Second Moment Method in last chapter.

Lemma 3.2 (Second Moment Method) *If X is a random variable, then*

$$\Pr(X = 0) \le \frac{E(X^2) - \mu^2}{\mu^2},$$

where $\mu = E(X)$. In particular, $\Pr(X = 0) \to 0$ *if* $E(X^2)/\mu^2 \to 1$.

A graph G with average degree d is called *balanced* if no subgraph of it has average degree greater than d. Complete graphs, cycles and trees are balanced.

Theorem 3.3 *Let F be a balanced graph with $k \ge 2$ vertices and $\ell \ge 1$ edges and let Q be the property that a graph contains F as a subgraph. Then $f(n) = n^{-k/\ell}$ is a threshold function for Q.*

Proof. To simplify the notation as before, we shall use $p = \frac{\gamma}{n^{k/\ell}}$ with $\gamma \to 0$ and $\gamma \to \infty$ to signify the function $\omega(n)$ in the denominator and numerator, respectively. Let $X = X(G)$ be the number of copies of F contained in $G = G_p \in \mathcal{G}(n,p)$. Denote by a for the number of graphs isomorphic to F on fixed k labeled vertices. Then

$$\mu = E(X) = \binom{n}{k} a p^\ell.$$

By noting the simple facts that $1 \le a \le k!$ with k and ℓ fixed, we have $E(X) \le n^k p^\ell = \gamma^\ell$ and $E(X) \sim \frac{a}{k!} n^k p^\ell = \frac{a}{k!}\gamma^\ell$. Thus

$$c_1\gamma^\ell \le c_2 n^k p^\ell \le \mu = E(X) \le \gamma^\ell,$$

where c_1 and c_2 henceforth c_i are positive constants. So the order of μ is $\gamma^\ell = n^k p^\ell$.

When $\gamma \to 0$, by Markov's inequality,

$$\Pr(G_p \in Q) = \Pr(X \ge 1) \le E(X) \to 0.$$

We then want to show that $\Pr(G_p \in Q) = \Pr(X \geq 1) \to 1$ as $\gamma \to \infty$. We turn to the Second Moment Method for help since the Markov's inequality does not work in this case.

For any k labeled vertices in $[n]$, we have $a = k!/|\mathcal{A}|$, where $\mathcal{A}$ is the automorphism group of F. Then there are total of $a\binom{n}{k}$ potential copies of F on $[n]$. Denote by

$$\mathcal{F} = \{F_1, F_2, \ldots\}$$

for the family of these copies. Denote by $F_i \cup F_j$ for the graph with vertex set $V(F_i) \cup V(F_j)$ and edge set $E(F_i) \cup E(F_j)$. The two critical observations are that most pairs F_i and F_j have no vertices in common, and if they have $s \geq 1$ common vertices and these s vertices contains t edges of F_j, then $t/s \leq \ell/k$ since F is balanced.

Let X_i be the indicator function of F_i. Then

$$E(X_i) = \Pr(X_i = 1) = \Pr(G \supset F_i).$$

Since $X = \sum_i X_i$,

$$E(X^2) = \sum_{i,j} E(X_i X_j) = \sum_{i,j} \Pr(G \supset F_i \cup F_j),$$

where the sum is taken over all pairs $\{i, j\}$ with $F_i, F_j \in \mathcal{F}$. Set

$$A_0 = \sum_{i,j\colon E(F_i) \cap E(F_j) = \emptyset} \Pr(G \supset F_i \cup F_j),$$

and for $s \geq 1$,

$$A_s = \sum_{i,\,j} \{\Pr(G \supset F_i \cup F_j) : \ |V(F_i) \cap V(F_j)| = s, E(F_i) \cap E(F_j) \neq \emptyset\}.$$

Then $E(X^2) = \sum_{s=0}^{k} A_s$. Note that if $E(F_i) \cap E(F_j) = \emptyset$, then

$$\Pr(G \supset F_i \cup F_j) = \Pr(G \supset F_i) \Pr(G \supset F_j)$$

from the independency of the events. We thus have $A_0 \leq \mu^2$ following from

$$A_0 = \sum_{V(F_i) \cap V(F_j) = \emptyset} \Pr(G \supset F_i \cup F_j)$$

$$
\begin{aligned}
&= \sum_{V(F_i)\cap V(F_j)=\emptyset} \Pr(G \supset F_i)\Pr(G \supset F_j) \\
&\le \Big(\sum_i \Pr(G \supset F_i)\Big)\Big(\sum_j \Pr(G \supset F_j)\Big) \\
&= E^2(X) = \mu^2.
\end{aligned}
$$

For $s \ge 1$ it is expected that A_s is much less than μ^2. Fix F_i, counting F_j that has s common vertices with F_i, in which these s common vertices contain t edges of $E(F_i)\cap E(F_j)$ with $t \le s\ell/k$ since F is balanced, we have

$$
\begin{aligned}
\sum_{j:\,|V(F_i)\cap V(F_j)|=s} \Pr(G \supset F_i \cup F_j) &\le \sum_{t\le s\ell/k} \binom{k}{s}\binom{n-k}{k-s} p^{2\ell-t} \\
&\le c_3 n^{k-s} \sum_{t\le s\ell/k} p^{2\ell-t}
\end{aligned}
$$

since k, s, ℓ are fixed and t is bounded. From the fact that there are $a\binom{n}{k}$ elements in $\mathcal{F}$, we obtain

$$
\begin{aligned}
A_s &\le a\binom{n}{k} c_3 n^{k-s} \sum_{t\le s\ell/k} p^{2\ell-t} \\
&\le c_4 n^{2k-s} \sum_{t\le s\ell/k} p^{2\ell-t} \le c_4 (n^k p^\ell)^2 n^{-s} \sum_{t\le s\ell/k} p^{-t} \\
&\le \frac{c_5 \gamma^{2\ell} n^{-s}}{p^{s\ell/k}} = \frac{c_5\gamma^{2\ell}}{(np^{\ell/k})^s} = \frac{c_5\gamma^{2\ell}}{\gamma^{s\ell/k}} \le \frac{c_6\mu^2}{\gamma^{s\ell/k}},
\end{aligned}
$$

where we used the fact that $n^k p^\ell$, γ^ℓ and μ have the same order. So for $s \ge 1$, we have $A_s/\mu^2 \le c_6/\gamma^{s\ell/k}$, and

$$
\frac{E(X^2)}{\mu^2} = \frac{A_0 + \sum_{s=1}^k A_s}{\mu^2} \le 1 + \sum_{s=1}^{k} \frac{c_6}{\gamma^{s\ell/k}} \le 1 + \frac{c_7}{\gamma^{\ell/k}}.
$$

By the Second Moment Method,

$$
\begin{aligned}
\Pr(X=0) &\le \Pr(|X-\mu| \ge \mu) \le \frac{\sigma^2}{\mu^2} \\
&= \frac{E(X^2)-\mu^2}{\mu^2} \le \frac{c_7}{\gamma^{\ell/k}},
\end{aligned}
$$

which tends to zero as $\gamma \to \infty$. □

In traditional definition of threshold function $f(n)$ of Erdős and Rényi $p_1/p_2 \to 0$ with $p_1(n) = f(n)/\omega(n)$ and $p_2(n) = f(n)\omega(n)$. On the other hand, if we have such $p_1(n)$ and $p_2(n)$, then $f(n) = \sqrt{p_1(n)p_2(n)}$ is a threshold function. In many cases it is just needed that p_1 is slightly less than p_2. A more precise definition of threshold function is as follows.

Let Q be a monotone increasing property of graphs. A function $p_\ell = p_\ell(n)$ is called a *lower threshold function* (ltf) if almost no graphs in $\mathcal{G}(n, p_\ell)$ have Q; a function $p_u = p_u(n)$ is called an *upper threshold function* (utf) if almost all graphs in $\mathcal{G}(n, p_u)$ have Q.

A realistic situation is very interesting. In a conference, a pair of mathematicians unknown each other can found a common mathematician friend. The following result called distance two theorem gives us a good explanation for this small world phenomenon. The property being distance two is not strictly increasing, but it is except the addition of the last edge.

Theorem 3.4 *For any function $\omega(n) \to \infty$ with $\omega(n) < \log n$, set*

$$p_\ell = \sqrt{\frac{2\log n - \omega(n)}{n}}, \quad \textit{and} \quad p_u = \sqrt{\frac{2\log n + \omega(n)}{n}}.$$

Then p_ℓ and p_u are ltf and utf for the property of graph having distance two, respectively.

Proof. Enumerate of all pairs of vertices $\{u, v\}$ of $\mathcal{G}(n, p)$ as $e_1, e_2, \ldots, e_m$ with $m = \binom{n}{2}$. For $e_k = \{u, v\}$, let $d(u, v)$ be the distance between u and v. Define

$$X_k = X_k(G) = \begin{cases} 0 & d(u,v) \le 2 \\ 1 & \text{otherwise,} \end{cases}$$

and $X = \sum_{k=1}^{m} X_k$. A non-complete graph G has distance two if and only if $X = 0$. Since the event $d(u, v) \ge 3$ for a pair of non-adjacent vertices is equivalent to that none of other $n-2$ vertices is adjacent to both u and v, so

$$E(X_k) = \Pr(X_k = 1) = (1-p)(1-p^2)^{n-2}.$$

Set $\mu = E(X)$, then

$$\mu = E(X) = \binom{n}{2}(1-p)(1-p^2)^{n-2}.$$

(i) Let $p = p_u = \sqrt{(2\log n + \omega(n))/n}$. We have

$$\begin{aligned} \mu &\sim \frac{n^2}{2}(1-p^2)^n \sim \frac{n^2}{2}e^{-np^2} \\ &= \frac{1}{2}e^{-\omega(n)} \to 0. \end{aligned}$$

Thus $\Pr(X \geq 1) \leq E(X) \to 0$, which proves that a.a. graphs in $\mathcal{G}(n,p)$ have distance at most two hence a.a. graphs in $\mathcal{G}(n,p)$ have distance two as almost no graph in $\mathcal{G}(n,p)$ is complete.

(ii) Let $p = p_\ell = \sqrt{(2\log n - \omega(n))/n}$. Suppose that $\omega(n) < \log\log n$ without loss of generality. Consider

$$E(X^2) = \sum_{i,j} E(X_iX_j) = A_0 + A_1 + A_2,$$

where $A_s = \sum_{|e_i \cap e_j| = s} E(X_iX_j)$, in which the sum is taken over all pairs $\{i,j\}$ with e_i and e_j having s vertices in common. Clearly

$$\mu = E(X) \sim \frac{n^2}{2}(1-p^2)^n \sim \frac{n^2}{2}e^{-np^2} = \frac{e^{\omega(n)}}{2} \to \infty,$$

and

$$A_0 = \sum_{|e_i \cap e_j| = 0} E(X_iX_j) \leq \binom{n}{2}\binom{n-2}{2}(1-p)^2(1-p^2)^{2(n-2)} < \mu^2.$$

Also

$$A_2 = \sum_{k=1}^{m} E(X_k) = \mu.$$

We now estimate A_1 that should not be big since A_0 counts most of pairs. For $e_i = \{u,v\}$ and $e_j = \{v,w\}$ with $|e_i \cap e_j| = 1$, the event $d(u,v) \geq 3$ and $d(v,w) \geq 3$ is equivalent to the following events happen:

B_1 : u and v, v and w are non-adjacent;

B_2 : none of vertices other than u, v, w is adjacent to all of the three vertices.

B_3 : none of distinct vertices is adjacent to both u and v, another is adjacent to both v and w;

We have

$$\begin{aligned} A_1 &= \sum_{|e_i \cap e_j|=1} E(X_i X_j) = \Pr(B_1 B_2 B_3) \leq \Pr(B_3) \\ &= 3\binom{n}{3}(1-p^2)^{2n-7} \\ &\sim \frac{n^3}{2}(1-p^2)^{2n} \sim \frac{n^3}{2} e^{-2np^2} = \frac{1}{2n} e^{2\omega(n)} \to 0. \end{aligned}$$

Hence

$$\sigma^2 = E(X^2) - \mu^2 = A_0 + A_1 + A_2 - \mu^2 < \mu + 1,$$

which and the Second Moment Method yield

$$\Pr(X = 0) \leq \Pr(|X - \mu| \geq \mu) \leq \frac{\sigma^2}{\mu^2} \to 0,$$

proving that almost no graphs in $\mathcal{G}(n, p)$ has distance two with $p = p_\ell$.
□

The further solution for diameter of random graphs is as follows. Let *diam*(G) be the diameter of G and let $d \geq 2$ be an integer. If

$$p = n^{1/d-1}(\log(n^2/x))^{1/d},$$

then

$$\Pr(\text{diam}(G_p) = d) \to e^{-x/2}$$

and

$$\Pr(\text{diam}(G_p) = d + 1) \to 1 - e^{-x/2}$$

an $n \to \infty$. See Bollobás (2001) for details. The above limit distribution implies that

$$p_{\ell,u} = n^{1/d-1}(2\log n \pm \omega(n))^{1/d}$$

are ltf and utf of graphs being distance d, respectively.

The distance two graphs are of interest in graph Ramsey theory. Let G_n be a Ramsey graph of order $n = r(3,k) - 1$. We suppose that G_n is edge maximal for triangle-freeness. Then G_n is a distance two graph. Since the order of n is $k^2/\log k$, so the maximum degree of G_n is bounded above by $k \le c\sqrt{n \log n}$. It is likely that the minimum degree of G_n has order $\sqrt{n \log n}$ hence the order of its edge density is $\sqrt{\log n/n}$ as that in above theorem.

Problem 3.1 *Let G_n be a Ramsey graph of order $n = r(3,k) - 1$ that is edge maximal. Determine the orders of the minimum and maximum degrees of G_n as $k \to \infty$.*

The following result gives threshold functions for the property of being connected. The deeper version of the result will be in the next section.

Theorem 3.5 *Let $\omega(n) \to \infty$ be a function with $\omega(n) < \log n$. Set*

$$p_\ell = \frac{\log n - \omega(n)}{n} \quad \textit{and} \quad p_u = \frac{\log n + \omega(n)}{n}.$$

Then p_ℓ and p_u are ltf and utf for graph in $\mathcal{G}(n,p)$ with the property of being connected, respectively.

Proof. Let Q be the family of connected graphs. Since Q is monotone increasing, we may assume that $\omega(n) \le \log\log n$ without loss of generality by Lemma 3.1. Let $X_k = X_k(G)$ be the number of components of $G \in \mathcal{G}(n,p)$ that have exactly k vertices.

(i) We first prove that $p = p_\ell$ is a ltf for Q. Set $\mu = E(X_1)$. Note that $(1-p)^n \sim e^{-np}$ since $np^2 \to 0$, we have

$$\begin{aligned} \mu &= E(X_1) = n(1-p)^{n-1} \\ &\sim n(1-p)^n \sim ne^{-np} = e^{\omega(n)} \to \infty. \end{aligned}$$

This may indicate that $\Pr(X_1 = 0) \to 0$ and $\Pr(X_1 \ge 1) \to 1$, so a.a. graphs have isolated vertices hence they are disconnected. In order to use the Second Moment Method, we need to estimate the variance $\sigma^2 = \sigma^2(X_1)$ hence $E(X_1^2)$. We first have

$$E[X_1(X_1 - 1)] = n(n-1)(1-p)^{2n-3},$$

which is the expected number of order pairs of isolated vertices. There are $n(n-1)$ ordered pairs of vertices, and such a pair are isolated if and only if they are neither adjacent each other nor adjacent to any other $n-2$ vertices, which count $2n-3$ edges. Then

$$\begin{aligned} E(X_1^2) &= E[X_1(X_1-1)] + E(X_1) \\ &= \mu + n(n-1)(1-p)^{2n-3}. \end{aligned}$$

We thus have

$$\begin{aligned} \sigma^2 &= \sigma^2(X_1) = E[(X_1-\mu)^2] = E(X_1^2) - \mu^2 \\ &= \mu + n(n-1)(1-p)^{2n-3} - n^2(1-p)^{2n-2} \\ &\le \mu + pn^2(1-p)^{2n-3}. \end{aligned}$$

Since $p = (\log n - \omega(n))/n$ with $\log\log n \ge \omega(n) \to \infty$ and $1-p \le e^{-p}$,

$$\begin{aligned} pn^2(1-p)^{2n-3} &\le (1+o(1))(\log n)ne^{-2np} \\ &= (1+o(1))(\log n)ne^{-2\log n + 2\omega(n)} \\ &= (1+o(1))\frac{\log n}{n}e^{2\omega(n)} \to 0. \end{aligned}$$

We thus obtain

$$\sigma^2 = \sigma^2(X_1) \le \mu + 1.$$

This and the Second Moment Method give

$$\Pr(G_p \in Q) \le \Pr(X_1 = 0) \le \Pr(|X_1 - \mu| \ge \mu) \le \frac{\sigma^2}{\mu^2} \to 0,$$

proving that p_ℓ is a ltf for property of being connected.

(ii) Now let $p = p_u = (\log n + \omega(n))/n$. Note that if G is not connected then it must contains a component of order at most $\lfloor n/2 \rfloor$. So

$$\Pr(G_p \notin Q) = \Pr\left(\sum_{k=1}^{\lfloor n/2 \rfloor} X_k \ge 1\right) \le E\left(\sum_{k=1}^{\lfloor n/2 \rfloor} X_k\right) = \sum_{k=1}^{\lfloor n/2 \rfloor} E(X_k).$$

Since if a set with k vertices induces a component, then any vertex in it is not adjacent to any vertex out of it. Thus

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

where we ignore condition that the set is connected. Therefore,

$$\Pr(G_p \notin Q) \leq \sum_{k=1}^{\lfloor n/2 \rfloor} \binom{n}{k} (1-p)^{k(n-k)}.$$

Let us split the sum into two parts S_1 and S_2. Note that $e^{kp} \leq 1 + \epsilon$ uniformly for $k \leq n^{3/4}$, and $ne^{-np} = e^{-\omega(n)/n}$, we have

$$\begin{aligned} S_1 &= \sum_{1 \leq k \leq n^{3/4}} \binom{n}{k} (1-p)^{k(n-k)} \leq \sum_{1 \leq k \leq n^{3/4}} \left(\frac{en}{k} e^{-np} e^{kp} \right)^k \\ &\leq \sum_{1 \leq k \leq n^{3/4}} \left((1+\epsilon) \frac{e^{1-\omega(n)}}{k} \right)^k \leq \sum_{1 \leq k \leq n^{3/4}} \left((1+\epsilon) e^{1-\omega(n)} \right)^k \\ &\leq (1+o(1))(1+\epsilon) e^{1-\omega(n)} \to 0. \end{aligned}$$

Note that for $n^{3/4} < k \leq n/2$, we have $\binom{n}{k} \leq (en/k)^k \leq (en^{1/4})^k$ and

$$(1-p)^{k(n-k)} \leq (1-p)^{kn/2} \leq e^{-knp/2} < \frac{1}{n^{k/2}},$$

hence

$$\begin{aligned} S_2 &= \sum_{n^{3/4} < k \leq \lfloor n/2 \rfloor} \binom{n}{k} (1-p)^{k(n-k)} \\ &\leq \sum_{n^{3/4} < k \leq n/2} \left(\frac{e}{n^{1/4}} \right)^k \leq (1+o(1)) \left(\frac{e}{n^{1/4}} \right)^{n^{3/4}} \to 0. \end{aligned}$$

Thus $S_1 + S_2 \to 0$, proving that a.a. graphs in $\mathcal{G}(n,p)$ are connected.
□

3.3 Poisson limit

In probability theory, we call a random variable X to have *Poisson distribution* if it takes non-negative integral values and $\Pr(X = k) = \frac{\mu^k}{k!} e^{-\mu}$ for some constant $\mu > 0$, which is the expectation of X (and the variance of X). An elementary fact is that if $X = \sum_{i=1}^n X_i \sim B(n,p)$ and $np \to \mu$ as $n \to \infty$, then $\Pr(X = k) \to \frac{\mu^k}{k!}e^{-\mu}$. This is because for fixed k

$$\begin{aligned}\Pr(X = k) &= \binom{n}{k}p^k(1-p)^{n-k} \sim \binom{n}{k}p^k(1-p)^n \\ &\sim \frac{n^k}{k!}p^k e^{-np} \to \frac{\mu^k}{k!}e^{-\mu}.\end{aligned}$$

In the last section, in order to show that a.a graphs in $\mathcal{G}(n,p)$ are disconnected for $p = (\log n - \omega(n))/n$, we in fact have proved that a.a. graphs in $\mathcal{G}(n,p)$ have isolated vertices. Let X be the number of isolated vertices in G_p of $\mathcal{G}(n,p)$, where $p = (\log n + x)/n$, then $X = \sum_{i=1}^n X_i$, where X_i is indicator of ith vertex being isolated. Define $p' = \Pr(X_i = 1)$. Then

$$p' = (1-p)^{n-1} \sim \exp(-np) = \frac{e^{-x}}{n} = \frac{\mu}{n},$$

where $\mu = e^{-x}$, so $np' \to \mu$. The distribution of X is close to $B(n,p')$ in the sense that $X_1, X_2, \cdots, X_n$ are "almost" mutually independent, so we are expecting that X has limit Poisson distribution.

The approach to the Poisson paradigm introduced in this section is called *Brun's sieve* for its user T. Brun in number theory. Let us begin with a basic identity called *inclusion-exclusion formula*.

In a probability space Ω, let $X_1, X_2, \ldots, X_\ell$ be 0-1 random variables and set

$$X = X_1 + X_2 + \cdots + X_\ell.$$

As usual, denote by $[\ell]$ for $\{1, 2, \ldots, \ell\}$. Define $S_0 = 1$ and

$$S_r = \sum_{[\ell]^{(r)}} \Pr(X_{i_1}X_{i_2}\ldots X_{i_r} = 1),$$

where the sum is taken over elements $\{i_1, i_2, \ldots, i_r\} \in [\ell]^{(r)}$, the family of all $r-$subsets of $[\ell]$. Note the elements $\omega \in \Omega$ such that $X_{i_1}X_{i_2}\cdots X_{i_r} = 1$ are what such that $X_{i_1} = 1, X_{i_2} = 1, \ldots, X_{i_r} = 1$, and $S_r = 0$ for $r > \ell$. For general r,

$$S_r = \sum_{\omega\in\Omega} \binom{X(\omega)}{r} \Pr(\omega)$$

as an element ω of the sample space for which $X(\omega) = t$ contributes to $\binom{t}{r}$ of the terms defining S_r. Here and in what follow, we write the formulas appropriate for finite sample space. Following standard notation, we define falling factorials by $(X)_0 = 1$ and

$$(X)_r = X(X-1)\cdots(X-r+1).$$

Then

$$S_r = \sum_{\omega\in\Omega} \frac{(X)_r}{r!} \Pr(\omega) = \frac{E((X)_r)}{r!}.$$

The quantity $E((X)_r)$ is called the rth *factorial moment* of X.

Theorem 3.6 (Inclusion-Exclusion Formula) *For each integer $k \geq 0$,*

$$\Pr(X = k) = \sum_{r\geq 0} (-1)^r \binom{k+r}{r} S_{k+r}.$$

Moreover, for each integer $m \geq 0$,

$$\sum_{r=0}^{2m-1} (-1)^r \binom{k+r}{r} S_{k+r} \leq \Pr(X = k) \leq \sum_{r=0}^{2m} (-1)^r \binom{k+r}{r} S_{k+r}.$$

Proof. It is easy to see

$$\begin{aligned} \Pr(X = 0) &= \Pr(X_1 = 0, \ldots, X_\ell = 0) \\ &= 1 - \Pr(\exists\, i, X_i = 1) = S_0 - S_1 + S_2 - \cdots. \end{aligned}$$

For general k, using

$$S_{k+r} = \sum_{\omega\in\Omega} \binom{X(\omega)}{k+r} \Pr(\omega),$$

and interchanging orders of the summation, we obtain

$$\sum_{r\geq 0} (-1)^r \binom{k+r}{r} S_{k+r} = \sum_{\omega\in\Omega} \left\{ \sum_{r\geq 0} (-1)^r \binom{k+r}{r} \binom{X(\omega)}{k+r} \right\} \Pr(\omega).$$

For a fixed ω hence fixed $X = X(\omega)$, note that

$$\sum_{r\geq 0} (-1)^r \binom{k+r}{r} \binom{X}{k+r} = \binom{X}{k} \sum_{r\geq 0} (-1)^r \binom{X-k}{r}.$$

If $X < k$, all terms vanish. If $X = k$, then one term ($r = 0$) contributes and the sum is 1. Finally, if $X > k$, the sum vanishes since

$$\sum_{r\geq 0}(-1)^r\binom{X-k}{r} = (1-1)^{X-k} = 0.$$

Thus

$$\sum_{r\geq 0}(-1)^r\binom{k+r}{r}\binom{X}{k+r} = \begin{cases} 1 & \text{if } X = k \\ 0 & \text{otherwise,} \end{cases}$$

and

$$\sum_{r\geq 0}(-1)^r\binom{k+r}{r}S_{k+r} = \sum_{\omega\in\Omega}\delta_{k\,X(\omega)}\Pr(\omega) = \Pr(X=k),$$

where δ_{ij} is the Kronecker delta. To verify the second result, note for all $s \geq 0$ and $n \geq 1$,

$$\sum_{r=0}^{s}(-1)^r\binom{n}{r} = (-1)^s\binom{n-1}{s},$$

which can be proved easily by induction on s, hence

$$\begin{aligned}\sum_{r=0}^{s}(-1)^r\binom{k+r}{r}\binom{X}{k+r} &= \binom{X}{k}\sum_{r=0}^{s}(-1)^r\binom{X-k}{r} \\ &= \begin{cases} 0 & \text{if } \ X < k \\ 1 & \text{if } \ X = k \\ (-1)^s\binom{X}{k}\binom{X-k-1}{s} & \text{if } \ X > k. \end{cases}\end{aligned}$$

We thus have

$$\begin{aligned}&\sum_{r=0}^{s}(-1)^r\binom{k+r}{r}S_{k+r} \\ = \ &\sum_{\omega\in\Omega}\left\{\sum_{r=0}^{s}(-1)^r\binom{k+r}{r}\binom{X(\omega)}{k+r}\right\}\Pr(\omega) \\ = \ &\sum_{X(\omega)=k}\Pr(\omega) + \sum_{X(\omega)>k}(-1)^s\binom{X}{k}\binom{X-k-1}{s}\Pr(\omega) \\ = \ &\Pr(X=k) + (-1)^s\sum_{X(\omega)>k}\binom{X}{k}\binom{X-k-1}{s}\Pr(\omega).\end{aligned}$$

In the last line elements ω such that $X(\omega) > k$ make a positive or negative contribution depending on whether s is even or odd. □

Suppose that we have defined a sequence of probability spaces and that in the space $\Omega = \Omega_n$ we have the preceding situation with $\ell = \ell(n)$. If $E((X)_r) \to \mu^r$ as $n \to \infty$, we can make a precise statement about the limiting distribution of $X = \sum_{i=1}^{\ell} X_i$.

Theorem 3.7 (Poisson Limit) *Suppose that there is a positive number μ such that*

$$\lim_{n\to\infty} S_r = \frac{\mu^r}{r!},$$

equivalently $\lim_{n\to\infty} E((X)_r)) = \mu^r$*, for each fixed integer $r \geq 0$. Then*

$$\lim_{n\to\infty} \Pr(X = k) = \frac{\mu^k}{k!} e^{-\mu}.$$

Namely, the limiting distribution of X is Poisson with mean μ.

Proof. Refer to the inequalities in the last theorem we have

$$\frac{1}{k!} \sum_{r=0}^{2m-1} (-1)^r \frac{E((X)_{k+r})}{r!} \leq \Pr(X = k) \leq \frac{1}{k!} \sum_{r=0}^{2m} (-1)^r \frac{E((X)_{k+r})}{r!}.$$

Note that if m is fixed, we can make the limit (as $n \to \infty$) term by term to get

$$\frac{1}{k!} \sum_{r=0}^{2m-1} (-1)^r \frac{\mu^{k+r}}{r!} \leq \lim_{n\to\infty} \Pr(X = k) \leq \frac{1}{k!} \sum_{r=0}^{2m} (-1)^r \frac{\mu^{k+r}}{r!}.$$

Since m is arbitrary, the result follows. □

Theorem 3.8 *For any fixed real number x, let*

$$p = \frac{\log n + x}{n},$$

and let X be the number of isolated vertices in a graph of $\mathcal{G}(n, p)$, then

$$\lim_{n\to\infty} \Pr[X = k] = \frac{\mu^k}{k!} e^{-\mu},$$

where $\mu = e^{-x}$. In particular, the limiting probability that graph in $\mathcal{G}(n, p)$ has no isolated vertices is $\exp(-e^{-x})$*.*

Proof. Define X_i as an indicator that the vertex i is an isolated vertex, and define $X = \sum_{i=1}^{n} X_i$. Then X counts the number of isolated vertices in G_p and

$$S_1 = E(X) = n(1-p)^{n-1} \to e^{-x}$$

as $n \to \infty$. More generally,

$$S_r = \binom{n}{r}(1-p)^{r(n-r)+\binom{r}{2}} \sim \frac{n^r}{r!}(1-p)^{rn} \to \frac{\mu^r}{r!},$$

where $\mu = e^{-x}$. The limiting distribution of X follows from Poisson limit theorem as desired. □

Corollary 3.1 *Suppose* $\log n \geq \omega(n) \to \infty$. *Then* $p_\ell = \frac{\log n - \omega(n)}{n}$ *and* $p_u = \frac{\log n + \omega(n)}{n}$ *are ltf and utf for graph in* $\mathcal{G}(n,p)$ *having no isolated vertices, respectively.* □

We shall show that for the same $p = (\log n + x)/n$,

$$\lim_n \Pr(G_p \text{ has no isolates}) = \lim_n \Pr(G_p \text{ is connected}) = \exp(-e^{-x}).$$

So in almost every graph when the last isolated vertex disappears, the graph G_p becomes connected in the evolution of random graph as x increases. Sightly before it is connected, a giant component with only a bounded number vertices outside has formed. In fact, the giant component are formed by larger components and the smaller components have bigger chances to survive.

Theorem 3.9 *For any fixed real number* x*, let*

$$p = \frac{\log n + x}{n},$$

and let A *denote the event that outside of at most one non-trivial component, all vertices are isolated. Then*

$$\lim_{n\to\infty} \Pr(A) = 1$$

and

$$\lim_{n\to\infty} \Pr[G_p \textit{ is connected}] = \exp(-e^{-x}).$$

Proof. We begin by identifying the following events in $\mathcal{G}(n,p)$.

A: Outside of at most one non-trivial component, G_p has only isolated vertices.

B: G_p has no isolated vertices.

C: G_p is connected.

Then $C = A \cap B$ and

$$\Pr(B) = \Pr(C) + \Pr(\bar{A} \cap B).$$

To prove that $\Pr(C) \to \exp(-e^{-x})$ as $n \to \infty$, it suffices to show that $\Pr(\bar{A}) \to 0$ since we have known that $\Pr(B) \to \exp(-e^{-x})$. Let $X \subseteq [n]$ be the vertex set of the largest component of G and let $Y = V \setminus X$. We do not distinguish a vertex set and the subgraph induced by this set if no danger of confusion. If $\bar{A}$ holds, then for some $X \subseteq [n]$ with $|X| \geq 2$,

1. X is connected;
2. Y contains edge;
3. There is no $X - Y$ edges.

Note that these events are independent. Let the probabilities of the events 1, 2 and 3 be denoted P_X, P_Y and P_{XY}, respectively and let $|X| = k$ and $|Y| = m = n - k$. By distinguishing that $k \leq \lfloor n/2 \rfloor$ or $m \leq \lfloor n/2 \rfloor$, we have

$$\Pr(\bar{A}) \leq \sum_{k=2}^{\lfloor n/2 \rfloor} \binom{n}{k} P_X P_{XY} + \sum_{m=2}^{\lfloor n/2 \rfloor} \binom{n}{m} P_Y P_{XY}. \tag{3.1}$$

To bound $\Pr(\bar{A})$, we use the following facts:

1. $P_X \leq k^{k-2} p^{k-1}$;
2. $P_Y = 1 - (1-p)^{\binom{m}{2}}$
3. $P_{XY} = (1-p)^{mk}$.

The first fact follows since X must contain one of the k^{k-2} possible spanning trees. Consider then first term on the right hand side of (3.1), we have

$$\sum_{k=2}^{\lfloor n/2\rfloor}\binom{n}{k}P_X P_{XY} \le \sum_{k=2}^{\lfloor n/2\rfloor}\binom{n}{k}k^{k-2}p^{k-1}(1-p)^{k(n-k)}.$$

The term corresponding to any fixed $k \ge 2$ can be bounded above as

$$c_1 n^k p^{k-1}(1-p)^{kn} \le c_1 n^k p^{k-1} e^{-knp} = c_1 e^{-kx} p^{k-1} \to 0,$$

where c_1 henceforth c_i are positive constants. For any $k \le n/2$,

$$(1-p)^{n-k} \le e^{-(n-k)p} \le e^{-np/2} = \frac{e^{-x/2}}{\sqrt{n}}.$$

Thus

$$\binom{n}{k}k^{k-2}p^{k-1}(1-p)^{k(n-k)} \le \left(\frac{en}{k}\right)^k k^{k-2}p^{k-1}\left(\frac{e^{-x/2}}{\sqrt{n}}\right)^k$$
$$= \frac{1}{k^2 p}\left(enp\frac{e^{-x/2}}{\sqrt{n}}\right)^k \le \frac{n}{\log n + x}\left(\frac{c_2\log n}{\sqrt{n}}\right)^k.$$

It follows that

$$\sum_{k=2}^{\lfloor n/2\rfloor}\binom{n}{k}P_X P_{XY} \le o(1) + \sum_{k\ge 4}\frac{n}{\log n + x}\left(\frac{c_2\log n}{\sqrt{n}}\right)^k \to 0$$

as $n \to \infty$. Now setting

$$K = \lfloor \sqrt[4]{n} \rfloor, \quad \text{and} \quad M = \lceil 2\sqrt{n}\exp(1-x/2)\rceil,$$

we shall separate the second term on the right hand side of (3.1) into three parts by K and M. Using the facts that

$$P_{XY} = (1-p)^{m(n-m)} \le e^{-m(n-m)p} \le e^{-mnp/2} = \left(\frac{e^{-x/2}}{\sqrt{n}}\right)^m$$

for $m \le \lfloor n/2 \rfloor$ and $\binom{n}{m} \le (en/m)^m$, we have

$$\begin{aligned}
\sum_{m=M}^{\lfloor n/2 \rfloor} \binom{n}{m} P_Y P_{XY} &\le \sum_{m=M}^{\lfloor n/2 \rfloor} \binom{n}{m} P_{XY} \le \sum_{m \ge M} \left(\frac{ene^{-x/2}}{m\sqrt{n}} \right)^m \\
&\le \sum_{m \ge M} \frac{1}{2^m} = \frac{1}{2^{M-1}} \to 0
\end{aligned}$$

since $M \to \infty$ as $n \to \infty$. On the other hand, if $m < M$, we have $e^{mp} \le 2$ for all large n since $mp \le Mp \to 0$, so

$$P_{XY} = (1-p)^{m(n-m)} \le \left(e^{-np} e^{mp} \right)^m \le \left(\frac{2e^{-x}}{n} \right)^m .$$

It follows that

$$\sum_{m=K}^{M-1} \binom{n}{m} P_Y P_{XY} \le \sum_{m=K}^{M-1} \binom{n}{m} P_{XY} \le \sum_{m \ge K} \frac{(2e^{-x})^m}{m!} \to 0$$

since the sum in the last line is the tail of a convergent series.

Finally, for all $m < K$,

$$P_Y \le 1 - (1-p)^{\binom{K}{2}},$$

which tends to zero uniformly on $m < M$, and it follows that

$$\sum_{m=2}^{K-1} \binom{n}{m} P_Y P_{XY} = o\left(\sum_{m=2}^{K-1} \binom{n}{m} P_{XY} \right) \le o\left(\sum_{m \ge 2} \frac{(2e^{-x})^m}{m!} \right),$$

which tends to zero. Combining these results, we find that $\Pr(\bar{A}) \to 0$, completing the proof. □

Corollary 3.2 *Suppose that* $\log n \ge \omega(n) \to \infty$*. Then* $p_\ell = \frac{\log n - \omega(n)}{n}$ *and* $p_u = \frac{\log n + \omega(n)}{n}$ *are ltf and utf for graph in* $\mathcal{G}(n,p)$ *of being connected, respectively.*

□

3.4 References

N. Alon and J. Spencer, *The Probabilistic Method, 3rd ed.*, Wiley-Interscience, New York, 2008.

B. Bollobás, *Random Graphs, 2nd Edition*, Cambridge University Press, London, 2001.

P. Erdős, Graph theory and probability, *Canad. J. Math.*, **11** (1959), 34-38.

P. Erdős and A. Rényi, On the evolution of random graphs, *Publ. Math. Inst. Hungar. Acad. Sci.*, **5** (1960), 17-61.

S. Jason, T. Łuczak, and A. Ruciński, *Random Graphs*, Wiley-Interscience, New York, 2000.

Chapter 4

Quasi-random Graphs

Random graphs have been proven to be one of most important tools in modern graph theory. Their tremendous triumph raises the following general question: what are the essential properties and how can we tell when a given graph behaves like a random graph G_p in $\mathcal{G}(n,p)$? Here a typical property of random graphs is that almost all G_p satisfy. This leads us to a concept of *quasi-random graphs*. It was Thomason (1987) who introduced the notation of jumbled graphs to measure the similarity between the edge distribution of quasi-random graphs and random graphs. Quasi-random graphs are also called pseudo-graphs. A cornerstone contribution of Chung, Graham and Wilson (1989) showed that many properties of different nature are equivalent to the notation of quasi-random graphs. For a survey on quasi-random graphs, see Krivelevich and Sudakov (2006). This chapter focuses on quasi-random graphs. In recent years, there are some quasi-random families of graphs appearing, which are all constructed by finite fields. Their algebraic parameters are easier to compute, some of which are related to characters of finite fields and thus the third section is devoted to the topics. The last section is application for quasi-random graphs in Ramsey theory.

4.1 Properties of dense graphs

Speaking formally, a quasi-random G of order n is a graph that behaves like a random graph $G(n, p)$ with $p = e(G)/\binom{n}{2}$. For $0 < p < 1 \le \alpha$, a graph G is called (p, α)-*jumbled* if each induced subgraph H on h vertices of G satisfies that

$$\left| e(H) - p\binom{h}{2} \right| \le \alpha h.$$

Equivalently, G is (p, α)-jumbled if the average degree $d(H)$ of each induced subgraph H of G satisfies that

$$|d(H) - p(h-1)| \le 2\alpha.$$

The following result of Thomason (1987) contains a simple local condition of a graph of being jumbled.

Theorem 4.1 *Let G be a graph of order n with $\delta(G) \ge pn$. If any pair of vertices has at most $p^2 n + \ell$ common neighbors, where $\ell > 0$, then G is $(p, \sqrt{(p+\ell)n}/2)$-jumbled.*

Proof. Let H be an induced subgraph of G of order h with $d(H) = d$, where $h < n$. Write $V(G) = \{v_1, v_2, \ldots, v_n\}$ and $V(H) = \{v_1, v_2, \ldots, v_h\}$, say. Let d_i be the number of neighbors of v_i in H for $1 \le i \le n$. Then $\sum_{i=1}^{h} d_i = hd$ and

$$\sum_{j=h+1}^{n} d_j \ge \sum_{i=1}^{h} (pn - d_i) = h(pn - d).$$

Since any pair of vertices are covered by at most $p^2 n + \ell$ vertices, and at most that in H particularly, we have

$$\sum_{i=1}^{n} \binom{d_i}{2} \le \binom{h}{2}(p^2 n + \ell).$$

The above and the convexity of the function $\binom{x}{2}$ imply that

$$h\binom{d}{2} + (n-h)\binom{h(pn-d)/(n-h)}{2} \le \binom{h}{2}(p^2 n + \ell).$$

Equivalently,

$$(d - ph)^2 \le \frac{n-h}{n}\Big[(h-1)\ell + p(1-p)n\Big],$$

which gives that

$$|d - p(h-1)| \le \sqrt{(p+\ell)n}$$

as claimed. Finally, note that the same inequality holds for $h = n$. □

For given graphs G and H, let $N_G^*(H)$ be the number of labeled occurrences of H as an induced subgraph of G, which is the number of adjacency-preserving injections from $V(H)$ to $V(G)$ whose image is the set of vertices of an induced copy of H of G. Namely, these injections are both adjacency-preserving and non-adjacency-preserving. Let $N_G(H)$ be the number of labeled copies of H as a (not necessarily induced) subgraph of G. Then

$$N_G(H) = \sum_{H'} N_G^*(H'),$$

where H' ranges over all graphs on $V(H)$ obtained from H by adding a set of edges. For example, if $G = H = C_t$, then $N_G^*(H) = N_G(H) = 2t$, and if $G = K_n$ and $n \ge t \ge 4$, then $N_G^*(C_t) = 0$ and $N_G(C_t) = N_G^*(K_t) = (n)_t$. If $G = K_{n/2,n/2}$ and n is even, then $N_G(C_4) = 2\Big(\frac{n}{2}(\frac{n}{2} - 1)\Big)^2 \sim 2 \cdot \left(\frac{n}{2}\right)^4$ for large n.

Let G be a (p,α)-jumbled graph of order n, where $\alpha = \alpha_n = o(n)$ as $n \to \infty$. Then, as shown by Thomason, for fixed p and fixed graph H of order h

$$N_G^*(H) \sim p^{e(H)}(1-p)^{\binom{h}{2}-e(H)}n^h.$$

Let x and y be vertices of G. Denote by $s(x,y)$ the number of vertices of G adjacent to x and y the same way: either to both or none. Let λ_i be eigenvalues of G with $|\lambda_1| \ge |\lambda_2| \ge \cdots \ge |\lambda_n|$. Let $\lambda = \lambda(G) = |\lambda_2|$. For two (not necessarily disjoint) subsets B and C, let $e(B,C)$ denote the number of edges from B to C, in which each edge in $B \cap C$ is counted twice. If $B \cap C = \emptyset$, then $e(B,C)$ is simply the number of edges between B and C.

The *quasi-random graph* defined by Chung, Graham and Wilson is in fact a family of simple graphs, which satisfy any (hence all) of equivalent properties in the following theorem. It is remarkable that these properties ignore "small" local structures. The expressions of the properties are related to the edge density p, here $p = 1/2$.

Theorem 4.2 *Let $\{G\}$ be a sequence of graphs, where $G = G_n$ is a graph of order n. Then the following properties are equivalent:*

$P_1(h)$**:** *For any fixed $h \geq 4$ and graph H of order h, $N^*_G(H) \sim (\frac{1}{2})^{\binom{h}{2}} n^h$.*

$P_2(t)$**:** $e(G) \sim \frac{n^2}{4}$ *and* $N_G(C_t) \leq (\frac{n}{2})^t + o(n^t)$ *for any even* $t \geq 4$.

P_3**:** $e(G) \geq \frac{n^2}{4} + o(n^2)$, $\lambda_1 \sim \frac{n}{2}$ *and* $\lambda_2 = o(n)$.

P_4**:** *For each* $U \subseteq V(G)$, $e(U) = \frac{1}{2}\binom{|U|}{2} + o(n^2)$.

P_5**:** *For each* $U \subseteq V(G)$ *with* $|U| = \lfloor \frac{n}{2} \rfloor$, $e(U) \sim \frac{n^2}{16}$.

P_6**:** $\sum_{x,y} \left| s(x,y) - \frac{n}{2} \right| = o(n^3)$.

P_7**:** $\sum_{x,y} \left| |N(x) \cap N(y)| - \frac{n}{4} \right| = o(n^3)$.

Proof.* The steps of proof of Chung, Graham and Wilso are $P_1(h+1) \Rightarrow P_1(h)$ and

$$P_1(2h) \Rightarrow P_2(2t) \Rightarrow P_2(4) \Rightarrow P_3 \Rightarrow P_4 \Longleftrightarrow P_5 \Rightarrow P_6 \Rightarrow P_1(2h),$$

so that all but P_7 are proven to be equivalent. They then add P_7 to the equivalent chain by proving that

$$P_2(t) \Rightarrow P_7 \Rightarrow P_6.$$

Here, we omit some steps but keep most of them and preserve the typical flavor.

Fact 1. $P_1(h+1) \Rightarrow P_1(h)$, and $P_1(3)$ implies the property

$$P_0 : \sum_v \left| \deg(v) - \frac{n}{2} \right| = o(n^2).$$

Let us remark that P_0 is equivalent to that

$$P_0' : \text{ All but } o(n) \text{ vertices of } G \text{ have degree } (1 + o(n))\tfrac{n}{2}$$

by Cauchy-Schwarz inequality, and P_0 implies that

$$e(G) \sim \frac{n^2}{4}.$$

Assume that $P_1(h+1)$ holds. Let H be a graph of order h. There are 2^h ways to extend it to a graph H' of order $h+1$, and each copy of H is contained in $n-h$ subgraphs H' of order $h+1$. By $P_1(h+1)$, we have

$$N_G^*(H') \sim n^{h+1} 2^{-\binom{h+1}{2}},$$

thus

$$N_G^*(H) \sim n^{h+1} 2^{-\binom{h+1}{2}} \frac{2^h}{n-h} \sim n^h 2^{-\binom{h}{2}},$$

which is $P_1(h)$. Suppose that $\{G\}$ satisfies $P_1(3)$. Let H_i be the graph of order 3 and i edges, $1 \le i \le 3$. By counting how often each edge can contribute to the various $N_G^*(H_i)$, we have

$$(n-2)\sum_v \deg(v) = N_G^*(H_1) + 2N_G^*(H_2) + N_G^*(H_3) \sim \frac{n^3}{2},$$

thus $\sum_v \deg(v) \sim \frac{n^2}{2}$ and $e(G) \sim \frac{n^2}{4}$. Also

$$\sum_v \deg(v)(\deg(v)-1) = N_G^*(H_2) + N_G^*(H_3) \sim \frac{n^3}{4},$$

implying that $\sum_v \deg^2(v) \sim \frac{n^3}{4}$. Then, by Cauchy-Schwarz,

$$\begin{aligned} \sum_v \Big|\deg(v) - \frac{n}{2}\Big| &\le \sqrt{n}\Big(\sum_v \Big|\deg(v) - \frac{n}{2}\Big|^2\Big)^{1/2} \\ &= \sqrt{n}\Big(\sum_v \deg^2(v) - n\sum_v \deg(v) + \frac{n^3}{4}\Big)^{1/2}, \end{aligned}$$

which is $o(n^2)$.

Fact 2. $P_1(2t) \Rightarrow P_2(2t)$ $(t \ge 2)$. Fact 1 has proved that $e(G) \sim \frac{n^2}{4}$. We then show that

$$N_G(C_{2t}) = \sum_{H'} N_G^*(C_{2t}) \le (1+o(1))\Big(\frac{n}{2}\Big)^{2t}.$$

As H' ranges over all graphs on $V(H)$ obtained from H by adding to it a set of edges, the number of such sets is $2^{\binom{2t}{2}-2t}$. This and $P_1(2t)$ imply $P_2(2t)$.

Fact 3. $P_2(2t) \Rightarrow P_2(4) \Rightarrow P_3$. There is nothing to prove for the first implication and we prove the second. Let A be the adjacency matrix of G and d the average degree of G. We first claim that

$$\lambda_1 \geq d.$$

Let us verify that for any unit vector X, $\lambda_1 \geq X^t AX$. Let Λ be the diagonal matrix with diagonal entries $\lambda_1, \lambda_2, \ldots, \lambda_n$ and P a normal orthogonal matrix such that $PAP^t = \Lambda$. Then PX is a unit vector, and

$$\lambda_1 = \lambda_1 (PX)^t \cdot (PX) \geq (PX)^t \Lambda (PX) = X^t (P^t \Lambda P) X = X^t AX.$$

By taking $X = \frac{1}{\sqrt{n}} J$, where $J = (1, 1, \ldots, 1)^t$, we obtain that

$$\lambda_1 \geq \frac{1}{n} J^t AJ = \frac{1}{n} \sum_v \deg(v) = d$$

as claimed. This and $e(G) \sim \frac{n^2}{2}$ imply $\lambda_1 \geq \frac{n}{2} + o(n)$. Next, consider the trace of A^4. Clearly,

$$tr(A^4) = \sum_{i=1}^{n} \lambda_i^4 \geq \lambda_1^4 \geq (1 + o(1)) \frac{n^4}{16}.$$

On the other hand, as this trace is precisely the number of labeled and closed walks of length 4 in G, i.e., the number of sequences $v_0, v_1, v_2, v_3, v_4 = v_0$ such that $v_i v_{i+1}$ is an edge. This number is $N_G(C_4)$ plus the number of such sequences in which $v_2 = v_0$, and plus the number of such sequences in which $v_2 \neq v_0$. Thus

$$\sum_{i=1}^{n} \lambda_i^4 = N_G(C_4) + o(n^4) \sim \left(\frac{n}{2}\right)^4.$$

It follows that $tr(A^4) \sim \frac{n^4}{16}$, thus $\lambda_1 \sim \frac{n}{2}$ and $\sum_{i=2}^{n} \lambda_i^4 = o(n^4)$ hence $\lambda_2 = o(n)$ as desired.

Fact 4. $P_3 \Rightarrow P_4$. To simply the proof, we suppose that G is regular. Then the Fact 4 follows from Corollary 4.2 in the next section.

Fact 5. $P_4 \Longleftrightarrow P_5$. The implication $P_4 \Rightarrow P_5$ is immediate, so we show $P_5 \Rightarrow P_4$. By ignoring one vertex possibly, we assume that n is even so that $n/2$ is an integer. Suppose that for any subset S with $|S| = n/2$, $\left|e(S) - \frac{n^2}{16}\right| < \epsilon n^2$, where $\epsilon > 0$ is fixed. We shall show that for any subset T,

$$\left|e(T) - \frac{1}{2}\binom{t}{2}\right| < 20\epsilon n^2,$$

where $t = |T|$. Let us consider two cases, .

Case 1. $t = |T| \geq n/2$. By averaging over all $S \subseteq T$ with $|S| = n/2$, we have

$$e(T) = \frac{1}{\binom{t-2}{n/2-2}} \sum \left\{e(S) : \; S \subseteq T, |S| = n/2\right\}$$

as each edge is counted exactly $\binom{t-2}{n/2-2}$ times. Thus

$$e(T) \leq \frac{\binom{t}{n/2}}{\binom{t-2}{n/2-2}} \Big(\frac{n^2}{16} + \epsilon n^2\Big) \leq \binom{t}{2}\Big(\frac{1}{2} + 9\epsilon\Big).$$

Similarly,

$$e(T) \geq \binom{t}{2}\Big(\frac{1}{2} - 9\epsilon\Big).$$

Case 2. $t = |T| < n/2$. We shall show that the assumption

$$e(T) \geq \frac{1}{2}\binom{t}{2} + 20\epsilon n^2$$

leads to a contradiction. Set $\bar{T} = V \setminus T$. Then $|\bar{T}| = n - t > n/2$ and by Case 1, we have

$$\binom{n-t}{2}\Big(\frac{1}{2} - 9\epsilon\Big) < e(\bar{T}) < \binom{n-t}{2}\Big(\frac{1}{2} + 9\epsilon\Big).$$

Consider the average value A of $e(T \cup T')$, where T' ranges over all subsets of $\bar{T}$ with $|T'| = n/2 - t$ so that $|T \cup T'| = n/2$, so

$$A = \binom{n-t}{n/2-t}^{-1} \sum_{T'} \left\{e(T \cup T') : \; T' \subseteq \bar{T}, \; |T'| = n/2 - t\right\}.$$

Counting how much different edges contribute to the sum, we know that the sum equals to

$$e(T)\binom{n-t}{n/2-t} + e(\bar{T})\binom{n-t-2}{n/2-t-2} + e(T,\bar{T})\binom{n-t-1}{n/2-t-1}.$$

From the fact that $e(T,\bar{T}) = e(G) - e(T) - e(\bar{T})$, we obtain that

$$A = \frac{n/2}{n-t}e(T) - \frac{(n/2-t)n/2}{(n-t)(n-t-1)}e(\bar{T}) + \frac{n/2-t}{n-t}e(G),$$

which satisfies

$$\begin{aligned} A \geq \ & \frac{n/2}{n-t}\left\{\frac{1}{2}\binom{t}{2} + 20\epsilon n^2\right\} - \frac{(n/2-t)n/2}{(n-t)(n-t-1)}\binom{n-t}{2}\left(\frac{1}{2}+9\epsilon\right) \\ & + \frac{n/2-t}{n-t}\binom{n}{2}\left(\frac{1}{2}-9\epsilon\right) > \frac{n^2}{16} + \epsilon n^2. \end{aligned}$$

Similarly, the assumption

$$e(T) < \frac{1}{2}\binom{t}{2} - 20\epsilon n^2,$$

leads a contradiction to the property P_5, too. □

A property is called *a quasi-random property* for $p = 1/2$ if it is equivalent to any property in Theorem 4.2. It is surprised that $P_2(4)$, which seems to be weaker, is a quasi-random property for $p = 1/2$.

Theorem 4.3 *The property*

$$P_2(4): \ e(G) \sim \frac{n^2}{4} \quad \textit{and} \quad N_G(C_4) \leq \left(\frac{n}{2}\right)^4 + o(n^4)$$

is a quasi-random property for $p = 1/2$.

Proof. See Fact 3 in the proof of the last theorem. □

Some other properties can be added to the list, one of which is in the next theorem.

Theorem 4.4 *The property*

$$P_8: \text{ For all } U, V \subseteq V(G),\ e(U,V) = \frac{1}{2}|U||V| + o(n^2)$$

is a quasi-random property for $p = 1/2$.

Proof. Let us prove the result by $P_4 \Longleftrightarrow P_8$. It suffices to show that $P_4 \Rightarrow P_8$. Suppose that P_4 holds. If U and V are disjoint, then

$$\begin{aligned} e(U,V) &= e(U \cup V) - e(U) - e(V) \\ &= \frac{1}{4}(u+v)^2 - \frac{1}{4}u^2 - \frac{1}{4}v^2 + o(n^2) \\ &= \frac{1}{2}uv + o(n^2), \end{aligned}$$

where $u = |U|$ and $v = |V|$. In case U and V are not disjoint, write $|U \cap V| = x$, from P_4 and what we just proved, we know that $e(U,V)$ equals to

$$\begin{aligned} & e(U \setminus V, V \setminus U) + e(U \cap V, U \setminus V) + e(U \cap V, V \setminus U) + 2e(U \cap V) \\ &= \frac{1}{2}(u-x)(v-x) + \frac{1}{2}x(u-x) + \frac{1}{2}x(v-x) + o(n^2) \\ &= \frac{1}{2}uv + o(n^2), \end{aligned}$$

which is P_8. □

The following theorem is for general edge density p. However, $0 < p < 1$ is fixed.

Theorem 4.5 *Let* $\{G\}$ *be a sequence of graphs, where* $G = G_n$ *is a graph of order* n. *Let* $0 < p < 1$ *be fixed. Then the following properties are equivalent:*

$P_1(h)$**:** *For any fixed* $h \geq 4$ *and graph* H *of order* h,

$$N_G^*(H) \sim p^{e(H)}(1-p)^{\binom{h}{2}-e(H)}n^h.$$

$P_2(t)$**:** $e(G) \sim \frac{pn^2}{2}$ *and* $N_G(C_t) \leq (pn)^t + o(n^t)$ *for any even* $t \geq 4$.

P_3: $e(G) \geq \frac{pn^2}{2} + o(n^2)$, $\lambda_1 \sim pn$ *and* $\lambda_2 = o(\lambda_1)$.

P_4: *For each* $U \subseteq V(G)$, $e(U) = p\binom{|U|}{2} + o(n^2)$.

P_5: *For each* $U \subseteq V(G)$ *with* $|U| = \lfloor \frac{n}{2} \rfloor$, $e(U) \sim \frac{p}{8}n^2$.

P_6: $\sum_{x,y} \left| s(x,y) - (p^2 + (1-p)^2)n \right| = o(n^3)$.

P_7: $\sum_{x,y} \left| |N(x) \cap N(y)| - p^2 n \right| = o(n^3)$.

Let us conclude the section with an example.

Example 4.1 *The Paley graph P_q of order q.*

This graph is defined in Chapter 2, where $q \equiv 1 \pmod 4$ is a prime power. Then P_q is $(q-1)/2$-regular, and the distinct eigenvalues of P_q are $(q-1)/2$, $(\sqrt{q}-1)/2$ and $-(\sqrt{q}-1)/2$. Therefore,

$$e(P_q) = \frac{q(q-1)}{4} \sim \frac{q^2}{4},\ \lambda_1 = \frac{q-1}{2} \sim \frac{q}{2},\ \lambda = \frac{\sqrt{q}-1}{2} = o(q).$$

Thus P_q satisfies quasi-random property P_3 hence all other quasi-random properties with $p = 1/2$.

4.2 Graph with small second eigenvalue

The last section was devoted to the quasi-random graphs of fixed edge density. Let us now switch to the case of density $p = p(n) = o(1)$, which is more important for some applications.

In applications, we shall allow the graphs to be semi-simple, that is, each vertex is attached at most a loop. When $p \to 0$, the situation is significantly more complicated as revealed by Chung and Graham (2002). The first remarked fact is that the properties defined for quasi-random graphs with fixed edge density may be not equivalent anymore. Let E_q^o be the Erdő-Rényi graph of order $n = q^2 + q + 1$. The graph is $(q+1)$-regular, in which $q+1$ vertices have loops (each of such vertices has one). So the edge density $p \sim \frac{1}{\sqrt{n}}$. We have found in Chapter 9 that $\lambda_1 = q + 1 \sim pn$, and $\lambda = \sqrt{q} = o(d)$. So the property P_3 holds. However,

$$p^4(1-p)^2 n^4 \sim n^2,$$

and thus the property $P_1(4)$ does not hold as E_q^o does not contain C_4.

Recall that the quasi-random property P_3, the magnitude of $\lambda = \lambda(G)$ is a measure of quasi-randomness. As called by Alon, a graph G is an (n, d, λ)-graph if G is d-regular with n vertices and

$$\lambda = \lambda(G) = \max\{|\lambda_i| : \ 2 \le i \le n\},$$

where $\lambda_1 = d$, and $\lambda_2, \ldots, \lambda_n$ are all eigenvalues of G. Here he connected quasi-randomness to the eigenvalue gap. For sparse graphs with $p = o(1)$, Chung and Graham (2002) found some equivalent properties under certain conditions. One of the properties is that $\lambda_1 \sim pn$ and $\lambda = o(\lambda_1)$.

We shall have more results on (n, d, λ)-graphs, which are due to Alon et.al, particularly Alon and Spencer (2008). For two (not necessarily disjoint) subsets B and C, we have defined $e(B, C)$ as the number of ordered pairs (u, v) with $u \in B$ and $v \in C$. If G is simple, then $e(B, C)$ is the same as defined in the last section, i.e., it counts each edge from $B \setminus C$ to $C \setminus B$ once, and each edge in $B \cap C$ twice. When G is semi-simple, it also counts each loop in $B \cap C$ once. For disjoint subsets B and C in a random graph, $e(B, C)$ is expected to be $\frac{d}{n}|B||C|$, which is close to the right-hand side of the inequality in the following theorem if λ is much smaller than d.

Theorem 4.6 *Let $G = (V, E)$ be a semi-simple (n, d, λ)-graph. Then for each partition of V into disjoint subsets B and C,*

$$e(B, C) \ge \frac{(d - \lambda)|B||C|}{n}$$

Proof. Let A be the adjacency matrix of G and I the identity matrix of order n. Observe that for any real vector x of dimension n (as a real valued function on V), we have

$$((dI - A)x, x) = \sum_{u \in V} \Big(dx_u^2 - \sum_{v:\, uv \in E} x_v x_u \Big)$$

$$= \ d \sum_{u \in V} x_u^2 - 2 \sum_{uv \in E} x_v x_u = \sum_{uv \in E} \Big(x_u - x_v \Big)^2.$$

Set $b = |B|$ and $c = |C| = n - b$. Define a vector $x = (x_v)$ by

$$x_v = \begin{cases} -c & v \in B, \\ b & v \in C. \end{cases}$$

Note that $dI - A$ and A have the same eigenvectors, and that the eigenvalues of $dI - A$ are precisely $d - \mu$ as μ ranges over all eigenvalues of A. Also, d is the largest eigenvalue of A corresponding to the eigenvector $J = (1, 1, \ldots, 1)^t$ and $(x, J) = 0$. Hence x is orthogonal to the eigenvector of the smallest eigenvalue of $dI - A$.

Since $dI - A$ is a symmetric matrix, its eigenvectors are orthogonal each other and form a basis of the n-dimensional space and x is a linear combination of these eigenvectors other than that of $J/\sqrt{n}$. This together with the fact that $d - \lambda$ is the second smallest eigenvalue of $dI - A$, we have

$$((dI - A)x, x) \geq (d - \lambda)(x, x) = (d - \lambda)(bc^2 + cb^2) = (d - \lambda)bcn.$$

However, as B and C form a partition of V,

$$\sum_{uv \in E} (x_u - x_v)^2 = e(B, C)(b + c)^2 = e(B, C)n^2,$$

implying the desired inequality. □

The next theorem bounds some kind of variance. In a random d-regular graph, we expect that a vertex v has $\frac{d}{n}|B|$ neighbors in B. The theorem shows that if λ is small, then $|N_B(v)|$ is not too far from the expectation for most vertices v, where $N_B(v) = N(v) \cap B$.

Theorem 4.7 *Let $G = (V, E)$ be a semi-simple (n, d, λ) graph. Then for each $B \subseteq V$,*

$$\sum_{v \in V} \left(|N_B(v)| - \frac{d}{n}|B| \right)^2 \leq \lambda^2 \frac{|B|(n - |B|)}{n}.$$

Proof. Let A be the adjacency matrix of G. Define a vector $f : V \to R$ by

$$f_u = \begin{cases} 1 - \frac{b}{n} & u \in B, \\ -\frac{b}{n} & u \notin B, \end{cases}$$

where $b = |B|$. Then $\sum_u f_u = 0$, and f is orthogonal to the eigenvector $J = (1, 1, \ldots, 1)^t$ of the largest eigenvalue d of A. Thus f is a linear combination of eigenvectors other than J, and

$$(Af, Af) \leq \lambda^2 (f, f) = \lambda^2 \frac{b(n-b)}{n}.$$

Let A_v be the row of A corresponding to vertex v. Then the coordinate $(Af)_v$ of Af at v is

$$A_v f = \Big(1 - \frac{b}{n}\Big)|N_B(v)| - \frac{b}{n}(d - |N_B(v)|) = |N_B(v)| - \frac{db}{n},$$

and thus

$$(Af, Af) = \sum_v \Big(|N_B(v)| - \frac{db}{n}\Big)^2,$$

the desired inequality follows. □

Corollary 4.1 *Let $G = (V, E)$ be a semi-simple (n, d, λ)-graph. Then for every two subsets B and C of G, we have*

$$\Big|e(B, C) - \frac{d}{n}|B||C|\Big| \leq \lambda \sqrt{|B||C|}.$$

Proof. Set $b = |B|$ and $c = |C|$. Note that

$$\begin{aligned} \Big|e(B,C) - \frac{dbc}{n}\Big| &= \Big|\sum_{v \in C} \Big(|N_B(v)| - \frac{db}{n}\Big)\Big| \leq \sum_{v \in C} \Big||N_B(v)| - \frac{db}{n}\Big| \\ &\leq \sqrt{c}\Big[\sum_{v \in C} \Big(|N_B(v)| - \frac{db}{n}\Big)^2\Big]^{1/2}, \end{aligned}$$

where the Cauchy-Schwarz inequality is used. From Theorem 4.7, we have

$$\begin{aligned} \Big|e(B,C) - \frac{dbc}{n}\Big| &\leq \sqrt{c}\Big[\sum_{v \in V} \Big(|N_B(v)| - \frac{db}{n}\Big)^2\Big]^{1/2} \\ &\leq \lambda\sqrt{c}\sqrt{b(1 - \frac{b}{n})} \leq \lambda\sqrt{bc} \end{aligned}$$

as desired. □

Let $e(B)$ and $\ell(B)$ be the number of edges and loops in B, respectively. Then

$$e(B,B) = 2\,e(B) + \ell(B).$$

Note that $\ell(B) \le |B|$ if G is semi-simple.

Corollary 4.2 *Let $G = (V,E)$ be a semi-simple (n,d,λ) graph, and let B be a subset of G. Then*

$$\Big|e(B) - \frac{d}{2n}|B|^2\Big| \le \frac{\lambda+1}{2}|B|.$$

Remark. By setting $e(B) = 0$, we have $\alpha(G) \le \frac{\lambda+1}{d}n$, which is slightly weaker than a similar bound obtained in Chapter 9.

For an (n,d,λ)-graph $G = (V,E)$ and $B \subseteq V$, define $\overline{B}$ as the set of vertices u so that the proportion of $N(u)$ in B, which is $|N_B(u)|/|B|$, is at most half of that in V. Then $|B||\overline{B}|$ is at most $\Theta(n^2/d)$ if $\lambda = \Theta(\sqrt{d})$.

Corollary 4.3 *Let $G = (V,E)$ be a semi-simple (n,d,λ)-graph and $B \subseteq V$. Define*

$$\overline{B} = \{u \in V : \ |N_B(u)| \le \frac{d}{2n}|B|\},$$

where $N_B(u) = N(u) \cap B$. Then

$$|B||\overline{B}| \le \Big(\frac{2\lambda n}{d}\Big)^2.$$

Consequently, $|B \cap \overline{B}| \le \frac{2\lambda n}{d}$.

Proof. From Theorem 4.7, we have

$$\sum_{v\in V}\Big(|N_B(v)| - \frac{d}{n}|B|\Big)^2 \le \lambda^2\frac{|B|(n-|B|)}{n} \le \lambda^2|B|.$$

Each $v \in \overline{B}$ contributes to the left-hand side more than $\left(\frac{d|B|}{2n}\right)^2$, thus

$$|\overline{B}|\Big(\frac{d|B|}{2n}\Big)^2 \le \lambda^2|B|,$$

implying what as claimed.

For an (n, d, λ)-graph, the spectral gap between d and λ is a measure for its quasi-random property. The smaller the value of λ compared to d, the closer is edge distribution to the ideal uniform distribution. How small can be λ?

Theorem 4.8 *Let G be an (n, d, λ)-graph and let $\epsilon > 0$. If $d \leq (1-\epsilon)n$, then*

$$\lambda \geq \sqrt{\epsilon d}.$$

Proof. Let A be the adjacency matrix of G. Then

$$\begin{aligned} nd &= 2e(G) = tr(A^2) = \sum_{i=1}^{n} \lambda_i^2 \\ &\leq d^2 + (n-1)\lambda^2 \leq (1-\epsilon)nd + n\lambda^2, \end{aligned}$$

which follows by what claimed.

On this estimate, we can say, not precisely, that an (n, d, λ)-graph with $\lambda = \Omega(\sqrt{d})$ has good quasi-randomness. Recall a result in Chapter 2, if G is an $srg(n, d, \mu_1, \mu_2)$ with $n \geq 3$. Then, except $\lambda_1 = d$, the other eigenvalues are solutions of the equation

$$\lambda^2 + (\mu_2 - \mu_1)\lambda + (\mu_2 - d) = 0.$$

Thus when $\mu_1 - \mu_2$ is small compared to d, which implies that λ is close to $\sqrt{d}$, G has good quasi-randomness.

Example 4.2 *The Erdős-Rényi graph E_q^0.*

Let E_q^0 be the Erdős-Rényi graph of order q^2+q+1 in Chapter 9, which is semi-simple, $(q+1)$-regular. Its distinct eigenvalues are $q+1$, $\sqrt{q}$ and $-\sqrt{q}$. So $\lambda \sim \sqrt{d}$.

Example 4.3 *The projective norm graph $G_{q,t}^o$.*

Let $G_{q,t}$ be the projective norm graph of Alon et. al. in Chapter 9, whose order $n = q^{t-1}(q-1)$. Each vertex has degree $q^{t-1}-1$ or $q^{t-1}-2$. It can be semi-simple and d-regular, where $d = q^{t-1} - 1$, if we attach some vertices with loops. The distinct eigenvalues of $G_{q,t}$ will be find in the next section as $q^{t-1} - 1$, $q^{(t-1)/2}$, 1, 0, -1, $-q^{(t-1)/2}$. Thus $\lambda \sim \sqrt{d}$.

Example 4.4 *The Alon's graph A_k.*

The A_k of Alon giving a constructive proof for $r(3,n) \geq cn^{3/2}$ is as follows. Let k be a positive integer that is not divisible by 3 and $F(2^k)$ be the finite field of 2^k elements. The elements of $F(2^k)$ are represented by binary vectors of length k. If a, b and c are such vectors, let (a,b,c) denote their concatenation, whose coordinates are those of a, followed by those b and those of c. Let $a|_i$ be the *ith* coordinate of a. Define

$$S = \{\alpha \in F^*(2^k) : \left.\alpha^7\right|_1 = 0\}, \quad T = \{\alpha \in F^*(2^k)\ \left.\alpha^7\right|_1 = 1\},$$

where the powers are computed in the field. Then $|S| = 2^{k-1} - 1$ and $|T| = 2^{k-1}$.

The graph A_k is defined on vertex set $F^3(2^k)$, whose vertices are all $n = 2^{3k}$ binary vectors of length $3k$. Two vectors u and v are adjacent if and only if there exist $s \in S$ and $t \in T$ so that

$$u + v = (s, s^3, s^5) + (t, t^3, t^5).$$

Theorem 4.9 *If k is not divisible by 3, then*

1. *The order of A_k is $n = 2^{3k}$;*
2. *A_k is d-regular, where $d = 2^{k-1}(2^{k-1} - 1) \sim \frac{1}{4}n^{2/3}$;*
3. *A_k is triangle-free;*
4. *Each eigenvalue μ other than the largest one satisfies*

$$-9 \cdot 2^k - 3 \cdot 2^{k/2} - \frac{1}{4} \leq \mu \leq 4 \cdot 2^k + 2^{k/2} + \frac{1}{4}.$$

□

Hence, $\lambda(A_k) \leq O(\sqrt{d})$. Also, from the results in Chapter 2 or Chapter 9, we have $\alpha(A_k) \leq O(n^{2/3})$.

The next example is due to Delsarte and Goethals and to Turyn (unpublished, reported in Sseidel (1976)). Define a graph $G_q^{(k)}$ of order q^2 as follows, where q is a prime power. Let the vertex set of $G_q^{(k)}$ be F_q^2, the two dimensional vector space over F_q. Partition the $q+1$ lines through the origin of the space into two sets P and N, where $|P| = k$. Two vertices x and y of $G_q^{(k)}$ are adjacent if $x - y$ is parallel to a line in P. The following result is easy to verify.

Theorem 4.10 *Let q be a prime power. Then the graph $G_q^{(k)}$ is an*

$$srg\Big(q^2, k(q-1), q-2+(k-1)(k-2), k(k-1)\Big),$$

and its spectrum is

eigenvalue	$k(q-1)$	$q-k$	$-k$
multiplicity	1	$k(q-1)$	$(q-1)(q+1-k)$

For any fixed p with $0 < p < 1$, if we take $k \sim pq^2$ as $q \to \infty$, then $G_q^{(k)}$ is quasi-random with edge density p.

4.3 Applications of characters⋆

We shall find the spectrum of $G_{q,t}$ defined in Chapter 9. Let us define the characters of a finite field $F(q)$, which are group homomorphisms from $F(q)$ or $F^*(q)$ to

$$S^1 = \{z : |z| = 1\} = \{e^{i\theta} : 0 \le \theta < 2\pi\},$$

respectively, where S^1 is a multiplicative group of complex numbers. An *additive character* of $F(q)$ is a function $\psi : F(q) \to S^1$ such that for any $x, y \in F(q)$,

$$\psi(x+y) = \psi(x)\psi(y).$$

Clearly $\psi(0) = 1$ and $\psi(-x) = \overline{\psi(x)}$. The trivial function ψ_0 with $\psi_0(x) \equiv 1$ is also called the *principle additive character* of $F(q)$.

A *multiplicative character* of $F(q)$ is a function $\chi : F^*(q) \to S^1$ such that for any $x, y \in F^*(q)$,

$$\chi(xy) = \chi(x)\chi(y).$$

Clearly $\chi(1) = 1$ and $\chi(x^{-1}) = \overline{\chi(x)}$. The trivial function χ_0 with $\chi_0(x) \equiv 1$ is also called the *principal multiplicative character* of $F(q)$. It is often to extend the domain of a multiplicative character χ to all of $F(q)$ by $\chi(0) = 0$ if $\chi \ne \chi_0$, and $\chi_0(0) = 1$. The character χ with $\chi(x) = x^{(q-1)/2}$ is often called *quadratic residue character.*

In the following proofs, we shall not distinguish the elements of $F(p)$ from the integers of $\{0, 1, \ldots, p-1\}$.

Lemma 4.1 *The numbers of additive characters and multiplicative characters of $F(q)$ are q and $q-1$, respectively.*

Proof. Let us begin with the multiplicative group $F^*(q)$, which is a cyclic group of order $q-1$ with $F^*(q) = \{1, \mu, \ldots, \mu^{q-2}\}$, where μ is a primitive element of $F(q)$. Each multiplicative character χ of $F(q)$ is uniquely determined by $\chi(\mu)$. From $1 = \chi(\mu^{q-1}) = \chi(\mu)^{q-1}$, we have that $\chi(\mu) = \zeta_{q-1}^k$ for some $0 \le k \le q-2$, where $\zeta_{q-1} = e^{2\pi i/(q-1)}$. If we use χ_1 to signify the multiplicative character of $F(q)$ with $\chi_1(\mu) = \zeta_{q-1}$, then the set of all multiplicative characters of $F(q)$ is $\{\chi_1^k : 0 \le k \le q-2\}$, in which χ_1^0 is the trivial character χ_0. Thus $F(q)$ has $q-1$ multiplicative characters, forming a group isomorphic to $F^*(q)$.

Let $q = p^m$ and let $\zeta_p = e^{2\pi i/p}$. For each $a = (a_1, a_2, \ldots, a_m) \in F^m(p)$, set

$$\psi_a : F(q) \to S^1, \psi_a(x) = \zeta_p^{a_1x_1 + a_2x_2 + \cdots + a_mx_m},$$

where $x = (x_1, x_2, \ldots, x_m)$ is the unique expression of x as a vector of $F^m(p)$. Then ψ_a is an additive character of $F(q)$. For $a \ne a'$, we show that $\psi_a \ne \psi_{a'}$. It suffices to show that ψ_a is not the trivial character for $a \ne 0$. Since for $a \ne 0$ there is some k such that $1 \le a_k \le p-1$, so for $e_k = (0, \ldots, 0, 1, 0, \ldots, 0)$, the unit vector with 1 at the kth coordinate, $\psi_a(e_k) = \zeta_p^{a_k} \ne 1$. Thus the group of additive characters of $F(q)$ contains at least hence exactly q elements. □

As usual, the function $\delta(x, y)$ is the Kronecker's symbol defined as

$$\delta(x, y) = \begin{cases} 1 & \text{if } x = y \\ 0 & \text{otherwise.} \end{cases}$$

The proof of Lemma 4.1 implies the following result.

Lemma 4.2 *Let χ be a multiplicative character of $F(q)$. Then*

$$\sum_{t \in F(q)} \chi(t) = q\,\delta(\chi, \chi_0).$$

Let us define the Gaussian sum as

$$G(\chi, \psi) = \sum_{x \in F(q)} \chi(x)\psi(x).$$

Theorem 4.11 *Let χ be a multiplicative character of $F(q)$ and ψ be an additive character of $F(q)$. Then $G(\chi_0, \psi_0) = q$, and $G(\chi, \psi) = 0$ if exactly one of χ and ψ is trivial. Furthermore*

$$|G(\chi, \psi)| = \sqrt{q}$$

if none of χ and ψ is trivial.

Proof. The first two equalities are easy, and we shall verify the last. In order to simplify the proof, we prove it for the case that q is a prime p, which is the most important special case. The proof for general q will be given later, which is more involved and the readers are encouraged to skip.

Let ψ and χ be additive character and multiplicative character of $F(p)$, none of which is trivial. From the proof of Lemma 4.1, we have $\psi(x) = \zeta_p^{ax}$, where $\zeta_p = e^{2\pi i/p}$ and $a \neq 0$. Let $g_a(\chi) = \sum_{x \in F(p)} \chi(x)\zeta_p^{ax}$, which is $G(\chi, \psi)$ on $F(p)$.

We shall verify that $\overline{g_a(\chi)} = \chi(a)\overline{g_1(\chi)}$. This follows from that

$$\begin{aligned} g_a(\chi) &= \sum_{x \in F(p)} \chi(x)\zeta_p^{ax} = \sum_{y \in F(p)} \chi(a^{-1}y)\zeta_p^{y} \\ &= \chi(a^{-1}) \sum_{y \in F(p)} \chi(y)\zeta_p^{y} = \overline{\chi(a)} g_1(\chi), \end{aligned}$$

and

$$|g_a(\chi)|^2 = g_a(\chi)\overline{g_a(\chi)} = |\chi(a)|^2 |g_1(\chi)|^2 = |g_1(\chi)|^2.$$

That is to say, $|g_a(\chi)|^2$ have the same value for any $a \neq 0$. On the other hand, for any $a \in F(p)$

$$g_a(\chi)\overline{g_a(\chi)} = \sum_{x \in F(p)} \chi(x)\zeta_p^{ax} \sum_{y \in F(p)} \overline{\chi(y)}\zeta_p^{-ay} = \sum_{x,\, y \in F(p)} \chi(x)\overline{\chi(y)}\zeta_p^{a(x-y)}.$$

It is easy to see that $\sum_{a \in F(p)} \zeta_p^{a(x-y)} = p\,\delta(x, y)$ as $a(x-y)$ ranges all of $F(p)$ for $x - y \neq 0$, which and the fact that $\chi(0) = 0$ as $\chi \neq \chi_0$ imply that

$$\sum_{a \in F(p)} g_a(\chi)\overline{g_a(\chi)} = \sum_{x,\, y \in F(p)} \chi(x)\overline{\chi(y)}\delta(x, y)p = (p-1)p.$$

Since $g_0(\chi) = 0$ as $\chi \neq \chi_0$, we obtain that $(p-1)|g_1(\chi)|^2 = (p-1)p$ hence $|g_a(\chi)| = |g_1(\chi)| = \sqrt{p}$. □

The proof of Theorem 4.11 for general $q = p^m$

The forms of additive characters in the proof of Lemma 4.1 are simple, but we shall express them in the other way for proving Theorem 4.11 in general case.

For $\alpha \in F(q) = F(p^m)$, define the *trace* of α to be

$$tr(\alpha) = \alpha + \alpha^p + \alpha^{p^2} + \cdots + \alpha^{p^{m-1}}.$$

Lemma 4.3 *If $\alpha, \beta \in F(q)$ and $a \in F(p)$, then*
(a) $tr(\alpha) \in F(p)$.
(b) $tr(\alpha + \beta) = tr(\alpha) + tr(\beta)$.
(c) $tr(a\alpha) = a\, tr(\alpha)$.
(d) For fixed $\alpha \neq 0$, $tr(\alpha x)$ maps $F(q)$ onto $F(p)$.

Proof. The properties (a),(b) and (c) follow from the facts that $tr^p(\alpha) = tr(\alpha)$, $(\alpha+\beta)^p = \alpha^p + \beta^p$, $\alpha^q = \alpha$ and $a^p = a$. To show the property (d), consider the fact that the polynomial $tr(\alpha x)$ has at most p^{m-1} roots and αx ranges all p^m elements of $F(q)$, there is $x \in F(q) = F(p^m)$ such that $tr(\alpha x) = c \neq 0$, where $c \in F(p)$. If $b \in F(p)$, the using property (c) we see that $tr((b/c)\alpha x) = (b/c)tr(\alpha x) = b$. Thus the trace $tr(\alpha x)$ is onto. □

For fixed $\alpha \in F(q)$, we now define $\psi_\alpha : F(q) \to S^1$ by

$$\psi_\alpha(x) = \zeta_p^{tr(\alpha x)},$$

where $\zeta_p = e^{2\pi i/p}$. Note that ψ_0 is the trivial additive character of $F(q)$. For the case $q = p$, $\psi_\alpha(x) = \zeta_p^{\alpha x}$ is exactly what we have used.

Lemma 4.4 *The function ψ_α has the following properties.*
(a) $\psi_\alpha(x+y) = \psi_\alpha(x)\psi_\alpha(y)$ for any $x, y \in F(q)$.
(b) If $\alpha \neq 0$, then there is $x \in F(q)$ such that $\psi_\alpha(x) \neq 1$.
(c) If $\alpha \neq 0$, then $\sum_{x \in F(q)} \psi_\alpha(x) = 0$; if $x \neq 0$, then $\sum_{\alpha \in F(q)} \psi_\alpha(x) = 0$.

Proof. The property (a) follows from that $tr(\alpha(x+y)) = tr(\alpha x) + tr(\alpha y)$. The property (b) follows from the fact that $tr(\alpha x)$ is onto as $\alpha \neq 0$, so there $x \in F(q)$ such that $tr(\alpha x) = 1$. Then $\psi_\alpha(x) = \zeta_p \neq 1$. As $\psi_\alpha(x) = \psi_x(\alpha)$, we shall only verify the first equality in the property (c). Let $S = \sum_{x \in F(q)} \psi_\alpha(x)$. Choose y such that $\psi_\alpha(y) \neq 1$, thus

$$\psi_\alpha(y)S = \sum_{x \in F(q)} \psi_\alpha(x)\psi_\alpha(y) = \sum_{x \in F(q)} \psi_\alpha(x+y) = S,$$

thus $S = 0$. □

Lemma 4.5 *For any fixed $\alpha \in F(q)$, the function ψ_α is an additive character of $F(q)$, and any additive character of $F(q)$ is of such form. Furthermore, for any $x, y \in F(q)$,*

$$\sum_{\alpha \in F(q)} \psi_\alpha(x-y) = q\,\delta(x,y).$$

Proof. The first assertion follows the property (a) in Lemma 4.4. For the second, we shall verify that the number of such functions is q. It suffices to show that if $\alpha \neq \beta$, the functions ψ_α and ψ_β are distinct. If $\psi_\alpha(x) = \psi_\beta(x)$ for any $x \in F(q)$, then

$$\zeta_p^{tr((\alpha-\beta)x)} = \psi_{\alpha-\beta}(x) = 1$$

for any $x \in F(q)$, implying that $\alpha = \beta$ from the property (b) in Lemma 4.4.

Since

$$\sum_{\alpha \in F(q)} \psi_\alpha(x-y) = \sum_{\alpha \in F(q)} \zeta_p^{tr(\alpha(x-y))},$$

which is q for $x = y$. If $x \neq y$, equality follows from the fact that $\alpha(x-y)$ ranges over all of $F(q)$ and the property (c) in Lemma 4.4. □

We now write Gaussian sum in the form

$$G(\chi, \psi_\alpha) = \sum_{x \in F(q)} \chi(x)\psi_\alpha(x).$$

We shall prove that if $\chi \neq \chi_0$ and $\alpha \neq 0$, then

$$|G(\chi, \psi_\alpha)| = \sqrt{q}.$$

Proof. The proof is an analogy of that for the case $q = p$. For any $\alpha \neq 0$, we first verify that $\overline{G(\chi, \psi_\alpha)} = \chi(\alpha)\overline{G(\chi, \psi_1)}$. This is because that

$$\begin{aligned} G(\chi, \psi_\alpha) &= \sum_{x \in F(q)} \chi(x)\zeta_p^{tr(\alpha x)} = \sum_{y \in F(q)} \chi(\alpha^{-1}y)\zeta_p^{tr(y)} \\ &= \chi(\alpha^{-1}) \sum_{y \in F(q)} \chi(y)\zeta_p^{tr(y)} = \overline{\chi(\alpha)}G(\chi, \psi_1). \end{aligned}$$

Therefore, we have $|G(\chi, \psi_\alpha)|^2 = |G(\chi, \psi_1)|^2$ for $\alpha \neq 0$. On the other hand, for any $\alpha \in F(q)$,

$$\begin{aligned} G(\chi, \psi_\alpha)\overline{G(\chi, \psi_\alpha)} &= \sum_{x \in F(q)} \chi(x)\zeta_p^{tr(\alpha x))} \sum_{y \in F(q)} \overline{\chi(y)}\zeta_p^{-tr(\alpha y))} \\ &= \sum_{x,\, y \in F(q)} \chi(x)\overline{\chi(y)}\zeta_p^{tr(\alpha(x-y))}. \end{aligned}$$

Since $\chi(0) = 0$ as $\chi \neq \chi_0$,

$$\sum_{\alpha \in F(q)} G(\chi, \psi_\alpha)\overline{G(\chi, \psi_\alpha)} = \sum_{x,\, y} \chi(x)\overline{\chi(y)}\delta(x, y)q = q(q-1).$$

Observing that $G(\chi, \psi_0) = 0$ as $\chi \neq \chi_0$, we have

$$\sum_{\alpha \in F(q)} G(\chi, \psi_\alpha)\overline{G(\chi, \psi_\alpha)} = \sum_{\alpha \in F^*(q)} |G(\chi, \psi_1)|^2 = (q-1)|G(\chi, \psi_1)|^2,$$

yielding that $(q-1)|G(\chi, \psi_1)|^2 = q(q-1)$ hence $|G(\chi, \psi_\alpha)| = |G(\chi, \psi_1)| = \sqrt{q}$. The proof for general case of Theorem 4.11 is completed. □

The order of a multiplicative character χ is the smallest positive integer d such that $\chi^d = \chi_0$. A more sophisticated result on character sum is the Weil's theorem as follows. Let χ be the multiplicative character of $F_q = F(q)$ of order $d > 1$ and $f(x)$ a polynomial over F_q. If $f(x)$ has precisely s distinct zeros and it is not the form $c(g(x))^d$, where $c \in F_q$ and $g(x) \in F_q[x]$, then

$$\left| \sum_{x \in F(q)} \chi(f(x)) \right| \leq (s-1)\sqrt{q}. \tag{4.1}$$

In particular, the inequality holds when χ is the square residue character and $f(x)$ is not the form $cg^2(x)$, where $c \in F_q$ and $g(x) \in F_q[x]$. Similarly, for an additive character $\psi \neq \psi_0$, if $g(x)$ is a polynomial of degree $n < q$ with g.c.d. $(n, q) = 1$, then $\left|\sum_{x\in F(q)} \psi(g(x))\right| \leq (n-1)\sqrt{q}$.

The prepared results are enough for introducing the following result of Szabó (2003) on spectrum of $G^o_{q,t}$, which is constructed in Chapter 9. Note that $G^o_{q,t}$ is $q^{t-1}(q-1)$-regular, in which each vertex $(A, a) \in F(q^2) \times F^*(q)$ with $N(2A) = a^2$ has a loop.

Theorem 4.12 *Let $t \geq 2$ be an integer and q be an odd prime power. The spectrum of $G^o_{q,t}$ is as follows.*

eig.	$q^{t-1}-1$	$q^{(t-1)/2}$	1	0	-1	$-q^{(t-1)/2}$
mul.	1	$\frac{(q^{t-1}-1)(q-2)}{2}$	$\frac{q^{t-1}-1}{2}$	$q-2$	$\frac{q^{t-1}-1}{2}$	$\frac{(q^{t-1}-1)(q-2)}{2}$

Proof. Let M be the adjacency matrix of $G^o_{q,t}$. Let ψ be an additive character of $F(q^{t-1})$ and χ be a multiplicative character of $F(q)$. Let $V(\psi, \chi)$ be the column vector whose coordinates are labeled by vertices of $G^o_{q,t}$, whose entry at (X, x) is $\psi(X)\chi(x)$. Then the entry of the column vector $MV(\psi, \chi)$ at (A, a) is

$$\begin{aligned}
\sum_{\substack{(B,b)\in F(q^{t-1})\times F^*(q)\\ N(A+B)=ab}} \psi(B)\chi(b) &= \sum_{B\in F(q^{t-1})\setminus\{-A\}} \psi(B)\chi\left(\frac{N(A+B)}{a}\right) \\
&= \sum_{C\in F^*(q^{t-1})} \psi(C-A)\chi\left(\frac{N(C)}{a}\right) \\
&= \left(\sum_{C\in F^*(q^{t-1})} \psi(C)\chi(N(C))\right) \overline{\psi(A)}\ \overline{\chi(a)}.
\end{aligned}$$

Setting

$$\lambda = \lambda(\psi, \chi) = \sum_{C\in F^*(q^{t-1})} \psi(C)\chi(N(C)),$$

then $\overline{\lambda(\chi, \psi)} = \lambda(\overline{\psi}, \overline{\chi})$,

$$MV(\psi, \chi) = \lambda V(\overline{\psi},\ \overline{\chi}), \tag{4.2}$$

and $MV(\overline{\psi},\ \overline{\chi}) = \overline{\lambda}V(\psi,\chi)$. Thus we have

$$M^2V(\psi,\chi) = \lambda\,\overline{\lambda}V(\psi,\chi) = |\lambda|^2V(\psi,\chi).$$

Hence $V(\psi,\chi)$ is an eigenvector of M^2 with corresponding eigenvalue $|\lambda(\psi,\chi)|^2$.

Observe that in the multiplicative group consisting of additive characters, the inverse ψ^{-1} of ψ is $\overline{\psi}$, and the similar statement holds also for the multiplicative group consisting of multiplicative characters. We claim that the eigenvectors of the form $V(\psi,\chi)$ are pairwise orthogonal. Let $(\psi',\chi') \neq (\psi,\chi)$, and let $\psi'' = \psi'\psi^{-1} = \psi'\overline{\psi}$ and $\chi'' = \chi'\chi^{-1} = \chi'\overline{\chi}$. Then $(\psi'',\chi'') \neq (\psi_0,\chi_0)$, where ψ_0 and χ_0 are trivial additive character and trivial multiplicative character, respectively. The inner product of complex vectors $V(\psi',\chi')$ and $V(\psi,\chi)$ is

$$\begin{aligned} V^T(\psi',\chi')\overline{V(\psi,\chi)} &= V^T(\psi',\chi')V(\overline{\psi},\overline{\chi}) \\ &= \sum_{(X,x)\in F(q^{t-1})\times F^*(q)} \psi''(X)\chi''(x) \\ &= \sum_{X\in F(q^{t-1})} \psi''(X) \sum_{x\in F^*(q)} \chi''(x) = 0, \end{aligned}$$

as one of sum in the last row is 0. The number of vectors of form $V(\psi,\chi)$ is equal to the order of $G^o_{q,t}$ by Lemma 4.1, hence all eigenvalues of M^2 are of form $|\lambda(\psi,\chi)|^2$. Therefore, any eigenvalue of M is of form

$$\pm|\lambda(\psi,\chi)| = \pm\left|\sum_{C\in F^*(q^{t-1})} \psi(C)\chi(N(C))\right|.$$

When $\psi = \psi_0$ and $\chi = \chi_0$, the corresponding eigenvalue is $q^{t-1}-1$, which can be obtained from Perron-Frobenius Theorem with multiplicity 1.

Let μ be a primitive element of $F(q^{t-1})$, and let

$$A_k = \{\mu^{k+j(q-1)} : 0 \leq j \leq \ell - 1\},$$

where $\ell = (q^{t-1}-1)/(q-1)$. Then $A_0, A_1, \ldots, A_{q-2}$ form a partition of $F^*(q)$ with $|A_k| = \ell$. It is easy to see $N(x) = N(y)$ if x and y are in the same A_k. Therefore, when $\psi = \psi_0$ and $\chi \neq \chi_0$, as

$$|\lambda(\psi_0,\chi)| = \left|\sum_{C\in F^*(q^{t-1})} \chi(N(C))\right| = \left|\ell \sum_{c\in F^*(q)} \chi(c)\right| = 0,$$

thus 0 is an eigenvalue of M with multiplicity $q-2$ which is the number of multiplicative characters of $F(q)$ except χ_0.

When $\psi \neq \psi_0$ and $\chi = \chi_0$,

$$\lambda(\psi, \chi_0) = \sum_{C \in F^*(q^{t-1})} \psi(C) = -\psi(0) = -1.$$

So 1 is an eigenvalue of M^2 with multiplicity $q^{t-1} - 1$, then ± 1 are eigenvalues of M with the sum of the multiplicities being $q^{t-1} - 1$. Let $W(\psi) = V(\psi, \chi_0) + V(\overline{\psi}, \chi_0)$. It follows from (4.2) that for any $\psi \neq \psi_0$, $MW(\psi) = -W(\psi)$. For any $\psi, \psi' \neq \psi_0, \psi \neq \overline{\psi'}$, it is easy to see that the complex vectors $W(\psi)$ and $W(\psi')$ are orthogonal. So -1 is an eigenvalue of M with multiplicity at least $(q^{t-1} - 1)/2$. Similarly, by considering $V(\psi, \chi_0) - V(\overline{\psi}, \chi_0)$, we know that 1 is an eigenvalue of M with multiplicity at least $(q^{t-1} - 1)/2$, hence each multiplicity is exactly $(q^{t-1} - 1)/2$.

When $\psi \neq \psi_0$ and $\chi \neq \chi_0$, observing that χN is a non-trivial multiplicative character of $F(q^{t-1})$, by Theorem 4.11 on Gaussian sum,

$$|\lambda| = \left| \sum_{C \in F^*(q^{t-1})} \psi(C)\chi(N(C)) \right| = q^{(t-1)/2}.$$

Let S and T be the multiplicities of the eigenvalues $q^{(t-1)/2}$ and $-q^{(t-1)/2}$ of M, respectively. As $(q^{t-1} - 1)(q-2)$ is the number of vectors of form $V(\psi, \chi)$ with $\psi \neq \psi_0$ and $\chi \neq \chi_0$,

$$S + T = (q^{t-1} - 1)(q - 2).$$

By the definition, the graph $G^o_{q,t}$ has a loop at each vertex (A, a) if and only if $N(2A) = a^2$. Since exactly $(q-1)/2$ elements of $F^*(q)$ are squares and the equation $N(X) = y$ has $(q^{t-1} - 1)/(q-1)$ solutions in X for each fixed $y \in F^*(q)$, there are $(q^{t-1} - 1)/2$ elements $A \in F(q^{t-1})$ with $N(2A)$ being a non-zero square. Once $N(2A)$ is a non-zero square, there are two distinct elements $a, -a \in F^*(q)$ with $N(2A) = a^2 = (-a)^2$. Thus $G^o_{q,t}$ contains $q^{t-1} - 1$ loops, which is the trace of M. Hence

$$q^{t-1} - 1 = tr(M) = \sum_{j=1}^{q^{t-1}(q-1)} \lambda_j$$

$$= \quad q^{t-1} - 1 + \frac{q^{t-1} - 1}{2} - \frac{q^{t-1} - 1}{2} + q^{(t-1)/2}(S - T),$$

implying that $S = T = (q^{t-1} - 1)(q - 2)/2$. □

The above theorem has the following corollary, which and the lower bound in Chapter 9 imply $\alpha(G_{q,t}) = \alpha(G^o_{q,t}) = \Theta(q^{(t+1)/2})$.

Corollary 4.4 *Let $t \geq 2$ be an integer and q be an odd prime power. Then*

$$\alpha(G_{q,t}) \leq \frac{(q^t - q^{t-1})(q^{(t-1)/2} + 1)}{q^{t-1} + q^{(t-1)/2} - 1} \sim q^{(t+1)/2} \sim n^{(t+1)/(2t)},$$

where $n = q^{t-1}(q - 1)$ is the order of $G_{q,t}$.

Let us conclude this section with an algebraic construction that almost matches the probabilistic bound $r_k(K_{m,n}) \geq k^m n - n^{1/2+o(1)}$ in Chapter 5.

Theorem 4.13 *Let positive integers k and m be fixed. Then*

$$r_k(K_{m,n}) \geq k^m n - n^{0.525}.$$

for large n.

Proof. As the assertion is trivial for $k = 1$, we assume that $k \geq 2$. Let $p \equiv 1 \pmod{2k}$ be a prime and F_p the finite field of p elements. Let μ be a primitive element of F_p. Define a logarithmic-like function $\log_\mu(x) : F_p^* \to Z_{p-1} = \{0, 1, \dots, p - 2\}$ as

$$\log_\mu(x) = \ell \text{ if } x = \mu^\ell, \ 0 \leq \ell \leq p - 2.$$

For every j with $0 \leq j \leq k - 1$, define a graph H_j on vertex set F_p, in which x and y is adjacent in H_j if and only if

$$\log_\mu(x - y) \equiv j \ (\bmod k).$$

As $p \equiv 1 \pmod{2k}$ and $(-1)^2 = 1$, we have $-1 = \mu^{(p-1)/2}$, and thus $\log_\mu(x - y) \equiv \log_\mu(y - x) \pmod{k}$, so the definition is compatible. In case $k = 2$, the graph H_0 is the Paley graph.

Lemma 4.6 *Let $k \geq 2$ be an integer and $p \equiv 1 \pmod{2k}$ be a prime. Let H_j, $0 \leq j \leq k-1$, be the graph defined with respect to a primitive element of μ of F_p. Then these H_j are pairwise isomorphic.*

Proof. We shall verify that each H_j is isomorphic to H_0. Define a bijection ϕ on $F(p)$ as $\phi(z) = \mu^j z$. Then $\{x, y\}$ is an edge of H_0 if and only if $x - y = \mu^\ell$ for some $\ell \equiv 0 \pmod{k}$. As $\phi(x) - \phi(y) = \mu^{j+\ell}$, thus $\{x, y\}$ is an edge of H_0 if and only if $\{\phi(x), \phi(y)\}$ is an edge of H_j. Thus H_j is isomorphic to H_0. For any vertex x, its neighborhood in H_0 is

$$\{x + \mu^k, x + \mu^{2k}, \ldots, x + \mu^{p-1}\},$$

so the degree of x in H_0 is $((p-1)/k)$. This proves the lemma. □

Let $\zeta_k = e^{2\pi i/k}$. It is easy to see the following identity holds

$$(x - \zeta_k)(x - \zeta_k^2) \cdots (x - \zeta_k^{k-1}) = 1 + x + x^2 + \cdots + x^{k-1}. \qquad (4.3)$$

Define a function χ on F_p^* as

$$\chi(x) = \zeta_k^\ell, \quad \text{where } \ell \equiv \log_\mu x \pmod{k}.$$

Extend χ to all of F_p by $\chi(0) = 0$. Then χ is a multiplicative character of F_p of order k.

Let $U \subseteq F_p$ be a subset of vertices of the graph H_0 with $|U| = m$. Denote by $J(U)$ for $\cap_{u \in U} N(u)$. If $|J(U)| < n$ for any such U, then $r_k(K_{m,n}) > p$ from Lemma 4.6. For a fixed U, define a function $f(x)$ on $x \in F_p$ as

$$f(x) = \prod_{u \in U} \prod_{j=1}^{k-1} (\chi(x-u) - \zeta_k^j) = \prod_{u \in U} \sum_{j=0}^{k-1} \chi^j(x-u),$$

where we use the identity (4.3). For $x \notin U$, if $x \notin J(U)$, then $f(x) = 0$ as $\chi(x-u) = \zeta_k^j$ for some j with $1 \leq j \leq k-1$. If $x \in J(U)$, then $f(x) = k^m$ as $\log_\mu(x-u) \equiv 0 \pmod{k}$ hence $\chi(x-u) = 1$. Therefore, we have

$$\sum_{x \notin U} f(x) = k^m |J(U)|.$$

Set $U = \{u_1, u_2, \ldots, u_m\}$. Note that χ is multiplicative thus

$$\begin{aligned} f(x) &= \prod_{t=1}^{m}\Big(1+\chi(x-u_t)+\cdots+\chi^{k-1}(x-u_t)\Big) \\ &= \sum_{0\le j_1,\ldots,j_m\le k-1} \chi\Big((x-u_1)^{j_1}\cdots(x-u_m)^{j_m}\Big) \\ &= 1+\sum_{\substack{0\le j_1,\ldots,j_m\le k-1\\ j_1+\ldots+j_m\ge 1}} \chi\Big((x-u_1)^{j_1}\cdots(x-u_m)^{j_m}\Big). \end{aligned}$$

Applying the Weil's theorem for the the polynomial $(x-u_1)^{j_1}\cdots(x-u_m)^{j_m}$ with $j_1+\cdots+j_m\ge 1$, which is not the form $ch^k(x)$ with $c\in F_p$ and $h(x)\in F_p[x]$ as $j_1,\ldots,j_m<k$, from (4.1), we have

$$\Big|\sum_{x\in F_p}\chi\Big((x-u_1)^{j_1}\cdots(x-u_m)^{j_m}\Big)\Big| \le (m-1)\sqrt{p}.$$

Hence we obtain that

$$\begin{aligned} \Big|p-\sum_{x\in F_p} f(x)\Big| &= \Big|\sum_{x\in F_p}\sum_{\substack{0\le j_1,\ldots,i_m\le k-1\\ j_1+\ldots+j_m\ge 1}} \chi\Big((x-u_1)^{j_1}\cdots(x-u_m)^{j_m}\Big)\Big| \\ &= \Big|\sum_{\substack{0\le j_1,\ldots,j_m\le k-1\\ j_1+\ldots+j_m\ge 1}}\sum_{x\in F_p} \chi\Big((x-u_1)^{j_1}\cdots(x-u_m)^{j_m}\Big)\Big| \\ &\le \sum_{\substack{0\le j_1,\ldots,j_m\le k-1\\ j_1+\ldots+j_m\ge 1}} (m-1)\sqrt{p}. \end{aligned}$$

It is well-known that the number of solutions of nonnegative integers $(j_1,\ldots,j_m)$ of the equation

$$j_1+j_2+\cdots+j_m = s$$

is $\binom{s+m-1}{s}$ for a fixed integer s. Omitting the constraint that $j_1,\ldots,j_m\le k-1$, we obtain that

$$\Big|p-\sum_{x\in F_p} f(x)\Big| \le \sum_{s=1}^{m(k-1)}\binom{s+m-1}{s}(m-1)\sqrt{p} = A\sqrt{p},$$

where $A = A(k, m)$ is independent of p. Note that $|f(x)| \leq k^m$, thus $|\sum_{x \in U} f(x)| \leq mk^m$ and

$$\begin{aligned} \Big|p - k^m |J(U)|\Big| &= \Big|p - \sum_{x \notin U} f(x)\Big| \leq \Big|p - \sum_{x \in F_p} f(x)\Big| + \Big|\sum_{x \in U} f(x)\Big| \\ &\leq A\sqrt{p} + mk^m \leq (A+1)\sqrt{p} \end{aligned}$$

for large p, which implies that

$$k^m |J(U)| \leq p + (A+1)\sqrt{p}.$$

It is known that there are asymptotically $N/(\phi(2k)\log N)$ primes p in the form $p \equiv 1 \pmod{2k}$ between 1 and N, where $\phi(2k)$ is the number of integers from 1 to $2k$ that are relatively prime to $2k$. Let $p \equiv 1 \pmod{2k}$ be a prime between $k^m n - n^{0.525}$ and $k^m n - n^{0.525}/2$. The existence of such prime for large n is ensured by results for estimating the difference between consecutive primes, see Baker, Harman and Pintz (2001). The constant 0.525 is in the process of improvement to $0.5 + o(1)$ implied by the famous Riemann hypothesis. By choosing such p, we have

$$|J(U)| \leq n - \frac{n^{0.525}}{2} + (A+1)\sqrt{k^m n} < n,$$

for large n. Thus H_0 contains no $K_{m,n}$, implying

$$r_k(K_{m,n}) > p \geq k^m n - n^{0.525}$$

as each H_i is isomorphic to H_0. □

The largest difference between consecutive primes is conjectured as $p^{1/2+o(1)}$. If so, we have $r_k(K_{m,n}) \geq k^m n - n^{1/2+o(1)}$, which is the same as that in Chapter 5.

4.4 References

N. Alon and V. Rödl, Sharp bounds for some multicolor Ramsey numbers, *Combinatoria*, **25** (2005), 125-141.

N. Alon and J. Spencer, *The Probabilistic Method, 3rd ed.*, Wiley-Interscience, New York, 2008.

R. Baker, G. Harman and J. Pintz, The difference between consecutive primes, II, *Proc. Lond. Math. Soc.*, **83** (2001), 532-562.

F. R. Chung, R. Graham, Sparse quasi-random graphs, *Combinatoria*, **22** (2002), 217-244.

F. R. Chung, R. Graham and R. Wilson, Quasi-random graphs, *Combinatoria*, **9** (1989), 345-362.

M. Krivelevich and B. Sudakov, Pseudo-random graphs, *Bolyai Soc. Math. Stud.*, **15** (2006), 199-262.

J. Seidel, A survey of two-graphs, in: Colloquio Internazionale sulle Teorie Combinatorie (Rome, 1973), vol I, Atti dei Convegni Lincei, No. 17, Accad. Naz. Lincei, Rome, 1976, 481C 511.

T. Szabó, On the spectrum of projective norm-graphs, *Inform. Process. Lett.*, **86** (2) (2003), 71-74.

A. Thomason, Pseudo-random graphs, in: *Proceedings of Random Graphs, Poznań 1985, M. Karoński, Eds.*, *Ann. Discrete Math.* **33** (1987), 307-331.

Chapter 5

Real-world Networks

Complex systems from various fields, such as physics, biology, or sociology, can be systematically analyzed using their network representation. A network (also known as a graph) is composed of vertices (or nodes) and edges, where vertices represent the constitutes in the system and edges represent the relationships between these constitutes.

We shall introduce some basic concepts for real-world networks in this chapter.

5.1 Data and empirical research

Big data is a term for large or complex data. The term often refers simply to that traditional data processing applications are inadequate, and seldom to a particular size of data set. Most big data come from real-world networks, and analysis of such data sets can find new correlations to spot business trends, prevent diseases, combat crime and so on.

Empirical research is research using empirical evidence, where the empirical evidence is the record of direct observations or experiences in form of data and big data. Through quantifying the data, a researcher can answer empirical questions from real-world. In particular, we are interested in the empirical research and the data from real-world networks.

It is usual that researchers aim the common case, so they often describe the behavior by ignoring some case that are not significant for the research. This is similar to the case in random graph, we describe an event by saying “almost all” to signify that the probability of event goes to 1. It is often that we are more concerned with the average of parameters since they are concentrated at the average in most cases. Average of some parameters may be more meaningful than the extremal value of them.

However, this is not always the case. When investigating a social networks, the nodes of large degrees, called “hubs” such as internet celebrities, attracted much attention as these nodes are very important for the structure of the networks.

Collecting data that needed is an challenge before analysis. For example, the data in Barabási and Albert (2009) came from a software designed to collect the links in World Wide Web pointing from one page to another. The data in Backstrom and Kleinberg (2014) and Ugander, Backstrom, Marlow and Kleinberg(2012) came from Facebook Inc. directly as several of co-authors are employees of the company.

5.2 Six degrees of separation

Six degrees of separation is the theory that any pair of persons is six or fewer steps away in the world as a network connected by friendship, which means the maximum distance of nodes in the network is at least six in language of graph theory. However, as claimed before, “any pair” for social networks for sociology may mean most pairs.

The term *small world* became famous since a paper of S. Milgram (1967) who is an American psychologist. Some seminal works have been conducted before Milgram took up the experiments reported as the small world problem, and the experiment is called “the small-world experiment”, in which Milgram and other researchers examined the average path length for social networks of people in the United States. The research suggested that human society is a small-world-type network, and the experiments are often associated with the phrase “six degrees of separation”, although Milgram did not use this term himself.

Milgram’s experiment developed out of a desire to learn more about the probability that two randomly selected people would know each other. This is one way of looking at the small world problem.

Though the experiment went through several variations, Milgram typically chose individuals in cities of Omaha, Nebraska, and Wichita, Kansas, to be the starting points and Boston, Massachusetts, to be the end point of a chain of correspondence. These cities were selected because they were thought to represent a great distance in US, both socially and geographically.

Information packets (a letter,a roster and postcards) were initially sent to randomly selected individuals in Omaha or Wichita. In the more likely case that the person did not personally know the target, then the person was to think of a friend or relative he knew personally who was more likely to know the target. He was then directed to sign his name on a roster in the information packet and forward the packet to that person. When and if the package eventually reached the contact person in Boston, the researchers could examine the roster to count the number of times it had been forwarded from person to person.

However, a significant problem was that often people refused to pass the letter forward, and thus the chain never reached its destination. In one case, only 64 of the 296 letters eventually did reach the target contact. Among these chains, the average path length fell around five and a half or six. Hence, the researchers concluded that people in US are separated by about six people on average.

Smaller communities, such as mathematicians and actors, have been found to be densely connected by chains of personal or professional associations. Mathematicians have created the Erdős number to describe their distance from Paul Erdős based on shared publications. A similar exercise has been carried out for the actor Kevin Bacon and other actors who appeared in movies together with him.

In 2001, D. Watts attempted to recreate Milgram's experiment on the Internet, using an e-mail message as the "package" that needed to be delivered, with $48,000$ senders and 19 targets (in 157 countries). Watts found that the average number of intermediaries was around six, reported in [**?**]. Today, the phrase "six degrees of separation" is often used as a synonym for the idea of the "small world" phenomenon.

Watts and Strogatz (1998) showed that the average path length between two nodes in a random network is equal to $\log N/\log K$, where

N is number of nodes and K is degree of acquaintances per node. Thus, assuming 10% of population of US is too young to participate and $N = 300,000,000$ (90% of the US population and $K = 30$, the Degrees of Separation 5.7. If $N = 6,000,000,000$ (90% of the World population) and $K = 30$, then Degrees of Separation 6.6.

However, the convenient way of communication in a social network will make the average distance smaller and smaller. Facebook's data team released data in online papers described that amongst all Facebook users at the time of research, the average distances of friendship links were 5.28 in 2008, 4.74 in 2011 and 3.57 in February 2016 (this year). The world changes from "six degrees of separation" to "four degrees of separation".

5.3 Clustering coefficient

An important measure of network topology, called *clustering coefficient*, assesses the triangular pattern as well as the connectivity in a vertex's neighborhood: a vertex has a high clustering coefficient if its neighbors tend to be directly connected with each other. The clustering coefficient c_v of a vertex v can be calculated as

$$c_v = \begin{cases} 0, & \text{if } d_v = 0 \\ \frac{e_v}{\binom{d_v}{2}} & \text{if } d_v \geq 2. \end{cases}$$

For $d_v = 1$, it is a convention to define $c_v \in [0, 1]$ depending on the situation. Thus $0 \leq c_v \leq 1$. The clustering coefficient c_v for $d_v \geq 2$ is the ratio of number of triangles and all possible triangles that share vertex v.

Let G^k be a graph obtained from G by adding new edges connecting vertices of distance at most k in G. It is to see if $n \geq 8$, then $c_v = 1/2$ for each v in circular lattice C_n^2.

For a graph G of order N (i.e., G contains N vertices) and minimum degree $\delta(G) \geq 2$, its *average clustering coefficient* is defined as

$$\overline{c}(G) = \frac{1}{N} \sum_{v \in V} c_v = \frac{2}{N} \sum_{v \in V} \frac{e_v}{d_v(d_v - 1)}.$$

Average clustering coefficient explains the clustering (triangulation) within a network by averaging the clustering coefficients of all its nodes. The idea of clustering coefficient is proposed (especially in the analysis of social networks) to measure the local connectivity or "cliqueness" of a social network. If a network has a high average clustering coefficient and a small average distance, it is often called a "small-world" network.

Let us label the vertices of G of order N as $v_1, v_2, \ldots, v_N$. Recall that $A = (a_{ij})_{N\times N}$ is the adjacency matrix of G, where

$$a_{ij} = \begin{cases} 1, & \text{if } v_iv_j \in E, \\ 0, & \text{otherwise.} \end{cases}$$

We also call the eigenvalues of A as eigenvalues of G. Let $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_N$ be eigenvalues of G in a non-increasing order. Set

$$\lambda = \lambda(G) = \max\{|\lambda_i| : \ 2 \leq i \leq N\}.$$

As called by Alon, a graph G is an (N, d, λ)-graph if G is d-regular with N vertices and $\lambda = \lambda(G)$. Note that a d-regular connected graph satisfies that $\lambda_1 = d$. For an (N, d, λ)-graph, the spectral gap between d and λ is a measure for its quasi-random property. The smaller the value of λ compared to d, the closer is the edge distribution to the ideal uniform distribution (i.e., it becomes a random graph). We may say, not precisely, that an (N, d, λ)-graph with $\lambda = O(\sqrt{d})$ has good quasi-randomness. Generally, this is a weak condition as most random graphs are such graphs.

Theorem 5.1 *Let G be an (N, d, λ)-graph that is connected. If $\lambda = O(\sqrt{d})$ as $d \to \infty$, then*

$$\overline{c}(G) \sim \frac{d}{N}.$$

Proof. Let A be adjacency matrix of G. Note that A is symmetric, and thus it is diagonalizable. Let $\lambda_1, \lambda_2, \ldots, \lambda_N$ be the eigenvalues of A. Then the eigenvalues of A^k are $\lambda_1^k, \lambda_2^k, \ldots, \lambda_N^k$. Note that the (i, j) element of A^k is the number of walks from vertex v_i to vertex v_j, and a closed walk of length 3 is exactly a triangle. Thus the *ith* diagonal element of A^3 is $2e_{v_i}$, and

$$\overline{c}(G) = \frac{2}{N}\sum_v \frac{e_v}{d_v(d_v-1)} = \frac{1}{Nd(d-1)}\sum_{i=1}^{N}\lambda_i^3 = \frac{1}{Nd(d-1)}\left(d^3 + \sum_{i=2}^{N}\lambda_i^3\right),$$

where we used the fact that $\lambda_1 = d$ as G is d-regular and connected. The assumption $\lambda = O(\sqrt{d})$ implies that

$$\frac{\left|\sum_{i=2}^{N} \lambda_i^3\right|}{Nd(d-1)} \leq \frac{N\lambda^3}{Nd(d-1)} = \frac{O(d^{3/2})}{d^2} \to 0.$$

Thus

$$\overline{c}(G) \sim \frac{d^2}{N(d-1)} \sim \frac{d}{N}$$

for large d. □

5.4 Small-world networks

The small-world phenomenon is typical for random graphs that have small maximum distances. A definition for small-world network describes it as a network, in which the typical distance L between two randomly chosen nodes grows proportionally to $\log N$, where N is the numbers of nodes in the network. Namely,

$$E(L) = \Theta(\log N),$$

which grows slowly as $N \to \infty$ and the average distance of nodes is small.

A certain category of small-world networks were identified as a class of random graphs by D. Watts and S. Strogatz in (1998). They measured that in fact many real-world networks have a small average distance, but also a clustering coefficient significantly higher than expected by random chance. They noted that graphs could be classified according to two independent structural features, namely the clustering coefficient, and average distance. Purely random graphs, built according to the Erdős-Rényi (ER) model, exhibit a small average distance (typically as $\Theta(\log N)$) along with a small clustering coefficient (typically d/N where d is the expected value of degrees).

Many biological,technological and social networks lie between completely regular and completely random. Typically, these networks have many vertices that are sparse in sense that the average degrees are much less than the number of vertices.

Watts and Strogatz modelled the small-world networks by starting at a graph C_n^k with

$$n \geq k \geq \log n \gg 1,$$

where $k \gg \log n$ guarantees that a random graph will be connected. Then, they choose vertices in order and edges adjacent to the chosen vertices and reconnect these edges to vertices chosen uniformly at random. In the process, the average clustering coefficient decreases slowly and the average distance decreases rapidly and thus they obtained a network between regular ring lattice C_n^k and completely random network. The obtained network has a large average clustering coefficient and a small average distance, which is called small-world network.

5.5 Power law and scale-free networks

Let X be a discrete random variables taking positive integers. If

$$\Pr(X = k) = \frac{c}{k^\gamma},$$

where $c, \gamma > 0$ are constants, then X is said to have a *power law* distribution. This distribution is also call Pareto distribution as economist Pareto originally used it to describe the allocation of wealth that a larger portion of the wealth of society is owned by a smaller percentage of the people (so called 80-20 rule). Contrast to exponential distribution that decreases rapidly, power law is also called heavy-tailed distribution.

Let $P(k)$ be the fraction P(k) of nodes in the network G that have degree k, namely

$$P(k) = \frac{|\{v : \deg(v) = k\}|}{N},$$

where N is number of nodes in G. If $P(k)$ is equal (or close) to power law, then the network G is said to be *scale-free*. For networks, it is typical $2 < \gamma \leq 3$.

The networks of citations between scientific papers are interesting. In 1965, D. Price (1965) found that vertices of degree k in such networks had a heavy-tailed distribution following a power law. He did not however use the term "scale-free network", which was not coined until some decades later. In 1976, Price also proposed a mechanism to explain the occurrence of power laws in citation networks, which he called *cumulative advantage* but which is today more commonly known under the name *preferential attachment*.

In 1999, A. Barabási and colleagues coined the term *scale-free network* when they found some nodes had a much bigger degrees than the that "expected" in random network, and they were surprised and used the term "scale-free", which now is used to describe the class of networks that exhibit a power-law degree distribution.

Albert L. Barabási is a physicist, best known for his work in the research of network theory, and Réka Albert, the co-author of paper (2009), is a professor of physics and biology.

Our earlier study (1999), Albert, Jeong and Barabási found that the World Wide Web is not a random network, but the number of links per node, often called the *degree distribution*, follows a power law. Subsequently, researchers found that not only the WWW, but many other networks, follow the same distribution. These different datasets together indicated that we are dealing with a potentially universal behavior, which might have a common explanation.

Barabási and Albert (2009) proposed a generative mechanism to explain the appearance of power-law distributions, which they called "preferential attachment" and which is essentially the same as that proposed by Price. Analytic solutions for this mechanism (also similar to the solution of Price) were presented earlier by Dorogovtsev, Mendes and Samukhin (2002). Finally, it was rigorously proved by mathematicians Bollobás, Riordan, Spencer and Tusnády (2001).

To explain this phenomenon, Barabási and Albert (2009) suggested the following random graph process as a model, called BA model.

Consider a random graph process in which vertices are added to the graph one at a time and joined to a fixed number of earlier vertices, selected with probabilities proportional to their degrees. Let $v_1, v_2, \ldots$ be a sequence of vertices. Assume that $m_0 \geq 2$ is the number of vertices to start at the process, and let $d(v_i)$ be the degree for the early vertex v_i in the existing graph.

They described the process to start with a small number m_0 of vertices, at every time step we add a new vertex with $m \leq m_0$ edges that link the new vertex to m different vertices already present in the system. If the new vertex is v_{t+1}, then the probability that v_{t+1} is adjacent to v_i is proportional to

$$\frac{d(v_i)}{\sum_{j=1}^{t} d(v_j)}.$$

The above probability signifies the new vertex to incorporate preferential attachment. Note that, to be clear to start, there should exist at least one edge in the first m_0 vertices. A question is if we connect each early vertex and new vertex randomly by above probability, then the excepted number of new edges is one.

The research in Barabási and Albert (2009) is empirical, and the proof is heuristic. The process defined in Bollobás et. al. (2001) preserves the idea of preferential attachment, and the description is much more complex, and power law has been shown for degrees at most $N^{1/15}$ with $\gamma = 3$.

On a theoretical level, some other abstract definitions of scale-free have been proposed. For example, Li et. al. (2005) offered a potentially more precise "scale-free metric". Let $G = (V, E)$ be a simple graph and $s(G) = \sum_{uv \in E} d(u)d(v)$ and $S(G) = s(G)/s_{\max}$, where $s_{\max}$ as the maximum value of $s(H)$ among simple graphs H on same vertex set V with degree distribution identical to G. The notation $S(G)$ gives a metric between 0 and 1, where a G with small $S(G)$ is "scale-rich", and G with $S(G)$ close to 1 is "scale-free". Note that $s(G)$ is maximized when high-degree nodes are connected to other high-degree nodes and $S(G)$ captures the notion of self-similarity implied in the name "scale-free".

Some properties are often listed as the characteristics of scale-free networks, which are as follows.

- Power-law degree distribution;
- Generated by certain random processes with preferential attachment;
- Highly connected hubs that hold the network together with the "robust yet fragile" feature of error tolerance, which is robust when attacked by removing some nodes randomly and fragile by removing some hubs deliberately;

- Generic in the sense of being preserved under random degree-preserving rewiring;
- Self-similar;
- Universal in the sense of not depending on domain-specific details.

5.6 Network Structure

As pointed by Newman (2003), the research on networks may provide new insight into the study of complex systems. Networks have many notable properties, such as the small-world property, the scale-free property, the community structure property, and the links between two objects usually display diversity.

By collecting data from mobile phones, Fagle, Pentland and Lazer (2009) found that the data have the potential to provide insight into the relational dynamics of individuals, and allow the prediction of individual-level outcomes such as job satisfaction.

The concept of contagion has expanded from its original grounding in epidemic disease to describe many processes that spread across networks such as fads, political opinions, the adoption of new technologies, and financial decisions, see, e.g. R. Pastor-Satorras and A. Vespignani (2001) and M. Newman, D. Watts and S. Strogatz (2002).

In traditional models of social contagion, the probability that an individual is affected by the contagion grows by monotonically with the size of neighborhood. By analyzing the growth of Facebook, Ugander, Backstrom, Marlow and Kleinberg (2012) find that the probability of contagion is tightly controlled by the number of connected components in an individual neighborhood, rather than by the actual size of the neighborhood.

A crucial task in the analysis of on-line social-networking systems is to identify important people liked by strong social ties. Drawing data from e-mail, Kossinets and Watts has developed a method of analyzing and estimating tie strength in on-line domains (2006), in which the key structure is *embeddedness*–the number $|N(u) \cap N(v)|$ of mutual friends of two people u and v, a quantity that typically increases with tie strength.

The embeddedness is not necessarily to be the most appropriate for characterizing particular types of strength ties. Backstrom and Kleinberg (2014) proposed a networks-based characterization for intimate relationships, those involving spouses or romantic partners. Using data from a large sample of Facebook users, they try to recognize these people with high accuracy. They found that embeddedness is in fact a comparatively weak means of characterizing romantic relations, and that an alternate network measure that they term *dispersion* is significantly more effective. Roughly, a link between two people has high dispersion when their mutual friends are not well connected to one another. Their research has an important contingent nature: given that a user has declared a relationship partner, they want to understand how effectively they can find partner.

Note that the links to a person's relationship partner or other closest friends may have lower embeddedness, but they often involve mutual neighbors from several foci, reflecting the fact that the social orbits of these close friends are not bounded within any one focus–consider, for example, a husband who knows several of his wife's co-workers, family members, and former classmates, even though these people being to different foci and do not know each other. Thus, Backstrom and Kleinberg proposed some definition as follows.

For a network $G = (V, E)$ and a pair of nodes u and v, denote by $C_{uv} = N(u) \cap N(v)$, the set of mutual friends of u and v, and $c_{uv} = |C_{uv}|$. Let $d(s, t, G)$ be the graph-theoretic distance between u and v in G. For distinct s and t in C_{uv}, define

$$d_{uv}(s,t) = \begin{cases} 1, & d(s,t,C_{uv}) \geq 3, \\ 0, & d(s,t,C_{uv} \leq 2. \end{cases}$$

Then, define the *absolute dispersion* of u and v as

$$disp(u,v) = \sum_{s,t \in C_{uv},\ s \neq t} d_{uv}(s,t).$$

Note that $disp(u, v)$ depends on both of C_{uv} and $d(s, t, C_{uv})$, and define

$$norm(u,v) = \frac{disp(u,v)}{c_{uv}},$$

which is called *normalized dispersion*. Predicting u's partner to be the individual v maximizing $norm(u, v)$ gives the correct answer in 48.0% of all instances.

There are two ways to strengthen normalized dispersion that lead to increased performance. The first is to rank pair of u and v by a function of the form

$$\frac{(disp(u,v)+b)^{\alpha}}{(c_{uv}+c)}.$$

Searching over choices α, b and c leads to maximum performance of 50.5% at

$$\alpha = 0.61, \quad b = 0, \quad c = 5.$$

The second way is by applying the idea of dispersion recursively. For a fixed node u, define first $x_v =$ for all neighbors v of u. Then, iteratively update each x_v to be

$$\frac{\sum_{w\in C_{uv}} x_w^2 + 2\sum_{s,t\in C_{uv}} d_{uv}(s,t)x_s x_t}{c_{uv}} \to x_v.$$

Note that after the first iteration, $x_v = 1 + 2 \cdot norm(u, v)$, and hence ranking nodes by x_v after the first iteration is equivalently to ranking nodes by $norm(u, v)$. Backstrom and Kleinberg found that the highest performance ranking nodes by values of x_v after the third iteration, call such x_v as *recursive dispersion*. The performance by embeddedness and recursive dispersion for romantic relationships is 24.7% and 50.6%, respectively; and that for (married) spouses is 32.1% and 60.7%, respectively.

5.7 References

R. Albert, H. Jeong and A.L. Barabási, Internet-diameter of the World-Wide Web, *Nature*, 401 (6749) (1999),130-131.

L. Backstrom and J. Kleinberg, Romantic partnerships and the dispersion of socialties: A network analysis of relation status on Facebook, Proc. 17th ACM conference on computer supported cooperative work and social computing, 2014.

A. Barabási and R. Albert, Emergence of scaling in random networks, *Science*, 286 (5439) (2009), 509-512.

B. Bollobás, O. Riordan, J. Spencer and G. Tusnády, The degree sequence of a scale-free random graph process, *Random Struct. Algor.*, 18 (2001), 279-290.

S. Dorogovtsev, J. Mendes, Evolution of networks, *Advances in Physics*, 51 (4) (2002),1079.

N. Fagle, A. Pentland and D. Lazer, Infering friendship network structure by using mobile phone data, *Proc. Natl. Acad. Sci. USA*, 106 (36) (2009), 15274-15278.

G. Kossinets and D. Watts, Empirical analysis of an evolving social network, *Science*, 311 (2006), 88-90.

L. Li, D. Alderson, J. Doyle and W. Willinger, Towards a theory of Scale-free graphs: Definitions, properties and implications, *Internet Math.*, 2 (4) (2005), 431-523.

S. Milgram, The small world problems, *Psychology Today*, 2 (1967), 60-67.

M. Newman, The structure and function of complete networks, *SIAM Review*, 45 (2003), 167-256.

M. Newman, D. Watts and S. Strogatz, Random graph model for social networks, *Proc. Natl. Acad. Sci. USA*, 99 (Suppl 1) (2002), 2566-2572.

R. Pastor-Satorras and A. Vespignani, Epidemic spreading in scale-free networks, *Phys. Rew. Lett.*, 86 (2001), 3200-3203.

D. Price, Networks of scientific papers, *Science*, 149 (3683) (1965), 510-515.

J. Ugander, L. Backstrom, C. Marlow and J. Kleinberg, Structural diversity in social contagion, *Proc. Natl. Acad. Sci. USA*, 109 (16) (2012), 5962-5966.

D. Watts and S. Strogatz, Collective dynamics of ‘small-world’ networks, *Nature*, 393 (6684) (1998), 440-442.