LTCC Course: Graph Theory 2022/23
§4 Probabilistic Methods and Random Graphs

Jan van den Heuvel and Yani Pehova

28 November 2022

References

The recommended book for this lecture's material is "The probabilistic method" by Alon and Spencer. Other classic texts on random graphs are:

- B. Bollobás, *Random Graphs*, Cambridge University Press, 2nd Edition (2001).
- S. Janson, T. Łuczak and A. Ruciński, *Random Graphs*, Wiley (2000).

Random Methods

A classical result from Ramsey theory says the following

Lemma 1. *Every 2-edge colouring of K_{4^n} contains a monochromatic clique on n vertices.*

This can be proved by induction on n and in fact we will do it next week. Alternatively, we can say that every 2-edge colouring of K_n contains a monochromatic clique on $0.5 \log n$ vertices (actually, $\log_4 n$ is a bit more than this). Even more alternatively, we can say that every n vertex graph contains a clique or an independent set of size (at least) $0.5 \log n$.

Having proved this, as mathematicians we immediately ask: is this best possible? Can we construct a graph on n vertices which contains no, say, clique or independent set of size $\log n$? $10 \log n$? $1000 \log n$?

In this case, one of the two required properties demands that the graph has rather few edges, and the other demands rather many. It's hard to strike a balance, and any attempt to base a construction around some nice structure seems doomed to failure[1].

In this lecture we see that by far the best way to solve such construction problems is not to give an explicit construction at all, but instead to "construct" the graph "at random", and show that, with positive probability, the random graph constructed has the required combination of properties.

To illustrate this idea let us show that Lemma 1 is tight up to the constant. Mind that in Ramsey theory we very much care about this constant, as well as lower order terms. The problem of determining the Ramsey number of the clique is notoriously difficult.

Lemma 2. *There exists an n-vertex graph G with no clique or independent set of size $2 \log_2 n$.*

[1] In fact there are clever constructions of 'pseudorandom' graphs which solve the last two questions, but these usually use deep theorems from algebraic geometry or number theory in their proofs — and they are still not good enough!

Proof. We build G as a random graph on vertex set $V = [n] = \{1, \ldots, n\}$. For each pair of vertices, independently, we toss a fair coin, and put an edge between the pair if we get a head.

Let's calculate the probability that a particular set C of c vertices forms a clique. This means that each of the $\binom{c}{2}$ coin-tosses corresponding to the pairs in the clique was a head, an event with probability $2^{-\binom{c}{2}}$.

For each set $C \subseteq V$ of size c, let Y_C be the indicator random variable taking value 1 if C is a clique and 0 otherwise. Then $\mathbb{P}(Y_C = 1) = 2^{-\binom{c}{2}}$.

Now, the *expected* number of cliques of size c is given by:

$$\mathbb{E} \sum_C Y_C = \sum_C \mathbb{E} Y_C = \sum_C \mathbb{P}(Y_C = 1) = \binom{n}{c} 2^{-\binom{c}{2}} \leq \Big(\frac{ne}{c 2^{(c-1)/2}} \Big)^c$$

using the inequality $\binom{n}{k} \leq (en/k)^k$. If $c \geq 2 \log_2 n$, then $n \leq 2^{c/2}$, and so the expected number of c-cliques is at most $(e\sqrt{2}/c)^c$. For $c \geq 5$, this expectation is at most $1/3$. This implies that the probability that the random graph has a clique of size c is at most $1/3$. The same calculation shows that the probability that the graph has an independent set of size c is also at most $1/3$.

So, with probability at least $1/3$, the random graph on $n \geq 5$ vertices has neither a clique nor an independent set of size as large as $2 \log_2 n$. $\square$

> *The idea: We can calculate expectation using 'linearity of expectation' and estimate probabilities from expectations.*

This delightfully simple argument was first given by Paul Erdős in 1947, and it is one of the first instances of the successful use of the *probabilistic method* in combinatorics. One can, and indeed Erdős did, recast this entire proof as a *counting argument*. Of the $2^{\binom{n}{2}}$ graphs with vertex set $[n]$, the number of them in the class A_C with a clique or independent set on the set C of size c is $\ldots$, and then effectively the same calculation shows that there are some graphs in none of the sets A_C.

We will talk more about *Ramsey numbers* next week, but for now, what we just showed is that the *Ramsey number* $R(c, c)$ is at least $\frac{c}{e} 2^{(c-1)/2}(1 - o(1))$. (Here the $o(1)$ refers to a term that tends to zero as $c \to \infty$.)

Some probabilistic inequalities

So far we only used very simple probabilistic tools:

- the linearity of expectation,
- the fact that if the expectation of a (non-negative) random variable is smaller than one, then it takes the value zero with non-zero probability.

Before we can turn to more difficult results, we need to introduce some important probabilistic tools. Note that in these notes, and usually in probabilistic combinatorics in general, we will work with a finite probability space. This means that the 'usual' difficulties of measure theory disappear; we can assume all sets are measurable, we can replace integrals with finite sums and hence there is no question of convergence. If you aren't familiar with those terms: don't worry, the last two sentences just say you don't need to be.

Lemma 3 (Markov's inequality)**.** *Let X be any non-negative real-valued random variable. For each $a > 0$, we have $\mathbb{P}(X \geq a) \leq \frac{\mathbb{E}X}{a}$.*

Proof. By definition, we have

$$\mathbb{E}X = \sum_{i \in \mathbb{R}_0} i\mathbb{P}[X = i] \geq a \cdot \mathbb{P}[X \geq a]$$

which since $a > 0$ completes the proof. $\square$

Note that in the above proof, the 'uncountably infinite' sum is simply a technical convenience. Since X is a random variable on a finite probability space, X takes one of only finitely many values, and the sum is really a finite sum.

Lemma 4 (Chebyshev's inequality)**.** *Let X be any real-valued random variable. For each $k > 1$, we have*

$$\mathbb{P}\Big(|X - \mathbb{E}X|^2 \geq k^2\big(\mathbb{E}X^2 - (\mathbb{E}X)^2\big)\Big) \leq \tfrac{1}{k^2}\,.$$

Proof. Observe that, letting $Z = \big(X - \mathbb{E}X\big)^2$, we have $\mathbb{E}Z = \mathbb{E}X^2 - \big(\mathbb{E}X\big)^2$, and Z is a non-negative random variable. Thus the desired statement is precisely Markov's inequality applied to Z. $\square$

For the third lemma we suppose that $X_1, \dots, X_n$ are independent Bernoulli random variables, with $\mathbb{P}(X_i = 1) = p_i$ (we do not need that the p_i are identical: but if they are we say X is a Binomial random variable). Let $X = \sum_i X_i$, so that by linearity of expectation we have $\mathbb{E}X = \sum_i p_i$.

Lemma 5 (a Chernoff bound)**.** *If X is the sum of independent Bernoulli random variables, then for each $0 < \delta < 3/2$ we have*

$$\mathbb{P}\big(|X - \mathbb{E}X| \geq \delta\mathbb{E}X\big) < 2e^{-\frac{\delta^2\mathbb{E}X}{3}}\,.$$

The proof of this is Exercise 1(a).

In all cases, the idea is supposed to be that a random variable is 'usually' 'not too far' from its mean. Markov's inequality provides only weak bounds on what 'usually' and 'not too far' should be, but in return we have only to compute expectation. Chebyshev's inequality often tells us much more, but we have to calculate not only the expectation of X but also of X^2, which can be tricky. Finally, Chernoff bounds (many different forms exist; the one quoted is often good) are very strong and easy to use, but only apply in the special case of a sum of independent Bernoulli random variables.

A very important fact, which we won't use in these lectures but which is central to a lot of modern research, is that the last sentence is not really true. The Chernoff bound also applies to Bernoulli random variables which are *not* independent, but where the dependencies are 'sequential'. For example, if we have $\mathbb{P}\big(X_i = 1 \big| X_1, \dots, X_{i-1}\big) \leq p_i$ for each i, then letting $S = \sum_i p_i$ we can conclude

$$\mathbb{P}\big(X \geq S + \delta S\big) < e^{-\frac{\delta^2 S}{3}}\,.$$

This can be proved using *coupling*. More excitingly, we can allow p_i to be not a constant but depend on the outcomes $X_1, \dots, X_{i-1}$. This means that $S = \sum_i p_i$ is no longer a constant; it too is a random variable (the 'observed expectation'). Suppose that we have $S \leq s$ almost surely. Then we can still conclude

$$\mathbb{P}\big(X \geq (1+\delta)s\big) < e^{-\frac{\delta^2 s}{3}}\,. \tag{1}$$

This is a *martingale concentration inequality* in disguise; proving it is Exercise 1(b). If you want to analyse a probabilistic process, you will likely need such things. A simple to state example is the *triangle free process.* This starts with the graph G_0 on n vertices with no edges. At each integer time $t \geq 1$, a pair (u, v) of vertices is chosen uniformly at random from the set of all pairs at distance three or more in G_{t-1} and the edge uv is added to form G_t. In other words, the constraint is that at each time we add an edge, we really add an edge (uv was not already present) and that edge does not create a triangle. The process stops when no such edges remain.

This process is easy to define, but hard to analyse (it was done in two independent papers, by Fiz Pontiveros, Griffiths and Morris, and by Bohman and Keevash, in 2013; both papers are quite long and technical). It turns out that the process is likely to run for about $\frac{1}{2\sqrt{2}} n^{3/2}\sqrt{\log n}$ steps, and the final graph is likely to have very large independent sets (as triangle-free graphs on n vertices go); this gives the best lower bound we know on the Ramsey number $R(3, k)$—we'll mention this again next week.

Girth and Chromatic Number

The *girth* of a graph G is the length of the shortest cycle in G. So a graph with high girth "locally" looks like a tree, hence has a rather simple structure. More precisely, suppose the girth is at least g, and we take $2k < g$ and let $N_k(v)$ be the number of vertices at distance at most k from vertex v. Then all the sets $N_k(v)$ for $v \in V(G)$ induce trees. In particular, all the induced subgraphs on the sets $N_k(v)$ are 2-colourable. Is it nevertheless possible for the whole graph to have a large chromatic number? In this section we show how to use Markov's inequality, also known as *the first moment method*, to show that the answer is yes, and in fact there exist graphs with arbitrarily high girth and chromatic number.

How can we prove that the chromatic number of a graph is large? One way is to show the graph has no large independent sets, as the set of vertices receiving any given colour is an independent set. Hence we will prove the following result.

Theorem 6. *For each $g, k \in \mathbb{N}$ and all sufficiently large n, there exists a graph on n vertices with girth greater than g and chromatic number at least k.*

The proof of this result (also due to Erdős, this time from 1959) is slightly more complicated than the one about Ramsey numbers. For a start, our random process involves the tossing of *biased* coins. Again we fix a vertex set $[n]$, but now we put an edge between each pair of vertices with probability p, all choices made independently. Here $p = p(n)$ is a function of n that we can choose to suit our needs. What this defines is the standard (binomial) model $\mathcal{G}(n, p)$ of *random graphs*. We use $G(n, p)$ to denote a random graph chosen according to this method.

So our goal is to find some p such that (with positive probability, for sufficiently large n) the random graph $G(n, p)$ has no cycle of length at most g and no independent set of size at least n/k (recall that a proper vertex colouring in k colours is a partition of the vertex set into k independent sets, one of which must have size at least n/k). Unfortunately there is no such p: if $p = 100/n$, it is very likely that $G(n, p)$ contains triangles (see the next section for more detail on this), and it is very likely that n/k vertices of $G(n, p)$ are isolated (and in particular form an independent set) for some (large) k. So if $p \leq 100/n$ we are likely to have too many isolated vertices, and if $p > 100/n$ then we are likely to have triangles in $G(n, p)$. However, the idea then will be to choose the parameter p in such a way that $G(n, p)$ 'almost' has the two properties we want so that a few vertex deletions will result in a graph with these properties.

Proof. Let X be the number of cycles of length at most g in $G(n, p)$. We have (again, by linearity of expectation) that

$$\mathbb{E}X = \sum_{i=1}^{g} \binom{n}{i} \frac{i!}{2i} p^i \leq \sum_{i} \frac{n^i}{2i} p^i \leq \frac{g}{2}(np)^g,$$

as long as $p \geq 1/n$. By Markov's inequality,

$$\Pr(X \geq n/2) \leq \frac{\mathbb{E}X}{n/2} \leq g n^{g-1} p^g.$$

This probability is at most 1/3 if we take $p = (1/3g)^{1/g} n^{1/g-1}$. Note that for such p, $np > 1$ for large n.

For ease of notation, let's assume that $2k$ divides n. Let Y be the number of independent sets of size $n/2k$. We did this calculation already, the expectation of Y is

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

Using the inequality $1 - p \leq e^{-p}$ and again the bound on the binomial coefficient $\binom{n}{k} \leq (en/k)^k$, we can show that this probability tends to 0 as $n \to \infty$. So, for large n, $\Pr(Y \geq 1) \leq \mathbb{E}Y \leq 1/3$ (this is a rather trivial application of Markov's inequality).

Since both $\Pr(Y \geq 1)$ and $\Pr(X \geq n/2)$ are at most 1/3, with positive probability $G(n,p)$ will satisfy $Y = 0$ and $X < n/2$. Fix one graph H (sampled from $G(n,p)$) which satisfies both these properties.

Finally, delete one vertex from each short cycle, removing at most $n/2$ vertices in total. The resulting graph has at least $n/2$ vertices, girth greater than g and no independent sets of size $n/2k$, meaning that its chromatic number is at least $(n/2)/(n/2k) = k$. □

The idea: For the p we chose, $G(n,p)$ is 'very far' from having independent sets of size n/k, and we can use the Chernoff bound which gives us very strong probability bounds to check this. On the other hand, even though $G(n,p)$ contains a few short cycles, it doesn't have many. We can use Markov's inequality to say that with a reasonable probability, $G(n,p)$ has 'not too many more than expectation' short cycles, and we can delete them.

Triangles in $G(n,p)$

The structure and properties of $G(n,p)$ in various regimes of p are of interest not only because they give us neat constructions for deterministic problems. The random graph $G(n,p)$ is a central model used in network theory.

In this section we see how we can apply Chebyshev's inequality to show that $p(n) = 1/n$ is a *threshold* for $G(n,p)$ to contain a triangle: if $p(n) = o(1/n)$, then $G(n,p)$ has no triangles, but if $p(n) = \omega(1/n)$, then $G(n,p)$ contains a triangle. Both behaviours happen *asymptotically almost surely* (a.a.s), meaning that as $n \to \infty$, the probability that the behaviour doesn't occur tends to 1.

Theorem 7. *$1/n$ is a threshold for $G(n,p)$ to contain a triangle.*

Proof. Let X be the number of triangles in G. We start by calculating the mean and variance of X. We easily have $\mathbb{E}X = \binom{n}{3}p^3 =: \mu$. Calculating the variance, however, is more work. Let

$X = \sum_S X_S$ be the sum of indicators over all 3-vertex sets S. As we have done previously, $X_S = 1$ if S induces a triangle. Then

$$\mathbb{E}(X^2) = \mathbb{E}\left(\sum_S X_S\right)^2 = \sum_{S,S'} \mathbb{E} X_S X_{S'}.$$

Note that for the diagonal terms we have $X_S^2 = X_S$ as each X_S takes values in $\{0,1\}$. So, we have that $\mathrm{Var}(X) = \sum_{S,S'} (\mathbb{E} X_S X_{S'} - \Pr X_S \Pr X_{S'})$.

For S, S' that share at most 1 vertex (i.e., no edges), we have $\mathbb{E} X_S X_{S'} = \Pr X_S \Pr X_{S'} = p^6$. For S, S' sharing 2 vertices (i.e. one edge), we have $\mathbb{E} X_S X_{S'} = p^5$ and $\Pr X_S \Pr X_{S'} = p^6$. For S, S' sharing 3 vertices (the diagonal terms), we have $\mathbb{E} X_S X_{S'} = p^3$ and $\Pr X_S \Pr X_{S'} = p^6$. This adds up to

$$\mathrm{Var} X = \binom{n}{3}(p^3 - p^6) + \binom{n}{3}(n-3)3(p^5 - p^6).$$

Suppose that $p = o(1/n)$. Then $\mu \leq n^3 p^3 \to 0$ as $n \to \infty$, so the graph is a.a.s. triangle-free. (Here we use again that $\Pr(X \geq 1) \leq \mu$.)

Suppose that $p = \omega(1/n)$. Then $\Pr(X = 0) \leq \Pr(|X - \mu| \geq \mu) \leq \mathrm{Var} X/\mu^2$ by Chebyshev's inequality. We have $\mathrm{Var} X \leq n^3 p^3 + n^4 p^5$ and $\mu^2 \geq n^6 p^6/40$. Thus

$$\Pr(X = 0) \leq \frac{40}{n^3 p^3} + \frac{40}{n^2 p}.$$

Both denominators tend to infinity with n, by choice of p, so $G(n,p)$ contains at least one triangle with probability tending to 1. $\square$

Designs

Finally, we want to turn to a major classical problem in combinatorics, which was only solved very recently. Consider the following task: Given a complete graph on n vertices, partition the edges into as many edge-disjoint triangles as possible plus a left-over set. Is it possible to do this in such a way that the left-over set is empty?

A design theorist will recognise this as asking for a *Steiner triple system* on n elements. These are known to exist if and only if n is congruent to 1 or 3 modulo 6 (to see that this is necessary, observe that all vertices have to have even degree, so n must be odd, and the number of edges has to be divisible by 3, which rules out 5 modulo 6).

But what happens if we ask for edge-disjoint copies of K_k instead of K_3 (and replaces "1 or 3 modulo 6" by appropriate other obvious divisibility conditions)? What happens if instead of partitioning the edges of a complete graph K_n we want to partition the edges of a complete t-uniform hypergraph $K_n^{(t)}$ into copies of $K_k^{(t)}$? These are all still good questions in design theory, but they are very hard. This problem dates back to Steiner in 1853, and it was a celebrated breakthrough when Keevash announced a solution at the beginning of 2014.

His solution, in fact, makes essential use of an *approximate* solution to the problem which was already given in 1985 by Rödl: He proved that the left-over set can be made an arbitrarily small fraction of the edge set (we do not need divisibility conditions in this case). To prove this approximate result Rödl used a (more advanced) probabilistic argument, which we shall discuss in the following.

Theorem 8 (Rödl, 1985)**.** *For any $1 \leq t \leq k \in \mathbb{N}$ and $\gamma > 0$, if n is large enough, there is a partition of the edges of $K_n^{(t)}$ into edge-disjoint copies of $K_k^{(t)}$ and a leftover set of size at most γn^t.*

To simplify the argument we will only consider the graph case for K_3-partitions here, that is, the case $t = 2, k = 3, n > 3$. Note that the $t = 1$ case is trivial, as is the $t = k$ case, as is the $k = n$ case - so this is the smallest non-trivial case, but actually the general case is not much harder. So what we want to prove is that

- there is a set of $(1-\gamma)n^2/6$ edge-disjoint triangles in K_n.

Let us try a first approach at using the probabilistic method for this[2]: We have to somehow construct the partition 'randomly', and the obvious way is to select triangles from K_n 'randomly' in some way. So suppose we selected triangles from K_n independently at random with some probability p. Then the expected number of triangles we select is $p\binom{n}{3}$, so we should choose p to be at least $(1-\gamma)/n$ in order for the expected number of triangles to be at least $(1-\gamma)n^2/6$. But it is easy to believe that these triangles will not be edge-disjoint: given any one triangle xyz, the expected number of (other) selected triangles containing xy is $p(n-3)$, so the expected number of selected triangles which share an edge with xyz is $3p(n-3)$, which is (at least) close to 3.

Maybe we can use a 'trick' like the one we used in proving Theorem 6, and delete some triangles in order to get a collection of edge-disjoint triangles? The only simple rule which will give us a collection of edge-disjoint triangles is: delete every triangle which shares an edge with another triangle. We can estimate how many selected triangles we have to delete by the number Y of pairs of selected triangles which share an edge. Hence $\mathbb{E}Y = p^2\binom{n}{3} \cdot n \le p^2n^4$, and by Markov's inequality we conclude that (with probability at least $9/10$) we only have to delete $10p^2n^4$ triangles. So we can find a collection of at least

$$p\binom{n}{3} - 10p^2n^4$$

edge-disjoint triangles. But unfortunately this is smaller than 0 if $p \ge (1-\gamma)/n$. Even if we take $p = \varepsilon/n$ for some $\varepsilon > 0$, this gives only approximately $(\varepsilon - 60\varepsilon^2)\frac{n^2}{6}$ edge-disjoint triangles. This of course is (well) below $n^2/6$, so we cannot hope to even get close to the desired number of edge-disjoint triangles. But, if ε is very small then we will delete only a tiny proportion of the triangles we selected.

Rödl's clever insight now was the following. If we perform this select-and-delete method with $p = \varepsilon/n$ for some very small $\varepsilon > 0$, then we will get a (small) collection of edge-disjoint triangles T_1, a (tiny) collection of 'waste' edges contained in the deleted triangles, and a (large) set of 'leftover' edges L_1. But this 'leftover' set will be very well-behaved - it will have the property that 'most' edges are contained in about the same number of triangles in L_1. We can repeat select-and-delete using only the edges of the leftover set to get another collection of edge-disjoint triangles T_2 and another leftover set L_2. By construction $T_1 \cup T_2$ is a collection of edge-disjoint triangles, and we can show that L_2 is still well-behaved. We can keep taking these 'Rödl Nibbles' until we get down to a tiny leftover set and the desired collection of edge-disjoint triangles. That is the basic idea of the proof that follows.

Note that to justify that this procedure works, there are two things we need to show. First, we need to show that the leftover set is really always nice enough to run the 'nibble' procedure (and to analyse it!). Second, we need to show that the total number of 'waste' edges does not get large. If we show both of these, then when we stop we have by construction a collection of edge-disjoint triangles, and the edges not in that collection are either 'leftover' or 'waste'; both sets are small and therefore most edges are in the triangles, as we want. These two properties are established by the following lemma.

[2] which will fail; then we will refine the method, which will also fail; and so on

Lemma 9. *Given $\gamma > 0$ and $\varepsilon > 0$, for all sufficiently small $\delta' > 0$ there exists $\delta > 0$ such that for all sufficiently large n the following holds. Suppose $D \geq \gamma^2 n/32$. Let G be a graph on n vertices in which all but at most $\delta e(G)$ edges are in $(1 \pm \delta)D$ triangles. Let S be a set of triangles selected independently with probability $p = \varepsilon/D$ from the triangles of G. Let T be the set of triangles in S which share no edge with any other triangle of S. Let G' be obtained from G by deleting any edge contained in a triangle of S. Then with probability at least $1/2$ we have*

(*a*) $|S| = (1 \pm \delta')\varepsilon e(G)/3$.

(*b*) $|S \setminus T| < 60\varepsilon|S|$.

(*c*) *All but at most $\delta' e(G)$ edges of G' lie in $(1 \pm \delta')(1-\varepsilon)^2 D$ triangles of G'.*

We will only prove the following simplified version of this lemma. The difference is that the assumption on G that most of its edges are in about D triangles in Lemma 9 is replaced by the stronger assumption that all edges are in exactly D edges. (In 'real' we need the weaker assumption, because this is what (*c*) guarantees and we want to apply the lemma successively). Lemma 10 is not conceptually easier to prove than Lemma 9. But the stronger assumption makes the calculations substantially shorter.

Lemma 10. *Given $\gamma > 0$ and $\varepsilon > 0$, for all sufficiently small $\delta > 0$, if n is large enough, the following holds. Suppose $D \geq \gamma^2 n/32$. Let G be a graph on n vertices in which every edge is in D triangles. Let S be a set of triangles selected independently with probability $p = \varepsilon/D$ from the triangles of G. Let T be the set of triangles in S which share no edge with any other triangle of S. Let G' be obtained from G by deleting any edge contained in a triangle of S. Then with probability at least $1/2$ we have*

(*a*) $|S| = (1 \pm \delta)\varepsilon e(G)/3$.

(*b*) $|S \setminus T| < 30\varepsilon|S|$.

(*c*) *All but at most $\delta e(G)$ edges of G' lie in $(1 \pm \delta)(1-\varepsilon)^2 D$ triangles of G'.*

Before we prove Lemma 10 let us briefly (and sketchily) discuss how Lemma 9 implies Theorem 8. First, the bounds on $|S|$ and $|S \setminus T|$ yield bounds on $e(G')$: we find that $e(G')$ is very close to $(1-\varepsilon)e(G)$ (the number of edges deleted is at least $3|T|$ and at most $3|S|$). We conclude that after t steps—t nibbles—applying Lemma 9 starting from $G_0 = K_n$ we get to a graph G_t in which the number of edges is close to $(1-\varepsilon)^t n^2/2$, and most edges are in about $(1-\varepsilon)^{2t} n$ triangles. If we choose τ such that $\gamma/2 < (1-\varepsilon)^\tau < \gamma$, then we conclude that $e(G_\tau) < \gamma n^2/2$ and that for each $t \leq \tau$, most edges of G_t are in at least $(1-\varepsilon)^{2t} n/2 > \gamma^2 n/8$ triangles. In particular this means the conditions of Lemma 9 that $e(G)$ and D should not be too small are satisfied up to τ steps, so we are allowed to make τ nibbles.

Second, at each nibble step we partition the edges of G into three parts: the edges of T (which are edge-disjoint triangles), the edges in $S \setminus T$ (which are 'waste'), and the edges of G'. Because $|S \setminus T|$ is much smaller than $|T|$, we conclude that the number of 'waste' edges in each nibble is a tiny fraction of the number of edges covered by T. So after τ nibbles, we have partitioned $E(K_n)$ into a collection of edge-disjoint triangles, a waste set which is tiny by comparison, and $E(G_\tau)$ which we know has size at most $\gamma n^2/2$. Provided we chose ε sensibly, this means the waste set and $E(G')$ together account for only at most γn^2 edges, and we conclude that the collection of edge-disjoint triangles we found covers all but γn^2 edges of $E(K_n)$ as desired.

Proof of Lemma 10. Given $\gamma > 0$ and $\varepsilon > 0$, we choose any sufficiently small $\delta > 0$.

Let G be an n-vertex graph in which every edge is in D triangles. Let S be obtained by choosing independently triangles from G with probability $p = \varepsilon/D$. Let T be the set of triangles in S which share no edge with any other triangle of S, and let G' be the graph obtained from G by removing all edges in triangles of S.

First we will estimate $|S|$. Observe that G has $\frac{D}{3}e(G)$ triangles, and S is obtained by selecting independently from these triangles with probability p. So the expectation of $|S|$ is $\frac{D}{3}pe(G) = \varepsilon e(G)/3$. By the Chernoff bound we conclude that

$$\Pr\Big(\big||S| - \varepsilon e(G)/3\big| > \delta\varepsilon e(G)/3\Big) < 2e^{-\delta^2\varepsilon e(G)/3}$$

which (provided $e(G)$ is large enough—which it is unless G is empty) is smaller than 0.1. We conclude that with probability at least 0.9 the set S has size

$$|S| = (1 \pm \delta)\varepsilon e(G)/3\,,$$

which proves that (a) holds with probability at least 0.9.

Now we can estimate $|S \setminus T|$ by exactly the same method we tried to use above. We estimate how many pairs of triangles we find sharing an edge in S; for each pair we delete two triangles. The number of pairs of triangles sharing an edge in G is just $e(G)\binom{D}{2}$, where we first choose the shared edge e then the pair of triangles. The expected number of these which appear in S is then $p^2\binom{D}{2}e(G) < \varepsilon^2 e(G)/2$. So the expected size of $S \setminus T$ is at most $\varepsilon^2 e(G)$ (since we delete both triangles from each pair). So by Markov's inequality with probability at least 0.9 we have

$$|S \setminus T| \le 10\varepsilon^2 e(G) < 60\varepsilon|S|\,,$$

which proves that (b) holds with probability at least 0.9.

Finally we have to show that most edges of G' are in the 'right' number of triangles of G'. This is the most difficult part of the proof. First, consider a fixed edge uv of G. We condition on the event that uv is also an edge of G': that is, we assume no triangle containing uv is in S. We want to know T_{uv}, the number of triangles containing uv in G'. First we will try to find $\mathbb{E}T_{uv}$: for this we just need to find the probability that a given triangle uvw of G survives to G'. Now there are $D-1$ triangles of G (apart from uvw) using the edge uw, and another (different!) $D-1$ triangles using the edge vw. The triangle uvw survives if and only if none of these triangles are in S (uvw cannot be in S since uv is an edge of G'). The probability of that occurring is just $(1-p)^{2D-2}$. So, since $p = \frac{\varepsilon}{D}$, we have

$$\mathbb{E}T_{uv} = D(1-p)^{2D-2} \approx (1-\varepsilon)^2 D\,.$$

Now we need to show that T_{uv} takes a value close to its expectation (at least most of the time). It would be very nice if we could say that T_{uv} was a sum of D independent Bernoulli random variables (one for each triangle of G containing uv) and use the Chernoff bound, but this isn't true: these random variables are not independent, as it can happen that one triangle being selected affects two of the random variables. So we have to use Chebyshev's inequality, and this means we need to estimate $\mathbb{E}T_{uv}^2$.

Claim 11.

$$\mathbb{E}T_{uv}^2 \le \big(\mathbb{E}T_{uv}\big)^2 + \delta^5 D^2/2\,.$$

We will prove this later. Assuming the claim, we can complete the proof. By Chebyshev's inequality, we have

$$\mathbb{P}\Big(\big|T_{uv} - \mathbb{E}T_{uv}\big|^2 \ge 2\delta^{-1}\big(\mathbb{E}T_{uv}^2 - (\mathbb{E}T_{uv})^2\big)\Big) \le \delta/16\,.$$

Substituting $\mathbb{E}T_{uv}^2$, we get that the probability of $T_{uv} \neq \mathbb{E}T_{uv} \pm \delta^2 D$ is at most $\delta/4$. Since $\delta < 1/4$, and since $D < 2\mathbb{E}T_{uv}$, we conclude that the probability of $T_{uv} \neq (1 \pm \delta/2)\mathbb{E}T_{uv}$ is at most $\delta/4$. Finally, since $\mathbb{E}T_{uv}$ is approximately $(1-\varepsilon)^2 D$, we conclude that the probability of $T_{uv} \neq (1 \pm \delta)(1-\varepsilon^2)D$ is at most $\delta/4$.

Now let Y be the random variable counting edges uv of G' for which $T_{uv} \neq (1 \pm \delta)(1-\varepsilon^2)D$. The expectation of Y is at most $\delta e(G)/4$, so by Markov's inequality with probability at least $3/4$ we have $Y \leq \delta e(G)$. In other words, with probability at least $3/4$, all but at most $\delta e(G)$ edges of G' lie in $(1 \pm \delta)(1-\varepsilon)^2 D$ triangles of G', which is (c).

Concluding, the probability that any one of (a), (b), (c) fails is at most $0.1 + 0.1 + 0.25 < 0.5$, so with probability at least 0.5 all three items hold as desired. It remains to check the Claim.

Proof of Claim 11. Given two events A and B, we define the *covariance* of A and B to be $\mathrm{Cov}(A,B) = \mathbb{E}AB - \mathbb{E}A\mathbb{E}B$. Now if A and B are independent, their covariance is zero. This will help us evaluate $\mathbb{E}T_{uv}^2$.

For each w such that uvw is a triangle of G, let I_w be the random variable which takes value 1 if uvw is a triangle of G', and 0 otherwise. We have the following equality.

$$\mathbb{E}T_{uv}^2 - (\mathbb{E}T_{uv})^2 = \mathbb{E}\Big(\sum_w I_w\Big)^2 - \Big(\mathbb{E}\sum_w I_w\Big)^2 = \mathbb{E}\sum_w I_w^2 + \mathbb{E}\sum_{w\neq w'} I_w I_{w'} - \sum_w (\mathbb{E}I_w)^2 - \sum_{w\neq w'} \mathbb{E}I_w \mathbb{E}I_{w'}\,,$$

and this is bounded above by

$$\mathbb{E}\sum_w I_w^2 + \sum_{w\neq w'} \mathrm{Cov}(I_w, I_{w'}) = \mathbb{E}T_{uv} + \sum_{w\neq w'} \mathrm{Cov}(I_w, I_{w'})\,, \tag{2}$$

where we left out the third sum and rearranged. Our aim is to show that this last quantity is bounded above by $\delta^5 D^2/2$. Now consider some w and w' which make triangles of G with uv. The events I_w and $I_{w'}$ are independent unless there is a triangle of G which uses both w, w' and one of u and v. In other words, $\mathrm{Cov}(I_w, I_{w'}) = 0$ unless uww' and vww' are triangles of G. We get

$$\mathrm{Cov}(I_w, I_{w'}) = (1-\varepsilon/D)^{4D-6} - (1-\varepsilon/D)^{4D-4} \leq 2\varepsilon/D\,.$$

Now for each given w, there are at most $2D$ vertices w' such that one of uww' or vww' is a triangle of G. So (2) can be bounded above by

$$\mathbb{E}T_{uv} + \sum_w 2D(2\varepsilon/D) = \mathbb{E}T_{uv} + 4D < \delta^5 D^2/2\,,$$

where the final inequality is true since $D > \gamma^2 n/16$. □

That proves the Claim, and hence the proof of Lemma 10 is complete. □

1 Exercises

There are plenty more exercises in the textbooks!

Exercise 1. *(a) Prove Lemma 5. Hint: first evaluate* $\mathbb{E}e^{tX}$ *for* $t > 0$ *a fixed constant, apply Markov's inequality, and optimise* t*. Start with the case of independent identical Bernoulli variables (i.e.* $p_1 = p_2 = \dots$*. For the general case, you might want to look up 'Jensen's inequality'.*

(b) Try to generalise this to prove the martingale concentration inequality (1)*. The difficult part is figuring out how to evaluate* $\mathbb{E}e^{tX}$*.*

Exercise 2. *For $k \in \mathbb{N}$, a graph $G = (V, E)$ has* Property S_k *if, for every pair (A, B) of disjoint k-element subsets of V, there is a vertex x of the graph that is adjacent to every vertex of A and no vertex of B.*
(a) Find a graph with property S_1.
(b) Show that, for each $k \in \mathbb{N}$, there is a graph with property S_k.

Exercise 3. *A k-uniform* hypergraph *is a pair $H = (V, E)$, where V is a set of* vertices*, and E is a family of k-element subsets of V. (So a 2-uniform hypergraph is just a graph.) A hypergraph $H = (V, E)$ has* Property B *if V can be partitioned into two subsets V_1 and V_2 in such a way that no edge is entirely contained within one of the two sets.*
(a) Show that, if $H = (V, E)$ is a k-uniform hypergraph with $|E| < 2^{k-1}$, then H has property B.
(b) Show that, if $H = (V, E)$ is a k-uniform hypergraph such that each edge in E intersects at most d others, and $e(d + 1) \leq 2^{k-1}$, then H has property B.

Exercise 4. *(a) Let $p = n^{-t}$, for $0 < t < 1$, and let k be a fixed natural number. Write down an expression for the expected number of k-cliques in $G(n, p)$. Hence show that, if $t > 2/(k - 1)$, the probability that $G(n, p)$ contains a k-clique tends to zero as $n \to \infty$.*

It is also true that, if $t < 2/(k - 1)$, then the probability that $G(n, p)$ contains a k-clique tends to one as $n \to \infty$: to prove this, one needs to work with the variance of the number of k-cliques.
(b) Let H denote the graph on five vertices a, b, c, d, e with seven edges: a, b, c, d form a clique, and de is also an edge. For $p = n^{-7/10}$, find the expected number of copies of H in $G(n, p)$. What is

$$\lim_{n\to\infty} \mathbb{P}(G(n, p) \textit{ contains a copy of } H)?$$

(c) There is a parameter $b(H)$ of graphs such that, if $p = n^{-t}$ and $t > b(H)$, then the probability that $G(n, p)$ contains a copy of H as a subgraph tends to zero, while if $p = n^{-t}$ and $t < b(H)$, then this probability tends to 1. Based on the calculations in this question, what do you think this parameter $b(H)$ might be?

Exercise 5. *Set $p = n^{-2/5}$, and consider a random graph $G = G(n, p)$.*
(a) Show that the degree of any fixed vertex v has a Binomial distribution, and find an upper bound on the probability that this degree is greater than or equal to $n^{2/3}$.
(b) Show that the probability that the maximum degree of G is at most $n^{2/3}$ is at least $2/3$.
(c) Show that, with probability at least $2/3$, for every pair (U, V) of subsets of $V(G)$, with $|U|, |V| \geq n^{1/2}$, there is an edge from U to V.
(d) What can you deduce from (b) and (c)?

Exercise 6. *(a) Try to prove Lemma 9. You should find that the only difficult part is to prove that most edges are in the 'right' number of triangles. You will not be able to prove that T_{uv} behaves nicely for every edge $uv \in G$: you will need to assume both that uv happened to lie in about the 'right' number of triangles in G, and that 'most' of those triangles share edges with about the 'right' number of triangles in G. If you make this assumption, you should be able to modify the argument given to show that T_{uv} is likely to be about the 'right' size. Then you will need to show that there cannot be too many edges of G which don't satisfy the assumption.*
(b) Try to prove the special case of Theorem 8 from Lemma 9. The difficulty here is to find out how to set constants in order to make the argument work.
(c) Try to prove Theorem 8—or try to understand the argument given in e.g. Alon and Spencer!